Contributions to School Readiness Skill Development: A Focus on Non-Academic Skills during the Preschool-Kindergarten Transition

---

A Dissertation Defense

Presented to

The Faculty of the Curry School of Education

University of Virginia

In Partial Fulfillment

of the Requirements for the Degree of

Doctor of Philosophy

---

by

Shannon E. Reilly, M.Ed.

August 2020

Curry Programs in Clinical and School Psychology
Curry School of Education
University of Virginia
Charlottesville, Virginia

APPROVAL OF THE DISSERTATION DEFENSE

This dissertation ("Contributions to School Readiness Skill Development: A Focus on Non-Academic Skills during the Preschool-Kindergarten Transition") has been approved by the Graduate Faculty of the Curry School Education in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

____
Name of Chair (Jason Downer, Ph.D.)

____
Name of Committee Member (Amanda Williford, Ph.D.)

____
Name of Committee Member (Julia Blodgett, Ph.D.)

____
Name of Committee Member (Timothy Konold, Ph.D.)

____
Date

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Contributions to School Readiness Skill Development: A Focus on Non-Academic Skills during the Preschool-Kindergarten Transition

Shannon E. Reilly

Contributions to School Readiness Skill Development: A Focus on Non-Academic Skills during the Preschool-Kindergarten Transition

The transition from preschool to kindergarten is challenging for young children because there are novel and more complex interactions with new adults and peers, shifts in focus toward more academic instruction, and increases in demands and expectations (Rimm-Kaufman & Pianta, 2000). Thus, it is critical to prepare children – particularly those at risk of falling behind academically early on because of socioeconomic vulnerabilities – for success by helping them to develop foundational skills prior to kindergarten entry, particularly because emerging skills are especially malleable to intervention during this sensitive time period (Knudsen, Heckman, Cameron, & Shonkoff, 2006; McClelland, Geldhof, Cameron, & Wanless, 2015). Although it is important to build a strong foundation in all school readiness skills on which more complex abilities can be scaffolded (Masten & Cicchetti, 2010), promoting cognitive and behavioral skills that undergird academic skill development – namely, executive functioning and self-regulation – may most efficiently focus efforts and resources seeking to prepare children for elementary school (McClelland, Acock, Piccinin, Rhea, & Stallings, 2013; Pace, Alper, Burchinal, Golinkoff, & Hirsh-Pasek, 2018; Raver et al., 2011). Though there have been recent efforts to better operationalize and differentiate executive functioning and self-regulation (Jones, Bailey, Barnes, & Partee, 2016), there is still much to learn about how these non-academic skills develop longitudinally in early childhood. Moreover, ECE programs and teachers are as yet less equipped to promote these skills than academic skills such as literacy and math that have been at the forefront of early childhood research for decades (Markowitz, Bassok, & Hamre, 2018). This is problematic because executive functioning and self-regulation develop not only through the resources children bring into the preschool classroom (i.e., prior skills; Nix et al., 2013) but also through the interactions children experience with the individuals – including teachers – and tasks in their classrooms (Downer, Booren, Lima, Luckner, & Pianta, 2010; Hamre et al., 2013). The nuances of how various characteristics of children and classroom processes interactively contribute to skill development, particularly in non-academic domains, are not yet fully understood.

As such, this three-paper dissertation addressed this gap in the literature by taking an integrated approach to examining the trajectory of early childhood skill development in the context of child, family, and classroom factors that facilitate these skills. Understanding these processes relies upon three main premises: 1) early childhood is a sensitive period for skill development; 2) it is crucial to develop non-academic skills, such as executive functioning and self-regulation, during this window of opportunity; and 3) these skills develop through the interactions and individual experiences children have, as well as the resources children bring into the preschool classroom. Specifically, the present dissertation draws from and integrates aspects of the bioecological model of human development (Bronfenbrenner & Morris, 2006), the Ecological and Dynamic Model of Transition (Rimm-Kaufman & Pianta, 2000), developmental cascades theory (Masten & Cicchetti, 2010), and the Teaching through Interactions framework (Hamre et al., 2013) to inform research questions for the three manuscripts.

Preschool as an Optimal Context for Promoting Skill Development

Early childhood is a sensitive period for skill development, in large part because it is a sensitive period for brain development (Kolb & Teskey, 2012; Phillips & Shonkoff, 2000; Shonkoff, 2011; Zelazo, Blair, & Willoughby, 2016). Rapidly developing foundational abilities are shaped by experiences (Kolb & Teskey, 2012) and are particularly malleable to intervention during this time period (Heckman & Masterov, 2007; McClelland et al., 2015). Moreover, skill accumulation is not merely additive or linear over time. Rather, developmental cascades theory posits that "competence begets competence" (Masten and Cicchetti 2010), meaning that developing foundational skills early on contributes to later school success (e.g., Bull, Espy, & Wiebe, 2008; Duncan et al., 2007; Nix, Bierman, Domitrovich, & Gill, 2013).

Having underdeveloped skills in early childhood can set children on a disadvantaged developmental trajectory and put them at risk for academic, social, emotional, and behavioral difficulties (Bierman & Erath, 2006; Masten & Cicchetti, 2010; Sabol & Pianta, 2012). This vulnerability disproportionately affects children from low-income families, with especially pronounced disparities for children from families with incomes below the poverty line (Conway, Waldfogel, & Wang, 2018). This is because (1) economic and material hardship conferred by poverty can lead to increased parental stress and less sensitive parenting practices, which in turn negatively affect young children's skill development (Bradley & Corwyn, 2002; Gershoff, Aber, Raver, & Lennon, 2007; Newland, Crnic, Cox, & Mills-Koonce, 2013), and (2) families in poverty tend to have limited financial, educational, and time resources that they can allocate to promoting their children's development (Cunha, 2015; Kalil, Ryan, & Corey, 2012; Yeung, Linver, & Brooks-Gunn, 2002). Through these mechanisms, poverty contributes to substantial achievement gaps that begin in early childhood and fail to narrow over time (Little, 2017; Reardon, 2013; Reardon & Portilla, 2016; Rowe, 2008). Given these increased risks – and longstanding consequences – for children from low-income and impoverished backgrounds, studies from this dissertation focused on understanding development primarily within this population.

Rather than trying to reduce achievement gaps after they have solidified, intervening in early childhood is an optimal time to "interrupt negative or promote positive cascades" (Masten & Cicchetti, 2010). In fact, one of the highest long-term returns on investing in individuals' development comes from boosting skills in early childhood (Cunha & Heckman, 2007; Knudsen et al., 2006). Multiple state and federal early childhood programs – including publicly funded preschool – have been developed to address the achievement gap early on, and now over two-thirds of 4-year-olds are enrolled in preprimary programs (National Center for Education Statistics [NCES], 2018).

Garnering support to effectively build skills in preschool is essential to have a successful transition to kindergarten, as asserted in Rimm-Kaufman and Pianta's (2000) Ecological and Dynamic Model of Transition. This transition is challenging for most young children, because it is typically characterized by shifts in expectations, increases in demands and responsibilities, and changes in routine toward focusing more on whole-group instruction that emphasizes academic skills (Bassok, Latham, & Rorem, 2016; Rimm-Kaufman & Pianta, 2000). This transition has become amplified in the past two decades, as kindergarten teachers have increased both their expectations of children's abilities to regulate themselves (e.g., follow directions, sit still and pay attention, finish tasks, and not act disruptively) and time spent in activities (e.g., whole group) that require more self-regulation (e.g., sustained attention; Bassok et al., 2016). This increase in kindergarten expectations has occurred without an accompanying increase in children's skills in these areas across that same time period (Bassok & Latham, 2017). Specifically, comparisons between kindergarten teacher reports in 1998 and 2010 indicate that self-regulation skills at kindergarten entry have stagnated and children's ability to pay attention has actually decreased in recent decades (Bassok & Latham, 2017). Thus, it is likely that more children are entering kindergarten already behind in key foundational skills, such as self-regulation, that set the stage for later success. It is therefore important to better understand the development of these skills in early childhood and factors that might facilitate their growth.

Importance of Non-Academic Skill Development in Early Childhood

Skill levels at kindergarten entry across domains – including language, literacy, mathematics, executive functioning, and self-regulation skills – contribute to long-term academic success (Pace et al., 2018; Ursache, Blair, & Raver, 2012). There has been a longstanding research focus on the importance of language and literacy (Dickinson & Tabors, 2001; National Early Literacy Panel [NELP], 2008; Whitehurst & Lonigan, 1998) and mathematics in early childhood (Cross, Woods, & Schweingruber, 2009; Day-Hess & Clements, 2017; Sarama & Clements, 2009). Attention has turned more recently to the significance of emerging non-academic skills, including executive functioning and self-regulation. It is important to differentiate these two constructs conceptually.

Executive functioning has been operationalized as a multidimensional cognitive skillset that underlies and facilitates goal-directed behavior and is typically comprised of working memory, inhibitory control, and set shifting abilities (Blair, Zelazo, & Greenberg, 2005; Diamond, 2013). It tends to hang together as a single construct in early childhood (Willoughby, Blair, Wirth, & Greenberg, 2012) and becomes more differentiated into component skills through later childhood and adolescence (Best & Miller, 2010; Miyake et al., 2000). Self-regulation, on the other hand, is a general umbrella term applied to the ability to control one's attention, cognitions, emotions, and behaviors to achieve a goal (e.g., Ackerman & Friedman-Krauss, 2017; Jones et al., 2016; McClelland et al., 2018). It can be considered the ability to *apply* executive functioning skills to govern, modulate, and express emotions and behaviors in context (Eisenberg & Sulik, 2012; Jones et al., 2016; Ponitz, McClelland, Matthews, & Morrison, 2009) and to "maintain optimal levels of … arousal" (Diamond, 2013). For instance, a child may be able to hold classroom rules in mind (working memory) but have difficulty executing the "hands are for helping not hurting" rule when a peer takes a toy or the "walking feet" rule when transitioning between centers. In this way, executive functioning can be seen as necessary but not sufficient for children to demonstrate regulatory abilities in the classroom (Blair & Ursache, 2011; Jones et al., 2016).

There is a large body of work indicating that executive functioning and self-regulation undergird the development of pre-academic and academic skills (Raver et al., 2011; Sabol & Pianta, 2012; Shonkoff, 2011). In particular, having higher executive functioning in preschool has been linked to growth in language, literacy, and math by the end of kindergarten (McClelland et al., 2014), as well as long-term achievement and educational attainment decades later (McClelland et al., 2013). Children's ability to regulate their behavior prior to kindergarten entry has been linked to higher first-grade academic achievement (Sektnan, McClelland, Acock, & Morrison, 2010). Compared to children with low emotion regulation, young children who are able to regulate their emotions well in preschool have been found to have positive school adjustment (Herndon, Bailey, Shewark, Denham, & Bassett, 2013) and lower behavior problems in preschool (Garner & Waajid, 2012), whereas having underdeveloped emotion regulation abilities has been shown to lead to academic and social difficulties (Ursache et al., 2012).

Empirical consensus on the importance of these cognitive, behavioral, and emotional skills has been corroborated practically. When kindergarten teachers were asked to describe readiness for school, their ranked answers were largely based on self-regulation and executive functioning; they painted a picture of children who could sit still, pay attention, follow directions, communicate their needs, and persist in tasks (Blair & Raver, 2015). This suggests that teachers consider it more important that children enter kindergarten with skills that enable them to benefit from classroom learning opportunities than that they come in with a strong academic foundation. However, the recognition that these non-academic abilities are integral to a successful launch into school has not yet been matched by children's skill gains. Whereas children's teacher-reported math and literacy skills at kindergarten entry have increased in the past two decades, teachers have not seen a commensurate gain in non-academic skills, such as self-regulation, attention, and approaches to learning (Bassok & Latham, 2017).

Although there has been increased focus on executive functioning and self-regulatory skills as important early childhood constructs in the past decade, many questions remain unanswered. Gaps in the current literature include (1) how to define these skills conceptually and methodologically, (2) how to measure them in precise and developmentally appropriate ways, (3) how executive functioning and regulation are distinct from and related to one another, (4) how these skills are related to other emerging abilities such as language, and (5) how these skills manifest in young children and develop longitudinally in the context of early learning and early childhood classroom processes (Ackerman & Friedman-Krauss, 2017; Jones et al., 2016). It is essential to better operationalize executive functioning and self-regulation, as well as to obtain a more nuanced understanding of the trajectory of their development in early childhood. Given the importance of cultivating these skills prior to kindergarten entry, a complementary line of research is to further explore factors that facilitate – or, conversely, hinder – their growth.

Experiences in Early Childhood: Interactive Contributions to Skill Development

Children's skills develop through dynamic proximal processes and frequent interactions in their immediate environment (Bronfenbrenner & Morris, 1998, 2006; Vygotsky, 1978). This dissertation takes an integrated approach to factors that influence young children's learning and development by employing the seminal work of Bronfenbrenner and Morris (ecological systems theory, 1998; and its offshoot, the bioecological model of human development, 2006) as well as more specific theoretical models, namely the Ecological and Dynamic Model of Transition (Rimm-Kaufman & Pianta, 2000) and the Teaching through Interactions framework (Hamre et al., 2013).

Ecological systems theory (Bronfenbrenner & Morris, 1998) and the bioecological model of human development (Bronfenbrenner & Morris, 2006) assert that (1) individuals develop through "processes of progressively more complex reciprocal interactions," and (2) the ways in which proximal processes affect skill development depends on characteristics and resources of the individuals themselves and their families, contextual factors, and the developmental timing in which the processes occur

(Bronfenbrenner & Morris, 2006). For young children, “bioecological resources” include accumulated skills and abilities (Bronfenbrenner & Morris, 2006); thus, it is feasible that skills in one domain accumulated prior to kindergarten entry (e.g., language skills) may facilitate the development of skills in another domain (e.g., regulation) as children learn to use language to govern behavior (Vygotsky, 1978). In addition to the importance of developmental timing, tracking processes and skill development over time (i.e., longitudinally) is critical, in part because skills build upon skills (Masten & Cicchetti, 2010). Proximal processes may make a small impression on skill development in the short term, but their effects may become magnified over time (Bronfenbrenner & Morris, 2006).

More specific to early childhood, Rimm-Kaufman & Pianta’s Ecological and Dynamic Model of Transition (2000) posits that children’s transition to kindergarten – a sensitive time period for skill development – should be understood holistically as resulting from the combined influence of children’s characteristics and ongoing interactions in the various contexts they occupy (i.e., home, school, peers, neighborhood). Importantly, this conceptualization considers factors that both directly and indirectly influence children’s successful transition to kindergarten, as well as how these factors and their influence change over time. Characterized by both change and stability, the kindergarten transition represents an opportunity as children enter the new learning environment to develop fresh interaction patterns that can either continue to positively influence adaptive behaviors or intervene to reduce maladaptive behaviors (Masten & Cicchetti, 2010; Rimm-Kaufman & Pianta, 2000).

These foundational models of human development include several levels of influencing factors, including processes occurring in environments in which children find themselves daily, such as at home or school (microsystem), all the way to widespread cultural traditions and national laws (macrosystem; Bronfenbrenner & Morris, 1998). It would be beyond the scope of this dissertation to adequately address factors at all levels influencing young children. Thus, studies in this dissertation focus explicitly on the child and microsystem levels, and specifically on the child in the context of the preschool classroom. This focus aligns with the assertion that processes that are most proximal to children's experience (i.e., daily classroom interactions versus national laws) have the most direct influences on children's skill development (Bronfenbrenner & Morris, 2006).

Specific to the context of the early childhood classroom, the Teaching through Interactions framework (Hamre et al., 2013; Hamre & Pianta, 2007) asserts that interactions between teachers and students in the classroom are primary factors influencing children's learning and development. Specifically, the extent to which teachers provide quality emotional and instructional support for children and organize the classroom to meet their needs has been consistently related to skill gains across domains (Burchinal et al., 2008; Curby, Rimm-Kaufman, & Ponitz, 2009; Maier, Vitiello, & Greenfield, 2012; Mashburn et al., 2008; Merritt, Wanless, Rimm-Kaufman, Cameron, & Peugh, 2012). More recent work has established that learning and development is not only a function of the overall quality of teacher-child interactions in the classroom but that it is also a result of the frequency and quality with which individual children engage with the individuals and materials in their classrooms (Downer et al., 2010). Specifically, positively engaging with teachers, peers, and learning activities has been related to higher levels of academic and non-academic readiness skills in preschool and primary school

(Fantuzzo et al., 2007; Ladd & Dinella, 2009). Consistent with Bronfenbrenner and Morris' (1998, 2006) and Rimm-Kaufman and Pianta's (2000) theories, children's individual engagement and the general quality of interactions in the classroom interactively contribute to skill gains in preschool (Williford, Maier, Downer, Pianta, & Howes, 2013).

A common theme in these conceptual models is that processes influencing development are multifactorial and interactive and that developmental outcomes are best understood and modeled by considering how the confluence of multiple proximal processes, child and family characteristics, and contexts over time contribute to skill growth. Empirical questions remain about the nuances of how various proximal processes in the preschool classroom work together to contribute to children's skill development, and to what extent the context of the classroom affects skill growth trajectories for children entering the classroom with different characteristics. A strength of this dissertation is that it contextualizes children's learning and development in the preschool classroom and attempts to explain how the confluence of specific processes affects developmental outcomes for individual children.

A Three-Study Approach

This dissertation conducted a line of research that collectively investigated (a) skill development in early childhood, with a focus on non-academic readiness skills; and (b) various factors that facilitate growth in these foundational abilities. In response to calls in the literature, these three papers emphasized clarity of constructs in definition and measurement (Baggetta & Alexander, 2016). They also focused on how the trajectory of skill development varied for children with different characteristics and diverse experiences in their preschool classrooms, with the idea that “one size does not fit all.” The three complementary papers varied in focus, complexity of processes explored, and methodology. For instance, the papers transitioned from a more global investigation of interactive processes to a more nuanced, granular approach to examining individualized trajectories of a specific non-academic skill over time. Specifically, Study 1 took a person-centered approach to understanding preschoolers’ engagement and explored how two types of proximal processes (i.e., classroom-level interactions and individual children’s engagement) interactively contributed to preschoolers’ skill gains in multiple domains. A question resulting from this study was why some children enter school more “ready to learn” in terms of their engagement and skills at preschool entry. Thus, Study 2 explored the interactive contributions of two cognitive skills (i.e., executive functioning and language) at preschool entry to children’s self-regulation development during the preschool year. This study also conceptually and methodologically disaggregated executive functioning and self-regulation, and investigated links between the two constructs. Finally, Study 3 comprised a fine-grained analysis of the developmental trajectory of executive functioning across the preschool-to-kindergarten transition, as well as whether and how this trajectory differed depending on characteristics of children and families.

Collectively, work from this dissertation contributes to the literature with the following findings and implications. Early childhood is a sensitive developmental period such that children experience significant gains in foundational school readiness skills during preschool and kindergarten; EF in particular grows most rapidly during preschool. There are intra-individual differences in children’s skill development based on demographic characteristics, the cognitive resources with which children arrive at preschool, and the ways in which children interact in the classroom. Specifically, children in poverty, boys, and African American children tend to have more disadvantaged EF trajectories than peers; language skills are promotive of and protective for children's emotion regulation development; and children who have lower positive engagement and relatively higher negative engagement than peers tend to make fewer gains during preschool. Finally, findings point to ways to reduce achievement gaps early on by (1) identifying children who are at risk of slower growth trajectories who might benefit from additional opportunities to develop skills, (2) suggesting when skill-boosting interventions might be particularly beneficial, and (3) identifying classroom environments (e.g., classrooms characterized by consistent routines and positive, proactive behavior management strategies) that are particularly promotive of and protective for the learning and development of young children.

The remainder of this dissertation is comprised of the three manuscripts in full, along with their associated tables and figures.

References

Ackerman, D. J., & Friedman-Krauss, A. H. (2017). Preschoolers' executive function: Importance, contributors, research needs and assessment options. *ETS Research Report Series*, *2017*, 1-24. doi: 10.1002/ets2.12148.

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

Bassok, D., & Latham, S. (2017). Kids today: The rise in children's academic skills at kindergarten entry. *Educational Researcher*, *46*, 7-20. doi: 10.3102/0013189X17694161.

Bassok, D., Latham, S., & Rorem, A. (2016). Is kindergarten the new first grade?. *AERA Open*, *2*, 2332858415616358. doi: 10.1177/2332858415616358.

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

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

Blair, C., & Raver, C. C. (2015). School readiness and self-regulation: A developmental psychobiological approach. *Annual Review of Psychology*, *66*, 711-731. doi: 10.1146/annurev-psych-010814-015221.

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, (Vol. 2, pp. 300-320). New York, NY: Guilford.

Blair, C., Zelazo, P. D., & Greenberg, M. T. (2005). The measurement of executive function in early childhood. *Developmental Neuropsychology, 28,* 561-571. doi: 10.4324/9780203764244.

Bradley, R. H., & Corwyn, R. F. (2002). Socioeconomic status and child development. *Annual Review of Psychology*, *53*, 371-399. doi: 10.1146/annurev.psych.53.100901.135233.

Bronfenbrenner, U., & Morris, P. A. (1998). The ecology of developmental processes. In W. Damon (Series Ed.) & R. M. Lerner (Vol. Ed.), *Handbook of child psychology: Theory* (5th ed., Vol. 1). New York, NY: Wiley.

Bronfenbrenner, U. & Morris, P. A. (2006). The bioecological model of human development. In R. M. Lerner (Ed.), *Handbook of child psychology: Theoretical models of human development.* (6th ed., Vol. 1). Hoboken, NJ: Wiley.

Bull, R., Espy, K. A., & Wiebe, S. A. (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. doi: 10.1080/87565640801982312.

Burchinal, M., Howes, C., Pianta, R., Bryant, D., Early, D., Clifford, R., & Barbarin, O. (2008). Predicting child outcomes at the end of kindergarten from the quality of pre-kindergarten teacher-child interactions and instruction. *Applied Developmental Science, 12*, 140-153. doi: 10.1080/10888690802199418.

Conway, A., Waldfogel, J., & Wang, Y. (2018). Parent education and income gradients in children's executive functions at kindergarten entry. *Children and Youth Services Review, 91*, 329-337. doi: 10.1016/j.childyouth.2018.06.009.

Cross, C. T., Woods, T. A., & Schweingruber, H. (Eds.) (2009). *Mathematics learning in early childhood: Paths toward excellence and equity*. Washington, DC: National Academies Press.

Cunha, F. (2015). Subjective rationality, parenting styles, and investments in children. In P. R. Amato, A. Booth, S. M. McHale, & J. Van Hook (Eds.), *Families in an era of increasing inequality* (pp. 83-94). Cham, Switzerland: Springer International Publishing.

Cunha, F., & Heckman, J. (2007). The technology of skill formation. *American Economic Review*, *97*, 31-47. doi:10.1257/aer.97.2.31.

Curby, T. W., Rimm-Kaufman, S. E., & Ponitz, C. C. (2009). Teacher–child interactions and children's achievement trajectories across kindergarten and first grade. *Journal of Educational Psychology*, *101*, 912-925. doi: 10.1037/a0016647.

Day-Hess, C., & Clements, D. H. (2017). The DREME network: Research and interventions in early childhood mathematics. *Advances in Child Development and Behavior, 53*, 1-41. doi: 10.1016/bs.acdb.2017.03.002.

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

Dickinson, D. K., & Tabors, P. O. (2001). *Beginning literacy with language: Young children learning at home and school*. Baltimore, MD: Paul H. Brookes Publishing.

Downer, J. T., Booren, L. M., Lima, O. K., Luckner, A. E. & Pianta, R. C. (2010). The

Individualized Classroom Assessment Scoring System (inCLASS): Preliminary reliability and validity of a system for observing preschoolers' competence in classroom interactions. *Early Childhood Research Quarterly, 25*, 1–16. doi: 10.1016/j.ecresq.2009.08.004.

Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., ... & Japel, C. (2007). School readiness and later achievement. *Developmental Psychology*, *43*, 1428-1446. doi: 10.1037/0012-1649.43.6.1428.supp.

Eisenberg, N., & Sulik, M. J. (2012). Emotion-related self-regulation in children. *Teaching of Psychology*, *39*, 77-83. doi: 10.1177/0098628311430172.

Fantuzzo, J., Bulotsky-Shearer, R., McDermott, P., McWayne, C., Frye, D., & Perlman, S. (2007). Investigation of dimensions of social-emotional classroom behavior and school readiness for low-income urban preschool children. *School Psychology Review, 36*, 44-62.

Garner, P. W., & Waajid, B. (2012). Emotion knowledge and self-regulation as predictors of preschoolers' cognitive ability, classroom behavior, and social competence. *Journal of Psychoeducational Assessment*, *30*, 330-343. doi: 10.1177/0734282912449441.

Gershoff, E. T., Aber, J. L., Raver, C. C., & Lennon, M. C. (2007). Income is not enough: Incorporating material hardship into models of income associations with parenting and child development. *Child Development*, *78*, 70-95. doi: 10.1111/j.1467-8624.2007.00986.x.

Hamre, B. K., & Pianta, R. C. (2007). Learning opportunities in preschool and early elementary classrooms. In R. C. Pianta, M. J. Cox, & K. L. Snow (Eds.), *School* *readiness and the transition to kindergarten in the era of accountability* (pp. 49-83). Baltimore, MD: Paul H Brookes Publishing.

Hamre, B. K., Pianta, R. C., Downer, J. T., DeCoster, J., Mashburn, A. J., Jones, S. M., ... & Brackett, M. A. (2013). Teaching through interactions: Testing a developmental framework of teacher effectiveness in over 4,000 classrooms. *The Elementary School Journal*, *113*, 461-487. doi: 10.1086/669616.

Heckman, J. J., & Masterov, D. V. (2007). The productivity argument for investing in young children. *Applied Economic Perspectives and Policy, 29,* 446-493. doi: 10.1111/j.1467-9353.2007.00359.x.

Herndon, K. J., Bailey, C. S., Shewark, E. A., Denham, S. A., & Bassett, H. H. (2013). Preschoolers' emotion expression and regulation: Relations with school adjustment. *The Journal of Genetic Psychology*, *174*, 642-663. doi: 10.1080/00221325.2012.759525.

Jones, S. M., Bailey, R., Barnes, S. P., & Partee, A. (2016). *Executive function mapping project: Untangling the terms and skills related to executive function and self regulation in early childhood* (No. 2016-88). Cambridge, MA: U.S. Department of Health and Human Services, Office of Planning, Research, and Evaluation.

Kalil, A., Ryan, R., & Corey, M. (2012). Diverging destinies: Maternal education and the developmental gradient in time with children. *Demography*, *49*, 1361-1383. doi: 10.1007/s13524-012-0129-5.

Knudsen, E. I., Heckman, J. J., Cameron, J. L., & Shonkoff, J. P. (2006). Economic, neurobiological, and behavioral perspectives on building America's future workforce. *Proceedings of the National Academy of Sciences*, *103*, 10155-10162.

doi: 10.1073/pnas.0600888103.

Kolb, B., & Teskey, G. C. (2012). Age, experience, injury, and the changing brain. *Developmental Psychobiology*, *54*, 311-325. doi: 10.1002/dev.20515.

Ladd, G. W., & Dinella, L. M. (2009). Continuity and change in early school engagement: Predictive of children's achievement trajectories from first to eighth grade?. *Journal of Educational Psychology*, *101*, 190-206. doi: 10.1037/a0013153.

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*, 103-109. doi: 10.3102/0013189X17698700.

Maier, M. F., Vitiello, V. E., & Greenfield, D. B. (2012). A multilevel model of child-and classroom-level psychosocial factors that support language and literacy resilience of children in Head Start. *Early Childhood Research Quarterly*, *27*, 104-114. doi: 10.1016/j.ecresq.2011.06.002.

Markowitz, A. J., Bassok, D., & Hamre, B. (2018). Leveraging developmental insights to improve early childhood education. *Child Development Perspectives*, *12*, 87-92. doi: 10.1111/cdep.12266.

Mashburn, A. J., Pianta, R. C., Hamre, B. K., Downer, J. T., Barbarin, O. A., Bryant, D., ... & Howes, C. (2008). Measures of classroom quality in prekindergarten and children’s development of academic, language, and social skills. *Child Development*, *79*, 732-749. doi: 10.1111/j.1467-8624.2008.01154.x.

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

McClelland, M. M., Acock, A. C., Piccinin, A., Rhea, S. A., & Stallings, M. C. (2013). Relations between preschool attention span-persistence and age 25 educational outcomes. *Early Childhood Research Quarterly*, *28*, 314-324. doi: 10.1016/j.ecresq.2012.07.008.

McClelland, M. M., Cameron, C. E., Duncan, R., Bowles, R. P., Acock, A. C., Miao, A., & Pratt, M. E. (2014). Predictors of early growth in academic achievement: The head-toes-knees-shoulders task. *Frontiers in Psychology*, *5*, 1-14. doi: 10.3389/fpsyg.2014.00599.

McClelland, M. M., Geldhof, G. J., Cameron, C. E., & Wanless, S. B. (2015). Development and self-regulation. In R. M. Lerner (Ed.), *Handbook of Child Psychology and Developmental Science*, (7th ed., pp. 1-43). doi: 10.1002/9781118963418.childpsy114.

Merritt, E. G., Wanless, S. B., Rimm-Kaufman, S. E., Cameron, C., & Peugh, J. L. (2012). The contribution of teachers' emotional support to children's social behaviors and self-regulatory skills in first grade. *School Psychology Review*, *41*, 141-159.

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*, 49-100. doi: 10.1006/cogp.1999.0734.

National Center for Education Statistics. (2018). *The condition of education: Preschool and kindergarten enrollment*. Retrieved from: https://nces.ed.gov/programs/coe/indicator_cfa.asp.

National Early Literacy Panel. (2008). *Developing early literacy: Report of the National Early Literacy Panel.* Washington, DC: National Institute for Literacy. Retrieved from: https://lincs.ed.gov/publications/pdf/NELPReport09.pdf.

Newland, R. P., Crnic, K. A., Cox, M. J., & Mills-Koonce, W. R. (2013). The family model stress and maternal psychological symptoms: Mediated pathways from economic hardship to parenting. *Journal of Family Psychology*, *27*, 96-105. doi: 10.1037/a0031112.

Nix, R. L., Bierman, K. L., Domitrovich, C. E., & Gill, S. (2013). Promoting children's social-emotional skills in preschool can enhance academic and behavioral functioning in kindergarten: Findings from Head Start REDI. *Early Education & Development*, *24*, 1000-1019. doi: 10.1080/10409289.2013.825565.

Pace, A., Alper, R., Burchinal, M. R., Golinkoff, R. M., & Hirsh-Pasek, K. (2018). Measuring success: Within and cross-domain predictors of academic and social trajectories in elementary school. *Early Childhood Research Quarterly*, *46,* 112-125. doi: 10.1016/j.ecresq.2018.04.001.

Phillips, D. A., & Shonkoff, J. P. (Eds.). (2000). *From neurons to neighborhoods: The science of early childhood development*. Washington, DC: National Academies Press.

Ponitz, C. 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*, 605-619. doi: 10.1037/a0015365.

Raver, C. C., Jones, S. M., 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. doi: 10.1111/j.1467-8624.2010.01561.x.

Reardon, S. F. (2013). The widening income achievement gap. *Educational Leadership*, *70*, 10-16.

Reardon, S. F., & Portilla, X. A. (2016). Recent trends in income, racial, and ethnic school readiness gaps at kindergarten entry. *AERA Open*, *2*(3), 2332858416657343. doi: 10.1177/2332858416657343.

Rimm-Kaufman, S. E., & Pianta, R. C. (2000). An ecological perspective on the transition to kindergarten: A theoretical framework to guide empirical research. *Journal of Applied Developmental Psychology*, *21*, 491-511. doi: 10.1016/S0193-3973(00)00051-4.

Rowe, M. L. (2008). Child-directed speech: relation to socioeconomic status, knowledge of child development and child vocabulary skill. *Journal of Child Language*, *35*, 185-205. doi: https://doi.org/10.1017/S0305000907008343.

Sabol, T. J., & Pianta, R. C. (2012). Patterns of school readiness forecast achievement and socioemotional development at the end of elementary school. *Child Development*, *83*, 282-299. doi: 10.1111/j.1467-8624.2011.01678.x.

Sarama, J., & Clements, D. H. (2009). *Early childhood mathematics education research: Learning trajectories for young children*. New York, NY: Routledge.

Sektnan, M., McClelland, M. M., Acock, A., & Morrison, F. J. (2010). Relations between early family risk, children's behavioral regulation, and academic achievement. *Early Childhood Research Quarterly*, *25*, 464-479. doi: 10.1016/j.ecresq.2010.02.005.

Shonkoff, J. P. (2011). Protecting brains, not simply stimulating minds. *Science*, *333*, 982-983. doi: 10.1126/science.1206014.

Ursache, A., Blair, C., & Raver, C. C. (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. doi: 10.1111/j.1750-8606.2011.00209.x.

Whitehurst, G. J., & Lonigan, C. J. (1998). Child development and emergent literacy. *Child Development*, *69*, 848-872. doi: 10.1111/j.1467-8624.1998.tb06247.x.

Williford, A. P., Maier, M. F., Downer, J. T., Pianta, R. C., & Howes, C. (2013). Understanding how children's engagement and teachers' interactions combine to predict school readiness. *Journal of Applied Developmental Psychology*, *34*, 299-309. doi: 10.1016/j.appdev.2013.05.002.

Willoughby, M. T., Blair, C. B., Wirth, R. J., & Greenberg, M. (2012). The measurement of executive function at age 5: Psychometric properties and relationship to academic achievement. *Psychological Assessment*, *24*, 226-239. doi: 10.1037/a0025361.

Yeung, W. J., Linver, M. R., & Brooks–Gunn, J. (2002). How money matters for young children's development: Parental investment and family processes. *Child Development*, *73*, 1861-1879. doi: 10.1111/1467-8624.t01-1-00511.

Zelazo, P. D., Blair, C. B., & Willoughby, M. T. (2016). *Executive function: Implications for education (NCER 2017-2000)*. Washington, DC: National Center for Education Research, Institute of Education Sciences, U.S. Department of Education.

MANUSCRIPT ONE

Roles of Children's Engagement and Teacher-Child Interactions in Developing School Readiness Skills

Shannon E. Reilly, M.Ed., Jason T. Downer, Ph.D., and Amanda P. Williford, Ph.D.

Center for Advanced Study of Teaching and Learning, University of Virginia

Revision requested, *Journal of Applied Developmental Psychology*

Author Note

Shannon Reilly, Center for Advanced Study of Teaching and Learning, University of Virginia; Jason T. Downer, Center for Advanced Study of Teaching and Learning, University of Virginia; Amanda P. Williford, Center for Advanced Study of Teaching and Learning, University of Virginia.

This study was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R305A060021 to the University of Virginia — funding the National Center for Research on Early Childhood Education (NCRECE) as well as the National Institute of Child Health and Human Development and the Interagency Consortium on Measurement of School Readiness: R01 HD051498. The opinions expressed are those of the authors and do not represent views of the funding agencies. Special thanks to the participating teachers, children, and families.

Correspondence concerning this article should be addressed to Shannon Reilly, Center for Advanced Study of Teaching and Learning, University of Virginia, PO Box 400267, Charlottesville, VA 22904. Email: sr6hd@virginia.edu.

Abstract

Developing school readiness skills in preschool provides an important foundation for school success and may help to reduce achievement gaps. Two proximal factors affecting school readiness outcomes, children's engagement and the quality of teacher-child interactions in the classroom, provide promising routes to facilitate the development of these skills. Using a low-income, racially and ethnically diverse sample, the present study examined associations between children's patterns of engagement in their classrooms and their school readiness skill development across the preschool year, and to what extent these relationships depend on the quality of teacher-child interactions. Children who tended to be Negatively Engaged made the smallest gains across domains during the preschool year. Being in a classroom with higher positive management and routines bolstered gains specifically for children who tended to be Negatively Engaged. Findings have implications for how to support skill development in children at risk for low engagement in the preschool classroom.

Keywords: preschool; classroom quality; teacher-child interactions; engagement; school readiness

Roles of Children's Engagement and Teacher-Child Interactions in Developing School Readiness Skills

Developing foundational skills early on sets the stage for young children's later school success across developmental domains (Bull, Espy, & Wiebe, 2008; Duncan et al., 2007; Nix, Bierman, Domitrovich, & Gill, 2013). Socioeconomic gaps in these skills begin even prior to kindergarten entry (Rowe, 2008) and fail to narrow over time (Raver, McCoy, Lowenstein, & Pess, 2013; Reardon, 2013). Thus, preschool is an ideal context to promote these skills among children – especially those at risk for low achievement – because foundational skills are newly developing and malleable to intervention at this time (Heckman & Masterov, 2007; McClelland, Geldhof, Cameron, & Wanless, 2015; Whitehurst & Lonigan, 2001). Children's skills develop through dynamic processes and interactions in their immediate environment (e.g., Bronfenbrenner & Morris, 1998). Taking a child-centered approach, this process of skill development depends not only on the larger classroom environment but also on what children bring to the interactions (Downer, Booren, Lima, Luckner, & Pianta, 2010; Fredericks, Blumenfeld, & Paris, 2004). This study examines two types of proximal processes in the preschool classroom and how these processes interact to contribute to skill development.

First, the ways in which children engage with available individuals and materials in their classroom has been linked to their school readiness skill development (Bulotsky-Shearer, Fernandez, Dominguez, & Rouse, 2011; Vitiello & Williford, 2016; Williford, Maier, Downer, Pianta, & Howes, 2013). Importantly, experiences with teachers, peers, and tasks do not occur in isolation from one another. Children's positive and negative interactions across these contexts manifest as *patterns* (i.e., profiles) of engagement that have been associated with skill gains (Williford et al., 2013). Moreover, children's engagement does not exist in a vacuum; rather, it is dependent on support and opportunities provided in the classroom (Booren, Downer, & Vitiello, 2012; Hamre, Hatfield, Pianta, & Jamil, 2014). As such, it is essential to consider that children are embedded in classroom contexts characterized by interactions between teachers and children that vary widely in quality (Burchinal, Vandergrift, Pianta, & Mashburn, 2010). Thus, the second set of proximal processes of interest in the present study is the quality of these classroom-wide interactions, which has also been associated with young children's skill development (e.g., Maier, Vitiello, & Greenfield, 2012; Merritt, Wanless, Rimm-Kaufman, Cameron, & Peugh, 2012). Recent evidence suggests that children's individual engagement and the general quality of teacher-child interactions in the classroom contribute to skill gains independently (Sabol, Bohlmann, & Downer, 2018) and in an interactive way (Williford et al., 2013). What is not yet known is how patterns of children's engagement interact with the quality of *specific* types of teacher-child interactions to promote early skill development. This study addresses this limitation, with the goal of unpacking how targeted, high-quality interactions in the preschool classroom might act as promotive or protective factors for the development of children from a racially and ethnically diverse, low-income sample.

Young Children's Engagement in the Classroom Related to Skill Development

Engagement has been conceptualized as a multidimensional construct that can be observed behaviorally in the ways in which young children positively and negatively interact with the individuals and materials in their immediate environment (Fredericks et al., 2004). Notably, engagement represents more than an isolated characteristic of an individual child because of the dynamic, bidirectional nature of interactions; instead, it should be considered as the confluence of what children bring into the classroom and the opportunities their environment affords. As such, examining engagement is not meant to label children but rather to describe the ways in which each child tends to interact in a particular classroom comprised of specific individuals and materials. Positively engaging with teachers, peers, and learning activities has been linked with higher academic and non-academic readiness skills over time (e.g., Fantuzzo et al., 2007; Ladd & Dinella, 2009). Conversely, engaging negatively with teachers and peers and having an underdeveloped ability to manage one's own behaviors have been related to lower skill development (e.g., Bulotsky-Shearer et al., 2011). The following provides greater detail about children's engagement across teachers, peers, and tasks, and how engagement relates to other child characteristics.

Engagement with teachers. Positively engaging with teachers can be characterized as the degree to which young children are sociable with and communicate with teachers (Downer et al., 2010). Children who display high levels of engagement with teachers seek them out, share positive affect (e.g., smiling together), and consistently initiate and sustain conversations with them both to have their needs met and to interact socially. Positive engagement with teachers has been associated with higher compliance and executive functioning (Williford et al., 2013) as well as academic skills (Birch & Ladd, 1997; O'Connor & McCartney, 2007).

Engagement with peers. Positive engagement with peers in the preschool classroom involves similar interactions as with teachers, including being sociable and initiating and maintaining conversations (Downer et al., 2010). Another facet of engaging with peers is the degree to which children show positive leadership strategies by asserting their ideas and teaching peers. High levels of prosocial behaviors with peers and interactive peer play have been linked to the development of school readiness skills, whereas disconnection with peers has been related to lower skills (Bulotsky-Shearer, Bell, Romero, & Carter, 2012; Fantuzzo & McWayne, 2002).

Engagement with tasks. In addition to interacting with individuals in the preschool classroom, young children engage with tasks and learning activities. Children who display high levels of positive engagement with tasks sustain attention on and actively participate in activities and extend their own learning by challenging themselves (Downer et al., 2010). Positively approaching learning with attentiveness and persistence has been related to higher levels of academic skills (Bohlmann & Downer, 2016; Li-Grining, Votruba-Drzal, Maldonado-Carreno, & Haas, 2010; Vitiello & Williford, 2016).

Negative engagement in the classroom. Children can also engage with individuals and tasks in ways that are conflictual or negative. This includes not only overt verbal and physical aggression toward teachers and peers but also subtler instances of negative engagement, such as displaying negative affect (e.g., frowning), resisting connections, not complying with teacher requests, and being confrontational with peers (Downer et al., 2010). With tasks, children can exhibit difficulty matching classroom expectations or respecting others' personal space. Engaging in higher levels of problem behaviors with teachers, peers, and learning tasks in Head Start, a federally funded preschool program targeted toward low-income children, and resisting connection in elementary school have both been linked with lower reading skills and other academic outcomes over time (Bulotsky-Shearer et al., 2011; Ladd & Dinella, 2009).

Engagement and other child characteristics. The extent to which children engage in the classroom has been related to other demographic characteristics and concurrently developing skills. For instance, older children tend to have higher levels of positive engagement across multiple domains (Fantuzzo et al., 2007; Vitiello, Booren, Downer, & Williford, 2012). Gender (Fantuzzo et al., 2007; Vitiello et al., 2012) and socioeconomic status (Liu, 2016; Stipek & Ryan, 1997) have been more inconsistently linked with engagement. Conflict with teachers has been related to special education referral in elementary school, but engagement does not consistently predict special education status (Buckrop & LoCasale-Crouch, 2016). Thus, negative engagement is related to but distinct from behavior problems or special needs status. Engagement has been found to predict academic achievement across the elementary school years even when accounting for these and other demographic characteristics (e.g., Ladd & Dinella, 2009). However, given that children's engagement in the classroom is likely related to characteristics that they bring into the classroom, it is important to understand how engagement patterns relate to demographic information available about children.

A Person-Centered Approach to Young Children's Engagement

Although there has been much recent work on children's engagement and how it relates to skill development, the majority has focused on a single type of engagement or interaction in isolation, such as with teachers (e.g., O'Connor & McCartney, 2007), peers (e.g., Fantuzzo & McWayne, 2002), tasks (e.g., Li-Grining et al., 2010), or problem behaviors (e.g., Bulotsky-Shearer et al., 2011). Even when multiple dimensions are assessed, these engagement-related behaviors are often captured through retrospective teacher and/or parent report of specific, isolated behaviors (Arnold, Kupersmidt, Voegler-

Lee, & Marshall, 2012; Fantuzzo et al., 2007; Ladd & Dinella, 2009). This provides useful information but is also subject to biases that may result in a child being rated as globally “high” or “low” on measures. What is less well understood is how individual children’s positive and negative engagement manifests *across* contexts in the classroom. It is important to incorporate and integrate different aspects of engagement because it is a multidimensional construct (Fredricks et al., 2004). Moreover, individual children can display both positive and negative engagement, such as seeking out and emotionally connecting with teachers and peers but also experiencing conflict or confrontation with them. As another example, children can be actively engaged in a learning task but also have difficulty regulating their behavior to match classroom expectations. As such, examining both positive and negative engagement with all available individuals and materials *together* provides a more comprehensive picture of children’s experiences in the early childhood classroom. Taking a “person-centered” approach to children’s engagement by identifying patterns of engagement in the classroom is one way to address this issue. This enables comparison across individuals rather than across variables, and allows for exploration of both qualitative and quantitative differences in presentations of engagement across contexts in the classroom (Marsh, Ludtke, Trautwein, & Morin, 2009).

A limited body of literature takes a “person-centered” approach to engagement. Even with this approach, some studies have focused on patterns of engagement from a specific valence (e.g., through the lens of problem behaviors only; Bulotsky-Shearer et al., 2011) or in specific activity settings (Chien et al., 2010). Others have incorporated multiple aspects of engagement (e.g., feeling emotionally connected to school, participating cooperatively as a student) but were conducted with older children and relied on teacher and parent report (Ladd & Dinella, 2009). Little is yet known about children's patterns of observed positive and negative engagement across teachers, peers, and tasks in the preschool classroom.

One recent exception to this dearth of research is a study by Williford and colleagues (2013), which identified three patterns of engagement. The majority of children were observed to be "typically engaged," characterized by "relatively low positive engagement with teachers and peers, moderate engagement with tasks," and low negative engagement (Williford et al., 2013). About one quarter was classified as "positively engaged," with significantly higher engagement with teachers, peers, and tasks than their typically engaged counterparts. A small subset tended to be "negatively engaged" and had lower engagement with tasks and higher negative engagement than their peers. Moreover, these engagement patterns were related to academic skill gains across the preschool year such that being classified as "positively engaged" was associated with making relatively larger gains in vocabulary and executive functioning skills; children who tended to be negatively engaged made the smallest gains across groups. The current study sought to replicate these three patterns of engagement and their relationships to school readiness gains.

Children's Engagement and School Readiness Development in the Classroom Context

The ways in which individual children experience and engage in their classrooms are embedded in a larger classroom context, characterized by frequent interactions between teachers and children (Hamre & Pianta, 2005, 2007). Moreover, children's engagement depends in part on the opportunities and support provided by this context (Booren et al., 2012; Vitiello et al., 2012). There is a global literature (e.g., from the United States, Chile, and Finland) indicating that the quality of teacher-child interactions in the classroom is related to children's learning and development of readiness skills (Hamre et al., 2013; Leyva et al., 2015; Pakarinen et al., 2014). Recent research suggests that the relationship between children's engagement and skill development may in part be dependent on the quality of teacher-child interactions in the classroom (Williford et al., 2013). Thus, it is important to examine how the quality of teacher-child interactions in the classroom might intersect with engagement patterns to facilitate or hinder foundational skill development.

A bifactor approach to teacher-child interactions. The extent to which teachers provide emotional and instructional support for children and organize the classroom to meet their needs has been consistently linked with children's skill development in multiple domains (Bierman, Nix, Greenberg, Blair, & Domitrovich, 2008; Burchinal et al., 2008; Curby, Rimm-Kaufman, & Ponitz, 2009; Mashburn et al., 2008). More recently, teacher-child interaction quality has been conceptualized from a bifactor approach (Chen, Hayes, Carver, Laurenceau, & Zhang, 2012) that has identified a global, general factor of teacher responsiveness that promotes learning and development broadly, as well as two domain-specific factors (i.e., cognitive facilitation and positive management and routines; Hamre et al., 2014). Hamre and colleagues (2014) found that general teacher responsiveness was associated with child outcomes across developmental domains, whereas more targeted interactions were linked to specific skills. In particular, cognitive facilitation predicted language and literacy skills, and positive management and routines was associated with inhibitory control. Taking a bifactor approach allows for a more precise understanding of the roles of general and domain-specific aspects of teacher-child interactions in children's learning and development. However, it is yet unclear whether and how these newly conceptualized types of interactions facilitate skill development for children who exhibit different patterns of engagement with teachers, peers, and tasks in their classrooms.

Interactive contributions of children's engagement and classroom interactions. A growing body of research suggests that children's engagement and the quality of teacher-child interactions in the classroom independently predict school readiness outcomes. Recently, Sabol et al. (2017) found that aspects of individual children's engagement predicted gains in directly assessed language, literacy, and executive functioning skills across the preschool year, over and above contributions of the average quality of teacher-child interactions in the classroom. There is a longstanding and widely respected conceptual argument that these factors are not merely additive but that they interact to facilitate early skill development (Bronfenbrenner & Morris, 1998). Preliminary empirical evidence suggests that children who are less than positively engaged have increased skill gains in classrooms characterized by globally high-quality interactions (Ponitz, Rimm-Kaufman, Grimm, & Curby, 2009; Williford et al., 2013). For example, in classrooms with relatively high-quality teacher-child interactions, children classified as "typically engaged" exhibited gains across the year similar to those of peers who tended to be "positively engaged" (Williford et al., 2013).

However, to our knowledge no research has yet examined how *specific* aspects of teacher-child interaction quality in the classroom might matter differently for promoting school readiness skills of children who display distinct patterns of engagement in the classroom. For instance, classrooms characterized solely by high-quality general teacher responsiveness but lacking in more targeted interactions may not be sufficient to boost the trajectory of early skill development for children who exhibit relatively high levels of negative engagement and low levels of positive engagement with the teachers, peers, and tasks in their classrooms. Rather, specific high-quality interactions targeted toward facilitating cognitive development and positively managing the classroom may be necessary to augment skill gains of these at-risk preschoolers (Greenwood et al., 2011; Hemmeter, Ostrosky, & Fox, 2006; Powell, Dunlap, & Fox, 2006). Thus, a final goal of this paper is to understand whether and how the quality of general and domain-specific types of interactions facilitates skill gains for preschoolers who exhibit different patterns of engagement across teachers, peers, and tasks. Doing so may contribute to unpacking the complex developmental pathways through which the quality of preschool classrooms contributes to individual children's school readiness skill development.

The Present Study

The present study sought to replicate and extend previous work by examining how proximal processes in the preschool classroom are associated with children's skill gains across the preschool year. Specifically, the research questions are as follows:

1) What do individual children's patterns of engagement look like in the preschool classroom, and how are these patterns associated with demographic characteristics of children?

2) How are these engagement patterns associated with gains in school readiness skills across the preschool year, and do patterns of engagement continue to relate to school readiness skills when classroom quality is taken into account?

3) To what extent do associations between engagement patterns and skill gains depend on the quality of specific types of teacher-child interactions in the classroom?

We expected to replicate the findings of Williford and colleagues (2013) by identifying three profiles of children's engagement patterns, such that the majority of children would be "typically engaged," with moderate teacher, peer, and task engagement and low negative engagement; the next largest group would fall into the "positively engaged" profile, with higher teacher, peer, and task engagement than typically engaged peers; and a smaller subset would be considered "negatively engaged," with relatively lower teacher, peer, and task engagement and higher negative engagement than peers. We expected that children classified in the "negatively engaged" profile, which is arguably the most "at risk," would be more likely to have other demographic risk factors (e.g., lower income-to-needs ratio), but that the correlation would not be so high as to suggest that being classified as "negatively engaged" is merely a proxy for other risk factors. We expected that children classified as "positively engaged" would make the largest skill gains across the year and that those classified as "negatively engaged" would make the smallest gains (Williford et al., 2013). We hypothesized that these associations would be maintained when the quality of teacher-child interactions in the classroom was taken into account (Sabol et al., 2018; Williford et al., 2013). Finally, we expected that classrooms characterized by high-quality teacher-child interactions would be protective for the skill development of children classified as less than positively engaged (i.e., typically or negatively engaged; Williford et al., 2013). Specifically, we hypothesized that cognitive facilitation would be especially promotive for the language and literacy development of these children and that positive management and routines would facilitate development of their executive functioning skills (Greenwood et al., 2011).

Method

Sample

Data for the present study were collected as part of the National Center for Research on Early Childhood Education (NCRECE) Professional Development Study (Downer, Pianta et al., 2012; Hamre et al., 2012), a randomized controlled study that sought to improve classroom interactions and boost children's skills. Data for the present study were drawn from the third phase, a follow-up period during which no intervention took place. For full information about the intervention and its results, see work by Downer and colleagues (2013) and Hamre and colleagues (2012). The impact of the intervention was not of interest in the current study but was controlled for in analyses. The NCRECE Professional Development Study was conducted in 10 large community preschool and Head Start sites across eight states. Teachers were eligible for participation if they: (a) were the lead teacher in a classroom in which the majority of children were eligible for kindergarten the following school year, (b) conducted instruction in English for the majority of the day, and (c) had access to high-speed Internet. Children in participating classrooms without an IEP were eligible for the study; four consented children per classroom were randomly selected.

A total of 220 teachers participated in the post-intervention phase and were included in the present study. They were 96% female with a mean age of 42.55 years (*SD* = 10.59, range = 22-69); 33% were Caucasian, 46.7% were African American, 11.8% were Hispanic or Latino, and 8.5% were of other racial or ethnic backgrounds. On average, teachers had 15.90 years of education (*SD* = 1.62, range = 12-20) and 8.45 years of experience at their current facility (*SD* = 6.41, range: 1-35 years). Approximately 54 percent of classrooms were in public or non-public Head Start, 24% were housed in other public schools, and 23% were in a public agency, non-profit, or for-profit setting. Class sizes ranged from 7-32 children, with a mean of approximately 17 (*SD* = 2.85). On average, 48% of children per class were female, 15.5% had limited English proficiency (LEP; *SD* = 20.6%), and 9% had an IEP or IFSP (*SD* = 11.73%). Approximately 87% of children in a given classroom were from low-income backgrounds (*SD* = 23%).

Within these 220 classrooms, a total of 895 eligible and consented children participated in the larger NCRECE Professional Development Study. The analytic sample for the present study was restricted to 710 children in these classrooms on whom observational engagement data were collected. Approximately 50% of children in the analytic sample were male; 41% percent were African American, 14% were Caucasian, 35% were Hispanic/Latino, and 20% were of other racial or ethnic backgrounds. English was not the primary language for approximately 15% of children. At the beginning of preschool, children had a mean age of 49.60 months (*SD* = 6.01), and their mothers had on average 13 years of education (*SD* = 2.39, range: 8-20). Average household size was 4.42 (*SD* = 1.59, range = 2-15), and families had an average income-to-needs ratio of 1.10 (*SD* = 1.05, range = 0.05-5.07), meaning that the average family in the study had an annual income just above the poverty level for their household size. There was an average of 158 days between fall and spring assessments (*SD* = 35.10). Few statistically significant differences emerged between children included in the final analytic sample (*n* = 710) and those who were not because they were missing observational engagement data (*n* = 185). Children in the final sample were significantly older, had more time between assessments, had teachers with more years of education, and were more likely to be in non-public Head Start settings.

Procedures

Recruitment. Program administrators in 10 urban areas across the United States were contacted to participate in the study, and administrators and teachers were invited to attend study recruitment sessions. Interested and eligible teachers gave informed consent, completed personal and classroom demographic surveys, and allowed data collectors to observe their classrooms. Parents or guardians of children in participating classrooms were given a letter explaining the study, an informed consent, and a demographic survey. Of consented children who did not have an IEP or IFSP, four from each classroom were randomly selected to participate in the study.

Data collection. Parents and teachers completed demographic surveys in the fall. Trained data collectors administered direct child assessments of English expressive and receptive vocabulary, phonological awareness, print knowledge, inhibitory control, and working memory in the fall and spring. Children whose primary language was Spanish were also administered receptive and expressive vocabulary subtests in Spanish. In the present study, only English assessments were used for all children. Trained data collectors observed teacher-child interactions at the classroom level and engagement at the individual child level mid-year.

Observation training and protocol. Data collectors attended a two-day training session for each of two observational measures: the Classroom Assessment Scoring System (CLASS; Pianta, La Paro, & Hamre, 2008), a classroom-level measure of teacher-child interactions, and the Individualized Classroom Assessment Scoring System (inCLASS; Downer et al., 2010), a measure of individual children's engagement. Each training involved a detailed review of all measure dimensions and a process of watching, coding, and discussing five training videos. To achieve reliability and become certified on each measure, data collectors had to code five videos independently and score within one point of a master code on 80% of dimensions. At this training, data collectors coded within one of the master code on 84% to 91% of CLASS dimensions and on 90% to 94% of inCLASS dimensions. During data collection, they continued to code within one of the master code on 80% to 93% of dimensions on CLASS and inCLASS segments across five recalibration segments. During data collection, 20 percent of field observations were double coded. Inter-rater reliability, calculated using intraclass correlation coefficients, was acceptable for all four inCLASS domains: Teacher ICC was .93, Peer ICC was .91, Task ICC was .78, and Negative Engagement ICC was .76.

Live classroom observations began in the morning and lasted approximately four hours, with data collectors alternating observation and coding cycles for the CLASS and the inCLASS across two days. Data collectors observed the classroom using the standardized CLASS protocol for 15 minutes and coded for 10 minutes; they then alternated between observing two individual children for 10 minutes each using the standardized inCLASS protocol and coded the inCLASS for 5 minutes after each cycle, with the goal of completing a minimum of three CLASS cycles and three inCLASS cycles per child per visit. For inCLASS observations, observers focused on two target children in a class on the first observation day and the remaining two on the second day.

Measures

Quality of individual engagement. The *Individualized Classroom Assessment Scoring System* (inCLASS; Downer et al., 2010) is an observational measure of individual children's engagement with teachers, peers, and tasks in the classroom. The inCLASS is comprised of ten dimensions: Positive Engagement with Teachers, Teacher Communication; Peer Sociability, Peer Assertiveness, Peer Communication; Engagement with Tasks, Self-Reliance; Teacher Conflict, Peer Conflict, and Behavior Control (reverse coded). Each dimension is scored on a 1-7 scale from "low" to "high," where 1-2 represent low quality, 3-5 represent mid-range quality, and 6-7 represent high quality (except for Teacher Conflict, Peer Conflict, and Behavior Control reverse-coded, for which higher scores indicate lower quality). Children's scores were averaged across cycles. These dimensions have been shown through an initial exploratory factor analysis and validation studies (Bohlmann et al., 2019; Downer et al., 2010; Williford et al., 2013) to load onto four domains of engagement: Positive Engagement with Teachers, Positive Engagement with Peers, Positive Engagement with Tasks, and Negative Classroom Engagement. The inCLASS has evidence of criterion-related validity such that it is significantly associated with other established measures of children's skills and development. For instance, in the initial validation study (Downer et al., 2010), Positive Engagement with Teachers was significantly associated with teacher ratings of children's closeness with them (Student-Teacher Relationship Scale [STRS]), Positive Engagement with Peers was related to teacher ratings of children's social skills and assertiveness (Teacher Child Rating Scale [TCRS]), Positive Engagement with Tasks was linked with teacher ratings of children's skills across several domains (Academic Rating Scale [ARS]; TCRS), and Negative Engagement was associated with teacher ratings of conflict and behavior problems (STRS; TCRS).

Twenty percent of inCLASS observations were double-coded by independent observers to check reliability; out of 198 double-coded inCLASS cycles in Phase III, coders were within 1 of each other 90% of the time. Using the four inCLASS domains, Williford et al. (2013) used latent profile analysis (LPA) to identify three profiles of individual children's engagement: Positively Engaged, Typically Engaged, and Negatively Engaged. The present study sought to replicate these profiles (see Analytic Plan).

Classroom-level quality of teacher-child interactions. The *Classroom Assessment Scoring System* (CLASS; Pianta et al., 2008) measures teacher-child interactions at the classroom level and has been widely used with diverse populations (e.g., Downer et al., 2012). The measure is comprised of ten dimensions: Positive Climate, Negative Climate (reverse coded), Teacher Sensitivity, Regard for Student Perspectives; Concept Development, Quality of Feedback, Language Modeling; Behavior Management, Productivity, and Instructional Learning Formats. Each dimension is scored on a 1-7 scale, with the same quality benchmarks as the inCLASS. Children's scores were averaged across cycles. Twenty percent of CLASS observations were double-coded by independent observers to check reliability; out of 112 double-coded CLASS cycles in Phase III, coders were within 1 of each other 90% of the time.

Using the ten CLASS dimensions, Hamre et al. (2014) developed a bifactor model of teacher-child interaction quality composed of three uncorrelated domains: one global factor, Teacher Responsiveness; and two domain-specific factors, Cognitive Facilitation and Positive Management and Routines. This bifactor model fit the data better than the traditional three-factor model, two-factor model, or one-factor model (Hamre et al., 2014). This study sought to replicate this bifactor model of the CLASS (see Analytic Plan).

School readiness outcomes.

Language. Expressive vocabulary was measured by the Picture Vocabulary subtest of the Woodcock-Johnson Tests of Achievement, Third Edition (internal consistency reliability = .81; McGrew & Woodcock, 2001). In this assessment, children were required to name objects in a series of pictures. Receptive vocabulary was measured using the Peabody Picture Vocabulary Test, Third Edition (Dunn & Dunn, 1997), during which children were asked to point to one of four pictures corresponding to an orally presented word. Internal consistency reliability for this test was .95 and test-retest reliability was between .91 and .94. Criterion prediction validity with the Wechsler Intelligence Scale for Children was .90 in the initial validation study (Williams & Tang, 1997).

Literacy. Children were administered two subtests of the Test of Preschool Early Literacy (TOPEL): Phonological Awareness and Print Knowledge (Lonigan, Wagner, Torgesen, & Rashotte, 2007). Internal consistency reliability ranged from .86 to .96 and inter-rater reliability, from .96 to .98 in the initial validation study (Lonigan et al., 2007). In the Phonological Awareness subtest, children were asked to verbally manipulate the sounds in words; the Print Knowledge subtest required children to identify letters and connect letters to sounds.

Executive functioning. Inhibitory control was measured by the Pencil Tap subtest of the Preschool Self-Regulation Assessment (PSRA; Smith-Donald, Raver, Hayes, & Richardson, 2007). In this assessment, children were asked to tap their pencil once when the examiner tapped twice and twice when the examiner tapped once. Scores represent percentage of correct responses. This test shows acceptable concurrent and construct validity (Smith-Donald et al., 2007). Working memory was assessed using the Backward Digit Span subtest (Carlson, 2005). Children were asked to repeat increasingly longer sequences of digits backward. Scores represent the highest number of digits a child correctly repeated backward.

Covariates. Families provided information about their child's date of birth, sex, race/ethnicity, and primary language, as well as level of maternal education, household size, and annual family income. Annual income and household size were used to calculate the income-to-needs ratio, which is a family's income relative to the poverty level for their household size (U.S. DHHS, 2009). Preschool teachers provided information about themselves (e.g., years of education, years of experience at their current facility) and their classrooms, such as whether language and literacy was a focus in their curriculum, average income-to-needs ratio in the class, and classroom setting (public Head Start; non-public Head Start; other public school; or other public, non-profit, or for-profit agency). Teacher intervention condition in the NCRECE Professional Development Study was also recorded.

Analytic Plan

Data preparation. Missing data were as follows: Approximately 16% of the sample was missing fall and 7% were missing spring school readiness data; 18% were missing information about the length of time between assessments. All study children had valid information about observed classroom quality and individual engagement. Missing child, family, and classroom covariate data were generally minimal and are described as follows: child age (0%), sex (0%), race/ethnicity (2.25%), primary language (0.14%), years of maternal education (4.23%), family income-to-needs ratio (13.6%), household size (2.7%), teacher years of education (2.39%) and years at current facility (2.96%), curriculum focus on language and literacy (14.23%), percent low-income in class (.28%), classroom setting (4.37%), and teacher intervention condition (.42%). Patterns of missingness for study variables were explored using Little's test of missing completely at random (MCAR; Li, 2013; Little, 1988), which indicated that data were not MCAR. However, data can be considered missing at random (MAR) because missingness on outcome variables was significantly predicted by covariates (covariate-dependent missingness; Li, 2013). Thus, full information maximum likelihood (FIML) estimation was an appropriate way to address missing data (Allison, 2012; Enders & Bandalos, 2001; Li, 2013); this was conducted in Stata Version 14.2. To account for the nestedness of children within classrooms, all analyses used robust standard errors clustered by teacher.

Prior to testing the current study's research questions, it was necessary to fit a CLASS bifactor model. A confirmatory bifactor model was fit to the 10 CLASS dimensions using Mplus Version 7.4 (Muthén & Muthén, 1998-2015) in an effort to replicate the findings of Hamre and colleagues (2014) and identify uncorrelated factors of teacher-child interaction quality. Each dimension loaded onto a domain-general factor and one of two domain-specific factors (Reise, Moore, & Haviland, 2010). All ten CLASS dimensions comprised the domain-general Responsive Teaching factor. Cognitive Facilitation was composed of Concept Development, Quality of Feedback, and Language Modeling. Positive Management and Routines included Positive Climate, Negative Climate (reversed), Teacher Sensitivity, Regard for Student Perspectives, Behavior Management, Productivity, and Instructional Learning Formats. The bifactor model was estimated with the three factors (Responsive Teaching, Cognitive Facilitation, and Positive Management and Routines) constrained to be uncorrelated; one of the loadings in each of the factors was set to 1 (Hamre et al., 2014; Reise et al., 2010).

Research question 1. To characterize children's patterns of engagement in the preschool classroom, a latent profile analysis (LPA) was conducted using Mplus Version 7.4 (Muthén & Muthén, 1998-2015) to describe patterns of individual children's engagement across the four inCLASS domains. The multilevel nature of the data was taken into account by using the TYPE=COMPLEX MIXTURE command and clustering by teacher. Parameters were estimated using maximum likelihood estimation with robust standard errors (MLR); a sandwich estimator was used to ensure consistent parameter estimates (Carroll, Wang, Simpson, Stromberg, & Ruppert, 1998). The four inCLASS domains were allowed to covary with one another. Profile membership was regressed on a rich set of auxiliary variables comprised of the following: child- and family-level factors included child age, race, sex, primary language, years of maternal education, family income-to-needs ratio, and household size; teacher- and classroom-level factors included teacher years of education and experience, teacher intervention condition, percentage of the class in poverty, curriculum focus on language and literacy, and program type.

The expected three-profile solution (following Williford et al., 2013) was compared to one-, two-, and four-profile solutions to determine the best-fitting model. Fit was assessed using: (1) the Akaike Information Criterion (AIC; Akaike, 2011), Bayesian Information Criterion (BIC; Schwarz, 1978), sample-sized adjusted BIC (ABIC; Burnham & Anderson, 2004); (2) the Vuong–Lo–Mendell–Rubin likelihood ratio test (VLMR) and the Adjusted Lo–Mendell–Rubin likelihood ratio test (Adjusted LRT; Lo, Mendell, & Rubin, 2001); (3) entropy; and (4) the theoretical and practical applications of the profiles (Muthén, 2004; Williford et al., 2013). Relatively lower AIC, BIC, and ABIC values signify better fit. A significant *p*-value on the likelihood ratio tests indicates that a given model is preferable to a model with one fewer profile (Lo et al., 2001; Nylund, Asparouhov, & Muthén, 2007). Higher entropy values indicate a more accurate solution (Clark & Muthén, 2009; Hix-Small, Duncan, Duncan, & Okut, 2004). Importantly, however, unlike other types of models there are not "golden rules" in determining the best-fitting LPA model and it is important to consider theory and prior research, particularly if multiple profile solutions have similar fit (Marsh et al., 2009).

Once the final model was chosen using the above criteria, pairwise comparisons with Tukey's Honestly Significant Difference (HSD) post-hoc tests provided information about significant differences in inCLASS domains by children's most likely profile membership (Abdi & Williams, 2010). To explore how children's engagement patterns were associated with other characteristics of children, families, and classrooms, pairwise comparisons with Tukey's HSD post-hoc tests were also conducted with other key study variables (e.g., school readiness outcomes) and child-, family-, and classroom-level covariates.

Research question 2. To test how children's patterns of engagement were associated with skill gains across the preschool year, path analyses in an SEM framework were performed in Stata Version 14.2 to explore associations between engagement profile membership (Typically Engaged as omitted reference category) and gains in school readiness skills across the preschool year. All six school readiness outcomes (expressive and receptive language, phonological awareness, print knowledge, inhibitory control, and working memory) were included simultaneously in the model, which controlled for covariances among them.

Next, the three bifactor classroom quality predictors – Responsive Teaching, Cognitive Facilitation, and Positive Management and Routines – were included to determine whether associations between engagement patterns and skill gains remained when classroom quality was taken into account. When we attempted to run a final model including all predictors and outcomes, and accounting for covariances among all outcomes, the model would not converge. Thus, we separated the models into conceptual groupings, such that language outcomes (expressive and receptive, $r = 0.75$, $p < .001$), literacy outcomes (phonological awareness and print knowledge, $r = 0.53$, $p < .001$) and executive functioning outcomes (inhibitory control and working memory, $r = 0.42$, $p <$

.001) were regressed on all predictors in three separate models, each of which controlled for covariances among within-domain outcomes.

All models controlled for a rich set of covariates, comprised of the following: at the child level, age in months as of September 1, sex (male = 1), race/ethnicity (categorized as Black/African American, Hispanic/Latino, and other race/ethnicity, with White/Caucasian as the reference category), primary language (English = 1 compared to non-English), and continuous values of years of maternal education, income-to-needs ratio, household size, relevant fall entry scores, and time between fall and spring assessments. At the teacher/classroom level, we included teacher years of education and experience, teacher intervention condition, whether the curriculum was focused on language and literacy (yes = 1), average income-to-needs ratio in the class, and program setting (public Head Start; non-public Head Start; and other public school; with public, for-profit, or non-profit agency as the omitted reference category). All analyses included robust standard errors clustered by teacher to account for nestedness of children within classrooms.

Research question 3. To test whether and how relationships between individual engagement patterns and school readiness gains depended on the quality of teacher-child interactions in the classroom, we created six interaction terms by multiplying each of the two included engagement profile dummy codes and each of the three classroom quality factors (positively engaged (PE) x responsive teaching (RT), PE x cognitive facilitation (CF), PE x positive management and routines (PMR), negatively engaged (NE) x RT, NE x CF, NE x PMR). Aside from the inclusion of the six interaction terms, procedures for conducting path analyses were the same as for research question 2, with models that (1) regressed outcomes on all predictors and covariates, (2) were separated by outcome domains (language, literacy, and executive functioning), and (3) accounted for covariances between outcomes in each model.

When predictors were significantly associated with school readiness gains, effect sizes were calculated by multiplying the unstandardized coefficient and the standard deviation of the predictor, then dividing by the standard deviation of the relevant outcome (Gutman, Sameroff, & Cole, 2003; Mashburn, Justice, Downer, & Pianta, 2009; Williford et al., 2013).

Results

Data Preparation: Replicating a Bifactor Model of Teacher-Child Interactions

Prior to establishing profiles of engagement and examining predictive models, we needed to replicate the bifactor model of the CLASS (Hamre et al., 2014). Results of this bifactor model revealed that some fit indices were acceptable, including the standardized root mean square residual (SRMR = .036). Others, such as the comparative fit index (CFI = .937) fell below typical standards but approximated published fit for the same bifactor model in another sample (CFI = .94; Hamre et al., 2014). The root mean square error of approximation (RMSEA = .155) exceeded acceptable standards. This bifactor model fit the data better than both a global, one-factor model (CFI = .715, SRMR = .096, RMSEA = .279) and the original three-factor model comprised of Instructional Support, Emotional Support, and Classroom Organization (CFI = .845, SRMR = .077, RMSEA = .215). The factors in the original model were also highly correlated with one another: Instructional Support and Emotional Support (r = .787), Instructional Support and Classroom Organization (r = .826), and Emotional Support and Classroom Organization (r = .994).

Correlations among factors in the bifactor model were much smaller: Responsive Teaching and Cognitive Facilitation ($r = .062$), Responsive Teaching and Positive Management and Routines ($r = .067$), and Cognitive Facilitation and Positive Management and Routines ($r = -.153$). Given that two of the three fit indices, along with factor loadings, in the bifactor model aligned with findings by Hamre et al. (2014), and the fact that this model was preferable to alternative models in terms of both model fit and correlations among component factors, the bifactor model was retained and used in analyses. Standardized and unstandardized loadings and fit statistics are provided in Table 1.

Patterns of Children's Engagement in the Preschool Classroom

Taken together, LPA model fit statistics, along with considering theory, indicated that a three-profile solution best fit the data compared to one-, two-, and four-profile solutions (see Table 2). Specifically, the AIC, BIC, and ABIC decreased until the three-profile solution and then increased in the four-profile solution. Although the four-profile solution was also acceptable, the three-profile solution had preferable information criterion statistics. Because of this and to align with prior work by Williford et al. (2013), the three-profile solution was selected and used in models to answer subsequent research questions (Marsh et al., 2009).

The majority of the sample (60%; $n = 425$) in this study exhibited a pattern of low-to-moderate engagement with teachers and peers, moderate engagement with tasks, and low negative engagement; this profile was labeled "Typically Engaged." Almost 32% ($n = 227$) of the sample comprised the "Positively Engaged" profile, which had significantly higher engagement with teachers ($t = 8.57$, $p < .001$), peers, ($t = 26.51$, $p = <$

.001) and tasks (t = 19.32, p = < .001) compared to Typically Engaged peers. A subset (8.2%; n = 58) of children in this sample displayed lower positive engagement with tasks (t = - 7.24, p < .001) and higher negative engagement (t = 27.02, p < .001) compared to Typically Engaged peers. This group, classified as "Negatively Engaged" in the classroom, also had significantly lower positive engagement with teachers (t = -2.61, p = .03), peers (t = - 12.50, p < .001), and tasks (t = - 17.68, p < .001) and significantly higher negative engagement (t = 24.43, p < .001) than peers characterized as Positively Engaged. Figure 1 depicts a graphical representation of the three profiles, and Table 3 provides means and standard deviations of the four inCLASS domains and key study variables for the full sample and each profile separately.

The average probabilities for children's most likely class membership indicated adequate fit between study children and the profiles. Children characterized as Typically Engaged had a 91.5% probability of being assigned to that profile, with much smaller probabilities of being assigned to the Positively Engaged (7.6%) or Negatively Engaged (1.0%) profiles. Children classified as Positively Engaged had an 85.0% chance of being assigned to that profile, and only a 14.5% and a 0.5% chance of being assigned to Typically or Negatively Engaged profiles, respectively. Children classified as Negatively Engaged had an 86.1% probability of being assigned to that profile, and a much smaller probability of being assigned to either the Positively (2.8%) or Typically Engaged (11.1%) profiles. Children's most likely class membership was dummy-coded into two variables, Positively Engaged and Negatively Engaged, with Typically Engaged, the largest profile, serving as the omitted reference group in all subsequent analyses.

Engagement patterns and other characteristics. Table 4 presents descriptive information for child-, family-, and teacher/classroom-level covariates used in the study, for the full sample and for each profile separately. Significant differences in these variables across profiles are denoted in the table with superscripts. Compared to peers in the Typically Engaged group, children characterized as Positively Engaged were more likely to be older ($t = 2.47$, $p = .04$) and to primarily speak English ($\chi^2 = 19.33$, $p < .001$). They also were more likely to be advantaged in terms of family and classroom characteristics. Specifically, children classified as Positively Engaged tended to have more years of maternal education ($t = 4.75$, $p < .001$), a higher income-to-needs ratio ($t = 5.53$, $p < .001$), a smaller household size ($t = -4.64$, $p < .001$), a higher classroom average income-to-needs ratio ($t = 5.00$, $p < .001$), and more educated teachers ($t = 4.72$, $p < .001$) with more years of experience ($t = 3.52$, $p = .001$) than Typically Engaged Peers. They were also more likely to be in a public school Head Start setting ($t = 6.12$, $p = .01$).

There were fewer differences between children classified as Typically Engaged and Negatively Engaged: the latter were more likely to be male ($\chi^2 = 8.38$, $p = .004$) and to have less educated teachers ($t = -2.86$, $p = .01$). Notably, children classified as Negatively Engaged had many demographic characteristics that were similar (i.e., statistically non-significant) to those of Positively Engaged peers, including the following: race/ethnicity, primary language, years of maternal education, household size, fall-spring assessment window, teacher years of experience, curriculum focus on language and literacy, and program setting type.

In addition to descriptive information about inCLASS scores across profiles, Table 3 provides information about the quality of teacher-child interactions (CLASS) that children in different profiles experienced, as well as their fall and spring school readiness scores. There was a mix of findings across engagement profiles with regard to teacher-child interaction quality. Classrooms of Positively Engaged children were characterized by significantly higher Responsive Teaching than those of Typically (t = 4.10, p < .001) and Negatively Engaged peers (t = 3.85, p < .001). These children also tended to experience higher Cognitive Facilitation (t = 3.55, p = .001) and lower Positive Management and Routines (t = -6.50, p < .001) than Typically Engaged peers. Classrooms of children classified as Negatively Engaged were characterized by significantly lower Responsive Teaching than those of Positively Engaged peers (t = 3.85, p < .001) and lower Positive Management and Routines than those of Typically Engaged peers (t = -2.47, p = .04). Children classified as Positively and Negatively Engaged experienced similar (i.e., statistically non-significant difference) Cognitive Facilitation and Positive Management and Routines; the same was true for children characterized as Negatively and Typically Engaged for Responsive Teaching and Cognitive Facilitation.

Associations between Engagement Patterns and Skill Gains

Table 5 presents results of models regressing school readiness outcomes on engagement profiles, teacher-child interaction quality variables, and study covariates. The left column under each of the six outcomes displays standardized coefficients and robust standard errors when only the profiles were included, and the right column under each outcome provides this information when the profiles and factors of teacher-child interaction quality were included in models.

Path analysis results revealed that children's patterns of engagement in the classroom were associated with school readiness gains across the year, controlling for covariates. Children classified as Negatively Engaged made significantly smaller gains across the year than Typically Engaged peers in receptive language ($\beta = -.06$, $p = .007$, effect size = .06), print knowledge ($\beta = -.06$, $p = .02$, effect size = .05), inhibitory control ($\beta = -.07$, $p = .02$, effect size = .07), and working memory ($\beta = -.05$, $p = .04$, effect size = .05). There were no significant differences in skill gains between children characterized as Positively Engaged and those classified as Typically Engaged.

When the quality of teacher-child interactions in the classroom was taken into account, individual children's patterns of engagement maintained associations with skill gains in language and literacy. Specifically, children classified as Negatively Engaged continued to make smaller gains than Typically Engaged peers across the preschool year in receptive language ($\beta = -.06$, $p = .005$, effect size = .06) and print knowledge ($\beta = -.05$, $p = .02$, effect size = .05). Relationships between profile membership and executive functioning skills became statistically non-significant when classroom quality was included. This may be because being in a classroom characterized by higher levels of Positive Management and Routines was associated with more skill gains across the year in inhibitory control ($\beta = .09$, $p = .03$, effect size = .09) and working memory ($\beta = .09$, $p = .006$, effect size = .09) for all children, regardless of their engagement pattern.

Interactive Contributions of Engagement Patterns and Teacher-Child Interaction Quality

Finally, this study explored whether associations between children's patterns of engagement in the preschool classroom depended on the quality of specific types of teacher-child interactions. Results revealed that this was the case for Negatively Engaged children, though only for some school readiness outcomes and in the context of one component of teacher-child interactions. Specifically, in classrooms characterized by higher levels of Positive Management and Routines, children classified as Negatively Engaged tended to perform as well as or better than Typically Engaged peers in expressive language ($\beta = .04$, $p = .04$, effect size = .04), phonological awareness ($\beta = .07$, $p = .009$, effect size = .07), and inhibitory control ($\beta = .09$, $p = .005$, effect size = .09). Figures 2, 3, and 4 depict these interactions graphically. There were no significant interactions comparing the Positively Engaged group to the Typically Engaged group, nor for any interactions involving Responsive Teaching or Cognitive Facilitation interactions.

Discussion

The quality of classroom processes, such as children's engagement and teacher-child interactions, has been independently related to young children's development (Sabol et al., 2018) in skills that set the stage for later school success (Duncan et al., 2007). Recent evidence indicates that these processes interact (Williford et al., 2013); however, there has not yet been specificity around how targeted, high-quality interactions in the preschool classroom might act as promotive or protective factors influencing the development of children from a high-risk sample. The current study sought to address this limitation. Children in this sample were diverse in terms of race, ethnicity, and primary language; they also came from relatively low-income backgrounds. It is important to examine how specific classroom processes are related to changes in school readiness skills for these children, because evidence suggests that high-quality early childhood experiences are especially beneficial for children who are at risk due to socioeconomic, minority, or language minority status (Bloom & Weiland, 2015; Burchinal et al., 2010, 2011; Cooper & Lanza, 2014).

We first identified patterns of children's positive and negative engagement across the teachers, peers, and tasks in their classrooms and how these patterns related to characteristics of children, families, and classrooms. Findings aligned with expectations, replicating three profiles found by Williford and colleagues (2013): Positively Engaged, Typically Engaged, and Negatively Engaged. These profiles were generally related to child demographics in expected patterns. Specifically, children classified as Positively Engaged tended to be more demographically advantaged (e.g., older, higher income-to-needs ratio, higher teacher education) than peers; however, children classified as Negatively Engaged were largely similar in demographics to their Typically Engaged peers.

We then examined the extent to which these patterns were associated with school readiness skill gains across the preschool year, and whether these associations were maintained when the quality of teacher-child interactions in the classroom was considered. Engagement patterns were indeed related to gains during preschool, and these associations were largely maintained with the inclusion of teacher-child interaction quality. As hypothesized, children characterized as Negatively Engaged made smaller gains than peers across the year; unexpectedly, children classified as Positively Engaged did not make larger gains than peers. Finally, we examined whether associations between engagement patterns and readiness gains depended on the quality of teacher-child interactions in the classroom. In this sample, higher levels of Positive Management and Routines bolstered gains in executive functioning skills for all children but facilitated skills across domains significantly more for children classified as Negatively Engaged. Thus, this specific type of teacher-child interaction is promotive for all children *and* protective for those at risk due to the ways in which they engage in the classroom. Below, we discuss key results in the context of recent literature, as well as implications for understanding and improving preschool classroom experiences, particularly for high-need children.

Characterizing Children's Engagement in the Early Childhood Classroom

Children in this sample displayed distinct patterns, or profiles, of engagement with the teachers, peers, and tasks in their preschool classroom. We took a person-centered approach to further explore how children's engagement manifests *across* different contexts rather than examining each variable singly. These profiles are informative because patterns clustered together in ways that help to illuminate how children interact in the preschool classroom and where teachers might intervene specifically to promote engagement. The majority of children displayed a pattern suggesting that they only occasionally interact positively with teachers and peers throughout the school day and that they somewhat more regularly sustain active engagement with tasks. Although children classified as Positively Engaged exhibited higher positive engagement with teachers, peers, and tasks, their positive interactions with individuals – especially teachers – were by no means consistent. A practical implication of this finding is that *all* children could benefit from more frequent and higher quality positive interactions with teachers and peers, which have been shown to support children's positive development (Fantuzzo & McWayne, 2002; O'Connor & McCartney, 2007).

Regarding negative engagement in the classroom, even children classified as Negatively Engaged displayed relatively low levels compared to the full range of the inCLASS. It may be that because serious negative behaviors (e.g., physical aggression) are often low frequency, they were not captured by this measure *or* that these children do not display these serious negative interactions in the classroom. Regardless, even relatively small elevations in negative engagement are meaningful because this was associated with lower skill gains in multiple areas. Moreover, because this sample excluded children who had an IEP or IFSP, these findings provide evidence that there is a group of typically developing children who display a pattern of engagement that puts them at risk of falling behind in skills. Children classified in this profile displayed behaviors that are arguably normative for preschool-aged children (e.g., confrontation, attention-seeking, difficulty controlling actions) but they displayed them to a significantly greater extent than peers, to the detriment of their skill development. Important to note is that it was not the case that being classified as Negatively Engaged was merely a proxy for other indicators of risk (e.g., low socioeconomic status). Children classified as Typically and Negatively Engaged had similar demographic characteristics; in some cases, the latter group had characteristics that were more aligned with those of Positively Engaged peers, who tended to be more advantaged demographically in some ways. This indicates that children classified as Negatively Engaged are at risk of falling behind academically for reasons – above and beyond demographic characteristics – that have to do explicitly with the ways in which they interact in the classroom.

This finding has implications for teacher training and professional development, because these at-risk children are likely to have more positive school experiences and to make more skill gains if teachers can identify and intervene with them early on. This would require teachers to have both a nuanced understanding of how children engage positively and negatively with teachers, peers, and tasks in the preschool classroom and the tools to be able to observe children's interactions objectively. These teacher skills are likely best taught with explicit instruction and ongoing coaching (e.g., Downer et al., in press). Next, teachers would benefit from instruction and coaching around developing not only academic skills but also positive engagement that enables children to be more available for learning and skill development. This might be accomplished by helping teachers to provide specific, targeted interactions that meet the needs of individual students (Hemmeter & Fox, 2009; Landry, Anthony, Swank, & Monseque-Bailey, 2009; Pianta, Burchinal et al., 2014; Pianta, DeCoster et al., 2014; Webster-Stratton & Reid, 2004).

Independent Contributions of Engagement and Teacher-Child Interaction Quality

This study found significant independent contributions of individual children's engagement and classroom-level teacher-child interactions; in both cases, findings were mixed in terms of how they replicated prior work. Children's engagement patterns were associated with skill gains across the preschool year such that children classified as Negatively Engaged made smaller gains in receptive language, print knowledge, inhibitory control, and working memory. These findings are relatively consistent with prior work on children who tend to have lower levels of positive engagement and higher levels of negative engagement (Sabol et al., 2018; Williford et al., 2013). Associations between the Negatively Engaged profile and language and literacy skill gains were maintained when classroom quality was taken into account, indicating that it is not merely that exposure to a certain degree of classroom quality leads to skill gains; rather, it is also important to examine how the individual child engages in the classroom. It may be that children classified as Negatively Engaged, who tend to be less available for learning in the classroom due to their engagement patterns, are less exposed to instruction or enriching interactions and thus have fewer opportunities to gain skills such as receptive language and print knowledge that are often explicitly instruction-driven. Effect sizes were small but comparable to previous findings (Williford et al., 2013); this may be related to controls placed on the models, such as including a rich set of covariates and accounting for covariances among outcomes.

Contrary to expectations and prior literature (Williford et al., 2013), children classified as Positively Engaged generally made skill gains comparable to their Typically Engaged peers. This is not to say that they had similar skill levels, because children classified as Positively Engaged did have significantly higher skills both in the fall and in the spring; rather, the trajectory of skill gains was similar across these two groups. Based on this finding, it was hypothesized that perhaps engagement's link to school readiness skill development for children with relatively high positive engagement across teachers, peers, and tasks is conditional upon the provision of certain types of classroom experiences or a certain level of quality in classroom experiences. However, this hypothesis was not borne out in results testing interaction effects.

Another unexpected finding was how classroom-level teacher-child interactions were associated with skill gains for all children. Similar to prior work (Hamre et al., 2014), the quality of specific interactions characterized by positive behavior management strategies and consistent routines was promotive for the executive functioning skills of all children. However, associations were not found for general Teacher Responsiveness, which was hypothesized to relate to skill gains across domains, or for Cognitive Facilitation, which was expected to facilitate children's language and literacy skills (Hamre et al., 2014). Importantly, Hamre and colleagues solely examined the relationship between classroom-level teacher-child interaction quality and skills without taking into account individual children's engagement. When multiple aspects of engagement and teacher-child interaction quality are considered together, classroom-level teacher-child interaction quality has been shown to become non-significant across multiple outcomes (Sabol et al., 2018). However, it is important to understand not only how these factors contribute to skill gains when each is considered independently, but also how they interact given the dynamic nature of classroom processes.

Interactive Contributions of Engagement and Teacher-Child Interaction Quality

Examining individual children's patterns of engagement and the quality of classroom-level teacher-child interactions in tandem offered the opportunity to better understand the confluence of factors that influence preschool children's school readiness skill development. One expected pattern that emerged in this study was that these at-risk children benefited from high-quality, specific interactions aimed at promoting positive behavior management. It was hypothesized that this type of interaction would be protective for the executive functioning skills of children classified as Negatively Engaged. Importantly, this type of interaction was also protective for the language and literacy skill development of these children. Children classified as Negatively Engaged tend to be the least available for learning in the classroom, given their relatively lower positive engagement and their higher negative engagement, which is characterized at least in part by having more difficulty controlling their actions and matching classroom expectations. Thus, it makes sense that interactions that scaffold children's ability to manage their behaviors would better enable them to participate in the learning activities and interactions that lead to skill development. This may explain why Positive Management and Routines was related to not only within-domain executive functioning skills but also more explicitly instruction-driven skills in domains of language and literacy. The hypothesis that Cognitive Facilitation would promote these latter skills, especially for children classified as Negatively Engaged, was not supported. It may be that a certain threshold of quality – especially around teachers' abilities to use advanced language and ask children to explain their thinking to facilitate higher-level reasoning – is necessary to change the trajectory of more explicitly instruction-driven academic skills (Hatfield, Burchinal, Pianta, & Sideris, 2016; Zaslow et al., 2010).

High-quality, specific teacher-child interactions were not especially promotive for children who tended to be Positively Engaged. Williford et al. (2013) also found that in classrooms characterized by high-quality teacher-child interactions, it was the Typically Engaged, rather than the Positively Engaged, group that made relatively more skill gains. It may be that classrooms in this sample did not meet quality thresholds necessary to continue the trajectory of skill gains for children who tend to be more available for learning – and who tend to enter preschool with higher skills – particularly in terms of instructionally supportive interactions that facilitate cognition, such as modeling children's language, providing quality feedback, and helping children to gain in-depth understanding of concepts. Previous literature indicates that teachers' instructional support tends to be lower than their emotional support and classroom organization

(Mashburn et al., 2008) and that teacher-child interactions are more strongly associated with children's school readiness outcomes when classroom quality is higher (Burchinal et al., 2010; Hatfield et al., 2016; Zaslow et al., 2010).

Both this study and prior work (Williford et al., 2013) have found that the larger context of the quality of teacher-child interactions in the classroom has the ability to reduce skill gaps between children who are at risk of falling behind due to the ways in which they engage in the classroom and those who are more available for learning in the preschool classroom. This study extends prior research by providing more precision around this finding. Specifically, high-quality interactions that help children to manage their behavior and match classroom expectations enable children who struggle with these abilities to gain more from learning experiences in ways that promote their school readiness skills across domains. This has important implications for teacher training. Not only do teachers need to be able to identify and intervene with children who display at-risk patterns of engagement in the classroom, but they also might focus on increasing the frequency and effectiveness with which they use positive behavior management strategies and establish consistent routines in the classroom.

Limitations and Directions for Future Research

Despite its contributions, this study is characterized by a number of sampling, conceptual, and methodological limitations that provide directions for future research. Though the present study's sample was racially and ethnically diverse, children were mainly from low-income households. It is important to understand the experiences of this at-risk demographic, but this limits generalizability to socioeconomically disadvantaged children. Future research might leverage a mixed-income sample to understand the role of proximal classroom processes in skill development more broadly. This study excludes children with an IEP or IFSP, who arguably are at risk of displaying negative patterns of engagement in the classroom. Thus, findings may actually underestimate the proportion of children who might be classified as Negatively Engaged. Finally, this study and much of the literature in this area is U.S.-centric, although the CLASS has been used internationally. Future international research using the inCLASS would help to characterize children's engagement in the preschool classroom cross-culturally.

Conceptually, school readiness outcomes in this study were restricted to directly assessed language, literacy, and executive functioning skills because of data set limitations. Future research might explore how proximal processes interact to promote other foundational academic skills, such as mathematics, as well as social, emotional, and behavioral outcomes. It is possible that general Responsive Teaching, which was not significantly associated with skill gains in this study, would be related to gains in children's social, emotional, and behavioral development. This study controlled for several child-, family-, and classroom-level covariates but it is possible that omitted variable bias affected results. For instance, this study did not have access to information about children's experiences prior to preschool (e.g., child care) that may have impacted both their engagement and skill development. Future research should examine the role of family factors present prior to preschool entry in promoting or hindering preschoolers' school readiness skill development.

Predictors (i.e., engagement and teacher-child interactions) were operationalized to align with previous research, which led to limitations in the questions this study could answer. Specifically, children's negative engagement as measured by the inCLASS is aggregated across teacher, peer, and task contexts; this approach limits our ability to unpack whether conflict, resisting connections with others, and difficulty with behavior control play different roles in hindering skill development. The person-centered approach taken in this study identified how children's engagement manifests in the preschool classroom across contexts and how those patterns relate to children's skills; however, it did not allow for an exploration of heterogeneity within profiles or an understanding of the specific aspects of engagement associated with skill development (though this has been examined in prior research; see Sabol et al., 2018).

Classroom quality was operationalized as the quality of teacher-child interactions in the classroom generally, as measured by the CLASS. Content of instruction and domain-specific quality of instructional content delivery are also important predictors of school readiness skills that might be more closely aligned with skills in particular domains (Zaslow et al., 2010). Future research might explore the role of different types of classroom quality using domain-specific measures such as the Early Language and Literacy Classroom Observation (ELLCO; Smith & Dickinson, 2002) or the Classroom Observation of Early Mathematics: Environment and Teaching (COEMET; Sarama & Clements, 2007). Finally, this study found unidirectional associations between proximal processes and skill gains in preschool. However, it is possible that there is a bidirectional relationship between children's engagement and skill development that was not tested. Future research with multiple time points across the year(s) might test cross-lagged models to further explore this potential for bidirectionality (Viljaranta, Lerkkanen, Poikkeus, Aunola, & Nurmi, 2009).

Methodologically, this was a correlational study that cannot make causal claims. As in previous studies (Sabol et al., 2018; Williford et al., 2013), observations of children's engagement (inCLASS) and teacher-child interactions (CLASS) were completed across one day by a single set of observers. It is possible that this protocol did not provide the most reliable estimate of children's engagement and teacher-child interaction quality. However, prior research indicates that the inCLASS captures children's engagement in a variety of activity settings across cycles, even when observations take place only within a single day (Vitiello et al., 2012). There may have been some contamination between the two measures, such that some of the correlation between the two may be an artifact of observation procedures rather than a true relationship between children's engagement and teacher-child interaction quality. However, we would theoretically expect that these processes would be significantly related to one another regardless of the observation protocol. Moreover, these measures have standardized training and field data collection protocols and ongoing recalibration to maintain reliability to the measures.

Though not a limitation, there were several hypothesized relationships that were not found to be significant in this study. Specifically, unlike in prior research, children classified as Positively Engaged did not make larger gains than peers across the preschool year (Williford et al., 2013) and interactions characterized by Responsive Teaching and Cognitive Facilitation were not significantly related to children's skill development (Hamre et al., 2014). Future studies might unpack these findings by using a more fine-grained analysis of the concepts teachers impart (e.g., basic skills versus analysis-inference skills; Downer, Rimm-Kaufman, & Pianta, 2007), the dosage and quality of specific instructional activities, the extent to which these activities align with individual children's skill levels, and for whom and to what degree these activities promote gains in school readiness skills.

Conclusion

This study builds upon previous work (Hamre et al., 2014; Sabol et al., 2018; Williford et al., 2013) and provides additional evidence that dynamic processes in the preschool classroom, such as individual children's patterns of engagement and the quality of teacher-child interactions, are associated with the development of foundational school readiness skills. It identified patterns of children's engagement that were associated with both characteristics of children and their skill development. Children who tended to have lower positive engagement and relatively higher negative engagement than peers made smaller gains across the year. However, classrooms characterized by consistent routines and positive, proactive behavior management strategies were protective for the skill development of these children. Supporting teachers' ability to identify children at risk of negatively engaging in the classroom and to consistently provide high-quality examples of this specific type of interaction may help to more equitably develop children's school readiness skills in preschool and thus reduce the achievement gap at kindergarten entry.

References

Abdi, H., & Williams, L. J. (2010). Tukey's honestly significant difference (HSD) test. In N. Salkind (Ed.), *Encyclopedia of Research Design* (pp. 1-5). Thousand Oaks, CA: Sage.

Akaike, H. (2011). Akaike's information criterion. In M. Lovric (Ed.), *International Encyclopedia of Statistical Science* (pp. 25-25). Berlin, Germany: Springer-Verlag Berlin Heidelberg.

Allison, P. D. (2012, April). Handling missing data by maximum likelihood. In *SAS global forum* (Vol. 23). Haverford, PA, USA: Statistical Horizons.

Arnold, D. H., Kupersmidt, J. B., Voegler-Lee, M. E., & Marshall, N. A. (2012). The association between preschool children's social functioning and their emergent academic skills. *Early Childhood Research Quarterly*, *27*, 376-386. doi: 10.1016/j.ecresq.2011.12.009.

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

Birch, S. H., & Ladd, G. W. (1997). The teacher-child relationship and children's early school adjustment. *Journal of School Psychology*, *35*, 61-79. doi: 10.1016/S0022-4405(96)00029-5.

Bloom, H. S., & Weiland, C. (2015). *Quantifying variation in Head Start effects on young children's cognitive and socio-emotional skills using data from the National Head Start Impact Study*. Washington, DC: MDRC.

Bohlmann, N. L., & Downer, J. T. (2016). Self-regulation and task engagement as predictors of emergent language and literacy skills. *Early Education and Development*, *27*, 18-37. doi: 10.1080/10409289.2015.1046784.

Bohlmann, N. L., Downer, J. T., Williford, A. P., Maier, M. F., Booren, L. M., & Howes, C. (2019). Observing children's engagement: Examining factorial validity of the inCLASS across demographic groups. *Journal of Applied Developmental Psychology*, *60*, 166-176. doi: 10.1016/j.appdev.2018.08.007.

Booren, L. M., Downer, J. T., & Vitiello, V. E. (2012). Observations of children's interactions with teachers, peers, and tasks across preschool classroom activity settings. *Early Education & Development*, *23*, 517-538. doi: 10.1080/10409289.2010.548767.

Bronfenbrenner, U., & Morris, P. A. (1998). The ecology of developmental processes. In W. Damon & R. M. Lerner (Eds.), *Handbook of child psychology: Theoretical models of human development* (5th ed., Vol. 1, pp. 993-1028). Hoboken, NJ: John Wiley & Sons Inc.

Buckrop, J., Roberts, A., & LoCasale-Crouch, J. (2016). Children’s preschool classroom experiences and associations with early elementary special education referral. *Early Childhood Research Quarterly*, *36*, 452-461. doi: 10.1016/j.ecresq.2016.01.016.

Bull, R., Espy, K. A., & Wiebe, S. A. (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. doi: 10.1080/87565640801982312.

Bulotsky-Shearer, R. J., Bell, E. R., Romero, S. L., & Carter, T. M. (2012). Preschool interactive peer play mediates problem behavior and learning for low-income children. *Journal of Applied Developmental Psychology*, *33*, 53-65. doi: 10.1016/j.appdev.2011.09.003.

Bulotsky-Shearer, R. J., Fernandez, V., Dominguez, X., & Rouse, H. L. (2011). Behavior problems in learning activities and social interactions in Head Start classrooms and early reading, mathematics, and approaches to learning. *School Psychology Review*, *40*, 39-56.

Burchinal, M., Howes, C., Pianta, R., Bryant, D., Early, D., Clifford, R., & Barbarin, O. (2008). Predicting child outcomes at the end of kindergarten from the quality of pre-kindergarten teacher-child interactions and instruction. *Applied Developmental Science, 12*, 140-153. doi: 10.1080/10888690802199418.

Burchinal, M., McCartney, K., Steinberg, L., Crosnoe, R., Friedman, S. L., McLoyd, V., & Pianta, R. (2011). Examining the Black–White achievement gap among low-income children using the NICHD study of early child care and youth development. *Child Development*, *82*, 1404-1420. doi: 10.1111/j.1467-8624.2011.01620.x.

Burchinal, M., Vandergrift, N., Pianta, R., & Mashburn, A. (2010). Threshold analysis of association between child care quality and child outcomes for low-income children in pre-kindergarten programs. *Early Childhood Research Quarterly*, *25*, 166-176. doi: 10.1016/j.ecresq.2009.10.004.

Burnham, K. P., & Anderson, D. R. (2004). Multimodel inference: Understanding AIC and BIC in model selection. *Sociological Methods and Research*, 33, 261–304.

doi: 10.1177/0049124104268644.

Carlson, S. M. (2005). Developmentally sensitive measures of executive function in preschool children. *Developmental Neuropsychology*, *28*, 595-616. doi: 10.1207/s15326942dn2802_3.

Carroll, R. J., Wang, S., Simpson, D. G., Stromberg, A. J., & Ruppert, D. (1998). The sandwich (robust covariance matrix) estimator. *Unpublished manuscript*.

Chen, F. F., Hayes, A., Carver, C. S., Laurenceau, J. P., & Zhang, Z. (2012). Modeling general and specific variance in multifaceted constructs: A comparison of the bifactor model to other approaches. *Journal of Personality*, *80*, 219-251. doi: 10.1111/j.1467-6494.2011.00739.x.

Chien, N. C., Howes, C., Burchinal, M., Pianta, R. C., Ritchie, S., Bryant, D. M., ... & Barbarin, O. A. (2010). Children's classroom engagement and school readiness gains in prekindergarten. *Child Development*, *81*, 1534-1549. doi: 10.1111/j.1467-8624.2010.01490.x.

Clark, S. L., & Muthén, B. (2009). Relating latent class analysis results to variables not included in the analysis. Unpublished paper available at www.statmodel.com/download/relatinglca.pdf.

Cooper, B. R., & Lanza, S. T. (2014). Who benefits most from Head Start? Using latent class moderation to examine differential treatment effects. *Child Development*, *85*, 2317-2338. doi: 10.1111/cdev.12278.

Curby, T. W., Rimm-Kaufman, S. E., & Ponitz, C. C. (2009). Teacher–child interactions and children's achievement trajectories across kindergarten and first grade. *Journal of Educational Psychology*, *101*, 912-925. doi: 10.1037/a0016647.

Downer, J. T., Booren, L. M., Lima, O. K., Luckner, A. E. & Pianta, R. C. (2010). The Individualized Classroom Assessment Scoring System (inCLASS): Preliminary reliability and validity of a system for observing preschoolers' competence in classroom interactions. *Early Childhood Research Quarterly, 25*, 1–16. doi: 10.1016/j.ecresq.2009.08.004.

Downer, J. T., López, M. L., Grimm, K. J., Hamagami, A., Pianta, R. C., & Howes, C. (2012). Observations of teacher–child interactions in classrooms serving Latinos and dual language learners: Applicability of the Classroom Assessment Scoring System in diverse settings. *Early Childhood Research Quarterly*, *27*, 21-32. doi: 10.1016/j.ecresq.2011.07.005.

Downer, J. T., Pianta, R. C., Burchinal, M., Field, S., Hamre, B. K., LoCasale-Crouch, J., & Scott-Little, C. (2013). Coaching and coursework focused on teacher-child interactions during language/literacy instruction: Effects on teacher outcomes and children's classroom engagement. *Unpublished paper, University of Virginia*.

Downer, J. T., Rimm-Kaufman, S. E., & Pianta, R. C. (2007). How do classroom conditions and children's risk for school problems contribute to children's behavioral engagement in learning?. *School Psychology Review*, *36*, 413-432.

Downer, J.T., Williford, A., Bulotsky-Shearer, R., Vitiello, V., Bouza, J., Reilly, S., & Lhospital, (2018). Using data-driven, video-based early childhood consultation with teachers to reduce children's challenging behaviors and improve engagement in preschool classrooms. *School Mental Health, 10,* 226-242.

Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., ... & Japel, C. (2007). School readiness and later achievement. *Developmental*

Psychology, *43*, 1428-1446. doi: 10.1037/0012-1649.43.6.1428.

Dunn, L. M., & Dunn, L. M. (1997). *PPVT-III: Peabody picture vocabulary test*. Circle Pines, MN: American Guidance Service.

Enders, C. K., & Bandalos, D. L. (2001). The relative performance of full information maximum likelihood estimation for missing data in structural equation models. *Structural equation modeling*, *8*, 430-457. doi: 10.1207/S15328007SEM0803_5.

Fantuzzo, J., Bulotsky-Shearer, R., McDermott, P., McWayne, C., Frye, D., & Perlman, S. (2007). Investigation of dimensions of social-emotional classroom behavior and school readiness for low-income urban preschool children. *School Psychology Review, 36*, 44-62.

Fantuzzo, J., & McWayne, C. (2002). The relationship between peer-play interactions in the family context and dimensions of school readiness for low-income preschool children. *Journal of Educational Psychology*, *94*, 79-87. doi: 10.1037/0022-0663.94.1.79.

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research*, *74*, 59-109. doi: 10.3102/00346543074001059.

Greenwood, C. R., Bradfield, T., Kaminski, R., Linas, M., Carta, J. J., & Nylander, D. (2011). The response to intervention (RTI) approach in early childhood. *Focus on Exceptional Children*, *43*, 1-22.

Gutman, L. M., Sameroff, A. J., & Cole, R. (2003). Academic growth curve trajectories from 1st grade to 12th grade: Effects of multiple social risk factors and preschool child factors. *Developmental Psychology*, *39*, 777-790. doi: 10.1037/0012-1649.39.4.777.

Hamre, B., Hatfield, B., Pianta, R., & Jamil, F. (2014). Evidence for general and domain-specific elements of teacher–child interactions: Associations with preschool children's development. *Child Development*, *85*, 1257-1274. doi: 10.1111/cdev.12184.

Hamre, B. K., & Pianta, R. C. (2005). Can instructional and emotional support in the first-grade classroom make a difference for children at risk of school failure?. *Child Development*, *76*, 949-967. doi: 10.1111/j.1467-8624.2005.00889.x.

Hamre, B. K., & Pianta, R. C. (2007). Learning opportunities in preschool and early elementary classrooms. In R. C. Pianta, M. J. Cox, & K. L. Snow (Eds.), *School readiness and the transition to kindergarten in the era of accountability* (pp. 49-83). Baltimore, MD: Brookes.

Hamre, B. K., Pianta, R. C., Burchinal, M., Field, S., LoCasale-Crouch, J., Downer, J. T., …, & Scott-Little, C. (2012). A course on effective teacher-child interactions: Effects on teacher beliefs, knowledge, and observed practice. *American Educational Research Journal*, 49, 88–123. doi: 10.3102/0002831211434596.

Hamre, B. K., Pianta, R. C., Downer, J. T., DeCoster, J., Mashburn, A. J., Jones, S. M., ... & Hamagami, A. (2013). Teaching through interactions: Testing a developmental framework of teacher effectiveness in over 4,000 classrooms. *The Elementary School Journal*, *113*, 461-487. doi: 10.1086/669616.

Hatfield, B. E., Burchinal, M. R., Pianta, R. C., & Sideris, J. (2016). Thresholds in the association between quality of teacher–child interactions and preschool children's school readiness skills. *Early Childhood Research Quarterly*, *36*, 561-571. doi: 10.1016/j.ecresq.2015.09.005.

Heckman, J. J., & Masterov, D. V. (2007). The productivity argument for investing in young children. *Applied Economic Perspectives and Policy, 29,* 446-493. doi: 10.1111/j.1467-9353.2007.00359.x.

Hemmeter, M. L., & Fox, L. (2009). The Teaching Pyramid: A model for the implementation of classroom practices within a program-wide approach to behavior support. *NHSA DIALOG*, *12*, 133-147. doi: 10.1080/15240750902774718.

Hemmeter, M. L., Ostrosky, M., & Fox, L. (2006). Social and emotional foundations for early learning: A conceptual model for intervention. *School Psychology Review*, *35*, 583-601.

Hix-Small, H., Duncan, T., Duncan, S., & Okut, H. (2004). A multivariate associative finite growth mixture modeling approach examining adolescent alcohol and marijuana use. *Journal of Psychopathology and Behavioral Assessment*, 26, 255–270. doi: 10.1023/B:JOBA.0000045341.56296.fa.

Ladd, G. W., & Dinella, L. M. (2009). Continuity and change in early school engagement: Predictive of children's achievement trajectories from first to eighth grade?. *Journal of Educational Psychology*, *101*, 190-206. doi: 10.1037/a0013153.

Landry, S. H., Anthony, J. L., Swank, P. R., & Monseque-Bailey, P. (2009). Effectiveness of comprehensive professional development for teachers of at-risk preschoolers. *Journal of Educational Psychology*, *101*, 448-465. doi: 10.1037/a0013842.

Leyva, D., Weiland, C., Barata, M., Yoshikawa, H., Snow, C., Treviño, E., & Rolla, A. (2015). Teacher–child interactions in Chile and their associations with prekindergarten outcomes. *Child Development*, *86*, 781-799. doi: 10.1111/cdev.12342.

Li, C. (2013). Little's test of missing completely at random. *Stata Journal*, *13*, 795-809. doi: 10.1177/1536867X1301300407.

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*, 1062-1077. doi: 10.1037/a0020066.

Little, R. J. (1988). A test of missing completely at random for multivariate data with missing values. *Journal of the American Statistical Association*, *83*, 1198-1202.

Liu, A. (2016). *Non-cognitive skills and the growing achievement gap*. (Report 16-861). Ann Arbor, MI: Population Studies Center.

Lo, Y., Mendell, N., & Rubin, D. (2001). Testing the number of components in a normal mixture. *Biometrika*, 88, 767–778. doi: 10.1093/biomet/88.3.767

Lonigan, C. J., Wagner, R. K., Torgesen, J. K., & Rashotte, C. A. (2007). *Test of preschool early literacy.* Austin, TX: Pro-Ed.

Maier, M. F., Vitiello, V. E., & Greenfield, D. B. (2012). A multilevel model of child-and classroom-level psychosocial factors that support language and literacy resilience of children in Head Start. *Early Childhood Research Quarterly*, *27*, 104-114. doi: 10.1016/j.ecresq.2011.06.002.

Marsh, H. W., Lüdtke, O., Trautwein, U., & Morin, A. J. (2009). Classical latent profile analysis of academic self-concept dimensions: Synergy of person-and variable-centered approaches to theoretical models of self-concept. *Structural Equation Modeling*, *16*, 191-225. doi: 10.1080/10705510902751010.

Mashburn, A. J., Justice, L. M., Downer, J. T., & Pianta, R. C. (2009). Peer effects on children's language achievement during pre-kindergarten. *Child Development*, *80*, 686-702. doi: 10.1111/j.1467-8624.2009.01291.x.

Mashburn, A. J., Pianta, R. C., Hamre, B. K., Downer, J. T., Barbarin, O. A., Bryant, D., ... & Howes, C. (2008). Measures of classroom quality in prekindergarten and children's development of academic, language, and social skills. *Child Development*, *79*, 732-749. doi: 10.1111/j.1467-8624.2008.01154.x.

McClelland, M. M., Geldhof, J. G., Cameron, C. E. & Wanless, S. B. (2015). Development and self-regulation. In R. M. Lerner (Ed.), *Handbook of child psychology and developmental science* (7th ed., Vol. 1, pp. 1-43). Hoboken, NJ: John Wiley & Sons, Inc.

McGrew, K. S., & Woodcock, R. W. (2001). *Woodcock-Johnson III technical manual*. Rolling Meadows, IL: Riverside Publishing.

Merritt, E. G., Wanless, S. B., Rimm-Kaufman, S. E., Cameron, C., & Peugh, J. L. (2012). The contribution of teachers' emotional support to children's social behaviors and self-regulatory skills in first grade. *School Psychology Review*, *41*, 141-159.

Muthén, B. O. (2004). Latent variable analysis: Growth mixture modeling and related techniques for longitudinal data. In D. Kaplan (Ed.), *Handbook of quantitative* *methodology for the social sciences* (pp. 345–368). Newbury Park, CA: Sage.

Muthén, L. K., & Muthén, B. O. (1998–2015). *Mplus user's guide* (7th ed.). Los Angeles, CA: Muthén & Muthén.

Nix, R. L., Bierman, K. L., Domitrovich, C. E., & Gill, S. (2013). Promoting children's social-emotional skills in preschool can enhance academic and behavioral functioning in kindergarten: Findings from Head Start REDI. *Early Education & Development*, *24*, 1000-1019. doi: 10.1080/10409289.2013.825565.

Nylund, K. L., Asparouhov, T., & Muthén, B. O. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: A monte carlo simulation study. *Structural Equation Modeling-A Multidisciplinary Journal*, 14, 535–569. doi: 10.1080/10705510701575396.

O’Connor, E., & McCartney, K. (2007). Examining teacher–child relationships and achievement as part of an ecological model of development. *American Educational Research Journal*, 44, 340-369. doi: 10.3102/0002831207302172.

Pakarinen, E., Aunola, K., Kiuru, N., Lerkkanen, M. K., Poikkeus, A. M., Siekkinen, M., & Nurmi, J. E. (2014). The cross-lagged associations between classroom interactions and children’s achievement behaviors. *Contemporary Educational Psychology*, *39*, 248-261. doi: 10.1016/j.cedpsych.2014.06.001.

Pianta, R. C., Burchinal, M., Jamil, F. M., Sabol, T., Grimm, K., Hamre, B. K., ... & Howes, C. (2014). A cross-lag analysis of longitudinal associations between preschool teachers’ instructional support identification skills and observed behavior. *Early Childhood Research Quarterly*, *29*, 144-154. doi: 10.1016/j.ecresq.2013.11.006.

Pianta, R. C., DeCoster, J., Cabell, S., Burchinal, M., Hamre, B. K., Downer, J., ... & Howes, C. (2014). Dose–response relations between preschool teachers' exposure to components of professional development and increases in quality of their interactions with children. *Early Childhood Research Quarterly*, *29*, 499-508. doi: 10.1016/j.ecresq.2014.06.001.

Pianta, R. C., La Paro, K. M., & Hamre, B. K. (2008). *Classroom assessment scoring system manual, pre-k.* Baltimore, MD: Brookes.

Ponitz, C. C., Rimm-Kaufman, S. E., Grimm, K. J., & Curby, T. W. (2009). Kindergarten classroom quality, behavioral engagement, and reading achievement. *School Psychology Review*, *38*, 102-120.

Powell, D., Dunlap, G., & Fox, L. (2006). Prevention and intervention for the challenging behaviors of toddlers and preschoolers. *Infants & Young Children*, *19*, 25-35. doi: 10.1097/00001163-200601000-00004.

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*, 394-408. doi: 10.1111/desc.12027.

Reardon, S. F. (2013). The widening income achievement gap. *Educational Leadership*, *70*, 10-16. doi: 10.1080/00223891.2010.496477.

Reise, S. P., Moore, T. M., & Haviland, M. G. (2010). Bifactor models and rotations: Exploring the extent to which multidimensional data yield univocal scale scores. *Journal of Personality Assessment*, *92*, 544-559. doi: 10.1080/00223891.2010.496477.

Rowe, M. L. (2008). Child-directed speech: relation to socioeconomic status, knowledge of child development and child vocabulary skill. *Journal of Child Language*, *35*, 185-205. doi: 10.1017/S0305000907008343.

Sabol, T. J., Bohlmann, N. L., & Downer, J. T. (2018). Low-income ethnically diverse children's engagement as a predictor of school readiness above preschool classroom quality. *Child Development, 89*, 556-576. doi: 10.1111/cdev.12832.

Sarama, J., & Clements, D. H. (2007). *Manual for classroom observation of early mathematics: Environment and teaching (COEMET) version 3*. Unpublished version.

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, 6, 461–464.

Smith, M. W., & Dickinson, D. K. (2002). *Early Language & Literacy Classroom Observation (ELLCO) toolkit, research edition [with] user's guide*. Baltimore, MD: Brookes.

Smith-Donald, R., Raver, C. C., Hayes, T., & Richardson, B. (2007). Preliminary construct and concurrent validity of the Preschool Self-regulation Assessment (PSRA) for field-based research. *Early Childhood Research Quarterly*, *22*, 173-187. doi: 10.1016/j.ecresq.2007.01.002.

Stipek, D. J., & Ryan, R. H. (1997). Economically disadvantaged preschoolers: Ready to learn but further to go. *Developmental Psychology*, *33*, 711-723. doi: 10.1037/0012-1649.33.4.711.

U.S. Department of Health and Human Services (DHHS). (2009). *The 2009 HHS poverty guidelines.* Retrieved from https://aspe.hhs.gov/2009-hhs-poverty-guidelines. Office of the Assistant Secretary for Planning and Evaluation, U.S. DHHS.

Viljaranta, J., Lerkkanen, M. K., Poikkeus, A. M., Aunola, K., & Nurmi, J. E. (2009). Cross-lagged relations between task motivation and performance in arithmetic and literacy in kindergarten. *Learning and Instruction*, *19*, 335-344. doi: 10.1016/j.learninstruc.2008.06.011.

Vitiello, V. E., Booren, L. M., Downer, J. T., & Williford, A. P. (2012). Variation in children's classroom engagement throughout a day in preschool: Relations to classroom and child factors. *Early Childhood Research Quarterly*, *27*, 210-220. doi:10.1016/j.ecresq.2011.08.005.

Vitiello, V., & Williford, A. P. (2016). Relations between social skills and language and literacy outcomes among disruptive preschoolers: Task engagement as a mediator. *Early Childhood Research Quarterly*, *36*, 136-144. doi: 10.1016/j.ecresq.2015.12.011.

Webster-Stratton, C., & Reid, M. J. (2004). Strengthening social and emotional competence in young children—The foundation for early school readiness and success: Incredible years classroom social skills and problem-solving curriculum. *Infants & Young Children*, *17*, 96-113.

Whitehurst, G. J., & Lonigan, C. J. (2001). Emergent literacy: Development from prereaders to readers. In S. B. Neuman, & D. K. Dickinson (Eds.), *Handbook of early literacy research*, (Vol. 1, pp.11-29). New York, NY: Guilford.

Williams, K. T., & Wang, J. (1997). *Technical references to the Peabody Picture Vocabulary Test (PPVT-III)* (3rd ed.). Circle Pines, MN: American Guidance Service.

Williford, A. P., Maier, M. F., Downer, J. T., Pianta, R. C., & Howes, C. (2013).

Understanding how children's engagement and teachers' interactions combine to predict school readiness. *Journal of Applied Developmental Psychology*, *34*, 299-309. doi: 10.1016/j.appdev.2013.05.002.

Zaslow, M., Anderson, R., Redd, Z., Wessel, J., Tarullo, L., & Burchinal, M. (2010). *Quality dosage thresholds and features in early childhood settings: A review of the literature* (No. cdea2df43cdf4b10bb74af4189cbab8d). Washington, DC: Mathematica Policy Research.

Table 1
Standardized and Unstandardized Factor Loadings and Model Fit for a Bifactor Model of Teacher-Child Interactions

	Responsive teaching		Cognitive facilitation		Positive management and routines	
	β	(B)	β	(B)	β	(B)
CLASS dimensions						
Positive climate	0.85	(1.00)	-	-	0.35	(1.00)
Negative climate (reversed)	0.46	(0.25)	-	-	0.48	(0.64)
Teacher sensitivity	0.94	(1.25)	-	-	0.11	(0.35)
Regard for student perspectives	0.87	(1.15)	-	-	-0.15	(-0.47)
Behavior management	0.72	(0.73)	-	-	0.56	(1.39)
Productivity	0.64	(0.63)	-	-	0.49	(1.20)
Instructional learning formats	0.84	(1.03)	-	-	-0.12	(-0.37)
Concept development	0.50	(0.38)	0.65	(1.00)	-	-
Quality of feedback	0.75	(0.72)	0.58	(1.13)	-	-
Language modeling	0.77	(0.81)	0.53	(1.14)	-	-
Variances	1.00	(0.74)	1.00	(0.18)	1.00	(0.12)
Model fit						
CFI	0.937					
RMSEA	0.155					
SRMR	0.036					

Note. CLASS = Classroom Assessment Scoring System. CFI = Comparative fit index. RMSEA = Root mean square error of approximation. SRMR = Standardized root mean square residual.

Table 2

Latent Profile Analysis Fit Statistics across Models with Different Numbers of Profiles

	1 factor	2 factors	3 factors	4 factors
AIC	36658.14	32995.31	32797.28	32845.27
BIC	36868.14	34095.54	34007.07	34164.63
Sample-size adjusted BIC	36722.08	33330.31	33165.63	33246.99
Entropy	-	0.65	0.76	0.81
Vuong-Lo-Mendell-Rubin *p*-value	-	0.37	0.76	0.35
Lo-Mendell-Rubin Adjusted LRT *p*-value	-	0.37	0.76	0.35

Note. AIC = Akaike Information Criterion. BIC = Bayesian Information Criterion. LRT = Likelihood ratio test. Entropy, VLMR *p*-value, and Adjusted LRT *p*-value are not provided for the 1-factor model. Auxiliary variables used to create profiles included: child age, sex, race/ethnicity, primary language, years of maternal education, income-to-needs ratio, and household size (child level; teacher years of education and experience, intervention condition (teacher level); and classroom focus on language and literacy, and percent of classroom in poverty (classroom level).

Table 3

Descriptive Statistics of Key Study Variables for the Full Sample and Three Profiles

	Full Sample			Positively Engaged			Typically Engaged			Negatively Engaged		
	M/%	*SD*	Range	*M/%*	*SD*	Range	*M/%*	*SD*	Range	*M/%*	*SD*	Range
Engagement as measured by inCLASS												
Teacher Engagement	2.21	0.85	1 - 6.17	2.57[a]	0.97	1 - 6.17	2.00[b]	0.70	1 - 4.92	2.26[b]	0.82	1 - 4.25
Peer Engagement	2.49	2.49	1 - 6.56	3.40[a]	0.79	1.78 - 6.56	2.04[b]	0.52[b]	1 - 4.17	2.25[c]	0.56	1.33 - 3.67
Task Engagement	4.26	0.81	1.83 - 6.5	4.98[a]	0.58	3.67 - 6.5	3.99[b]	0.64	1.83 - 5.75	3.36[c]	0.66	1.83 - 4.67
Negative Engagement	1.37	0.38	1 - 3.67	1.32[a]	0.27	1 - 2.22	1.27[a]	0.24	1 - 2	2.28[b]	0.39	1.81 - 3.67
Classroom quality (standardized values)												
Responsive teaching	0	0.84	-2.35 - 1.91	0.20[a]	0.81	-2.35 - 1.91	-0.07[b]	0.83	-2.35 - 1.74	-0.26[b]	0.87	-1.75 - 1.5
Cognitive facilitation	0	0.38	-1.03 - 1.11	0.07[a]	0.39	-0.79 - 1.11	-0.04[b]	0.38	-1.03 - 1.08	.01[a,b]	0.32	-.76 - .75
Positive management and routines	0	0.30	-.88 - .78	-0.10[a]	0.27	-0.88 - .74	0.06[b]	0.31	-0.88 - .78	-0.04[a]	0.29	-.82 - .55
School Readiness												
Fall												
Expressive language	11.50	4.76	0 - 23	13.04[a]	3.94	1.0 - 23	10.61[b]	4.97	0 - 22	11.69[a,b]	4.72	1 - 22
Receptive language	39.12	20.73	1 - 101	46.15[a]	19.66	5.0 - 101	35.28[b]	20.03	1 -87	38.18[b]	22.82	6 - 90
Phonological awareness	11.97	5.33	0 - 27	13.10[a]	5.10	0 - 27	11.44[b]	5.29	0 - 27	10.62[b]	5.73	0 - 24
Print knowledge	12.60	10.38	0 - 36	16.89[a]	10.82	0 - 36	10.58[b]	9.52	0 - 36	9.61[b]	9.03	1 - 34
Inhibitory control	0.48	0.33	0 - 1	0.56[a]	0.31	0 - 1	0.45[b]	0.33	0 - 1	0.37[b]	0.33	0 - 1
Working memory	1.17	0.47	1 - 4	1.25[a]	0.57	1 - 4	1.12[b]	0.39	1 - 4	1.18[a,b]	0.52	1 - 3
Spring												
Expressive language	12.93	4.11	1 - 23	14.34[a]	3.38	3 - 23	12.18[b]	4.26	1 - 22	12.70[b]	4.26	2 - 23
Receptive language	48.85	20.27	5 - 101	55.40[a]	18.91	7 - 97	45.85[b]	20.30	5 - 94	44.45[b]	19.37	7 - 101
Phonological awareness	14.42	5.81	1 - 27	16.07[a]	5.48	3 - 27	13.70[b]	5.76	1 - 27	12.78[b]	6.09	2 - 24
Print knowledge	21.05	11.35	0 - 36	25.06[a]	9.92	2 - 36	19.62[b]	11.50	0 - 36	15.67[c]	10.94	1 - 36
Inhibitory control	0.62	0.32	0 - 1	0.70[a]	0.31	0 - 1	0.60[b]	0.32	0 - 1	0.47[c]	0.31	0 - 1
Working memory	1.35	0.67	1 - 5	1.51[a]	0.79	1 - 5	1.28[b]	0.60	1 - 4	1.18[b]	0.47	1 - 3

Note. inCLASS scores are on a scale of 1 (low quality) to 7 (high quality), with the exception of negative classroom engagement, for which higher ratings indicate more negative interactions. In each row, significant pairwise differences are denoted by different superscript letters ($p < 0.05$). If settings have corresponding superscript letters, then the pairwise difference is non-significant.

Table 4

Descriptive Statistics of Covariates for the Full Sample and Three Profiles

	Full Sample			Positively Engaged			Typically Engaged			Negatively Engaged		
	M/%	*SD*	Range	*M/%*	*SD*	Range	*M/%*	*SD*	Range	*M/%*	*SD*	Range
Child level												
Child age (months as of 9/1)	49.60	6.01	27.29 - 64.28	51.38[a]	4.72	34.06 - 59.80	48.95[b]	6.28	30.67 - 64.28	47.34[b]	6.82	27.29 - 58.38
Child is male	49.86%	-	-	47.14%[a]	-	-	48.71%[a]	-	-	68.97%[b]	-	-
Child is White/Caucasian	13.83%	-	-	12.50%[a]	-	-	14.29%[a]	-	-	15.79%[a]	-	-
Child is Black/African American	41.07%	-	-	45.54%[a]	-	-	38.01%[a]	-	-	45.61%[a]	-	-
Child is Hispanic/Latino	35.45%	-	-	29.91%[a]	-	-	38.98%[a]	-	-	31.58%[a]	-	-
Child is another race	9.65%	-	-	12.05%[a]	-	-	8.72%[a]	-	-	7.02%[a]	-	-
Child's primary language is English	84.63%	-	-	92.92%[a]	-	-	79.76%[b]	-	-	87.93%[a,b]	-	-
Years of maternal education	12.72	2.39	8 - 20	13.32[a]	2.65	8 - 20	12.39[b]	2.24	8 - 20	12.61[a,b]	1.83	8 - 18
Family income-to-needs ratio	1.10	1.05	.05 - 5.07	1.43[a]	1.30	0.07 - 5.07	0.93[b]	0.84	0.05 - 4.20	0.96[b]	0.98	0.08 - 4.49
Household size	4.43	1.59	2 - 15	4.05[a]	1.33	2 - 11	4.66[b]	1.71	2 - 15	4.32[a,b]	1.39	2 - 9
Fall - spring assessment window (days)	157.87	35.10	84 - 273	154.68[a]	35.24	90 - 244	160.06[a]	34.96	84 - 273	155.78[a]	35.13	98 - 243
Teacher/classroom level												
Teacher years of education	16.02	1.60	12 - 20	16.48[a]	1.48	13 - 20	15.87[b]	1.61	12 - 20	15.23[c]	1.50	12 - 18
Teacher years of experience	8.41	6.41	1 - 35	9.64[a]	7.57	1 - 35	7.77[b]	5.66	1 - 35	8.25[a,b]	6.02	1 - 35
Curriculum focused on language and literacy	32.18%	-	-	34.05%[a]	-	-	32.27%[a]	-	-	24.49%[a]	-	-
Average income-to-needs ratio in class	1.09	0.78	0.10 - 4.34	1.30[a]	0.93	0.14 - 4.34	0.98[b]	0.68	0.10 - 4.34	0.98[b]	0.72	0.10 - 3.82
Setting												
HS / public school	14.43%	-	-	9.77%[a]	-	-	17.11%[b]	-	-	12.73%[a,b]	-	-
Non-public HS	39.62%	-	-	36.28%[a]	-	-	40.59%[a]	-	-	45.45%[a]	-	-
Other public school	23.56%	-	-	27.91%[a]	-	-	22.00%[a]	-	-	18.18%[a]	-	-
Public, non-profit, or for-profit agency	22.39%	-	-	26.05%[a]	-	-	20.29%[a]	-	-	23.64%[a]	-	-
N (%) sample	710 (100%)			227 (31.97%)			425 (59.86%)			58 (8.17%)		

Note. Teacher intervention condition was also controlled for in models but is not shown here because it was not of interest in the current study. In each row, significant pairwise differences are denoted by different superscript letters ($p < 0.05$). If settings have corresponding superscript letters, then the pairwise difference is non-significant.

Table 5

Associations Between Children's Engagement, Teacher-Child Interaction Quality, and Spring School Readiness Outcomes

	Expressive Language		Receptive Language		Phonological Awareness		Print Knowledge		Inhibitory Control		Working Memory	
	β (SE)	β (SE)	β (SE)	β (SE)	β (SE)	β (SE)	β (SE)	β (SE)	β (SE)	β (SE)	β (SE)	β (SE)
Child engagement												
Positively engaged	< .01 (.02)	< .01 (.02)	0.01 (.03)	< -0.01 (.03)	0.06 (.04)	0.06 (.04)	0.02 (.03)	0.01 (.03)	0.04 (.04)	0.06 (.04)	0.02 (.04)	0.04 (.04)
Negatively engaged	-0.01 (.02)	-0.01 (.02)	-0.06** (.02)	-0.06** (.02)	-0.02 (.03)	-0.01 (.03)	-0.06* (.02)	-0.05* (.02)	-0.07* (.03)	-0.06 (.03)	-0.05* (.02)	-0.04 (.02)
T-C interaction quality												
Responsive teaching	-	0.01 (.02)	-	0.01 (.02)	-	0.05 (.04)	-	0.02 (.03)	-	-0.01 (.04)	-	0.05 (.04)
Cognitive facilitation	-	0.02 (.02)	-	0.02 (.02)	-	0.01 (.04)	-	0.04 (.03)	-	< -0.01 (.04)	-	.01 (.04)
Pos. management & routines	-	0.03 (.03)	-	0.03 (.02)	-	0.09 (.05)	-	0.02 (.04)	-	0.09* (.04)	-	0.09* (.03)

* $p < .05$, ** $p < .01$, *** $p < .001$.

Note. N = 710. T-C = Teacher-child. Standardized coefficients are presented. FIML was used to estimate missing data, and robust standard errors were clustered by teacher (N = 220). Covariates included but not shown: child age, sex, race/ethnicity, primary language, years of maternal education, income-to-needs ratio, household size, fall scores, and time between fall and spring assessments (child level); teacher years of experience, teacher years of education, classroom focus on language and literacy, program setting type, and dummy codes for teacher intervention condition (teacher/classroom level).

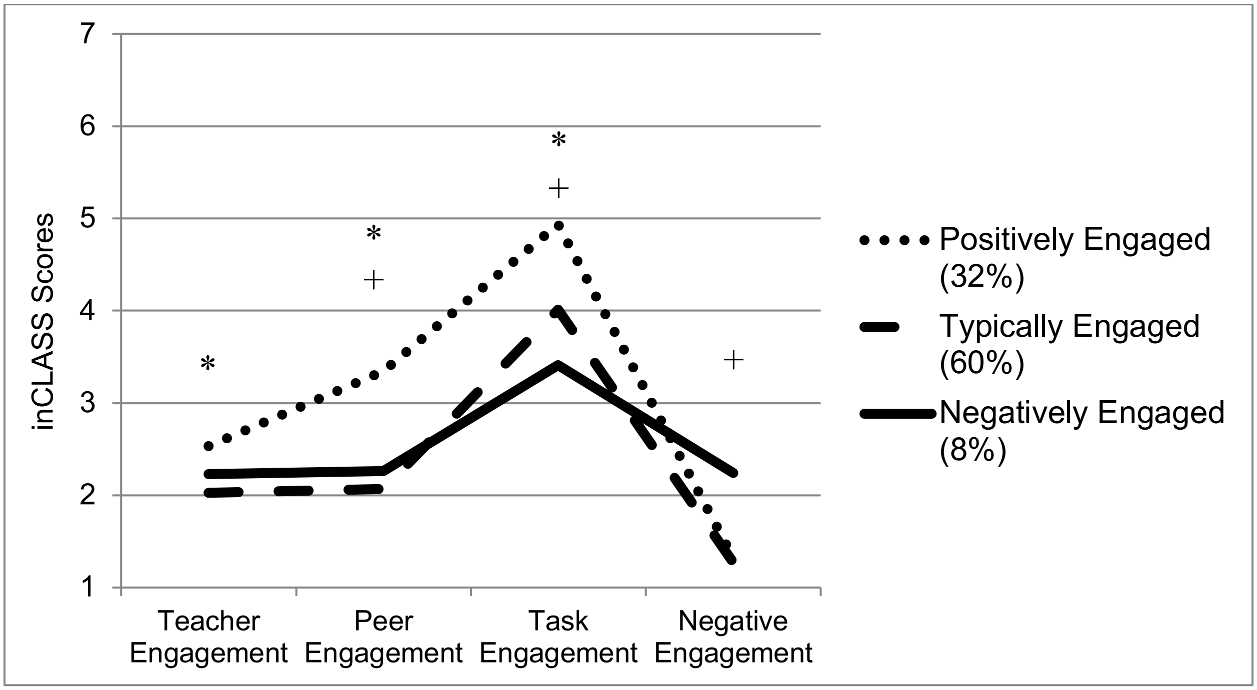

Figure 1. Graphical depiction of inCLASS engagement profiles. * denotes a statistically significant difference between the Positively Engaged and Typically Engaged profiles on a given inCLASS domain. + denotes a statistically significant difference between the Typically Engaged and Negatively Engaged profiles on a given inCLASS domain.

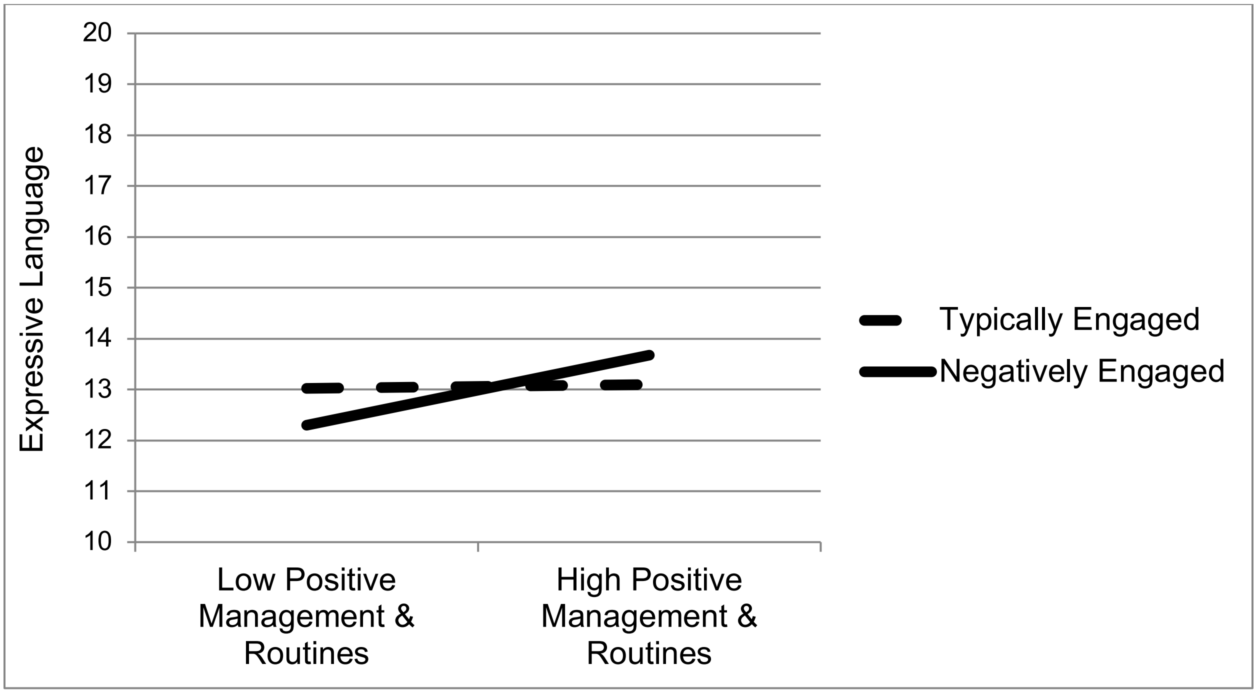

Figure 2. Graphical representation of the interaction effect between engagement profile membership and one standard deviation above (High) and below (Low) Positive Management and Routines on gains in expressive language across the preschool year.

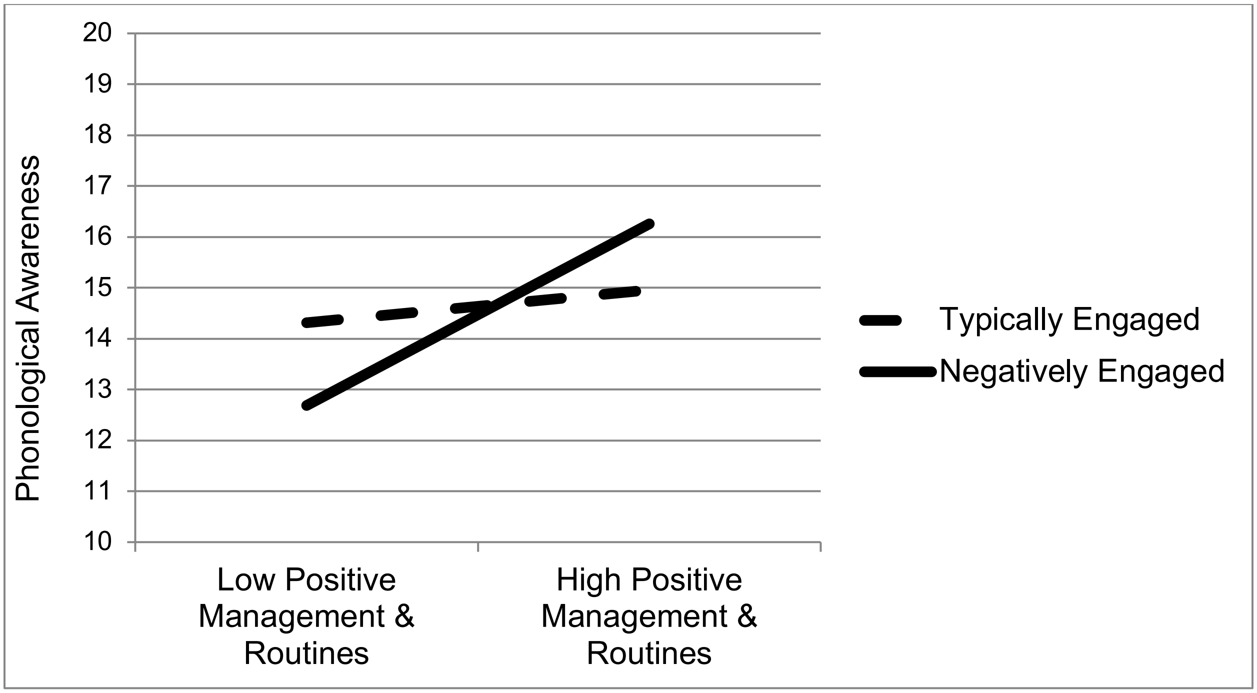

Figure 3. Graphical representation of the interaction effect between engagement profile membership and one standard deviation above (High) and below (Low) Positive Management and Routines on gains in phonological awareness across the preschool year.

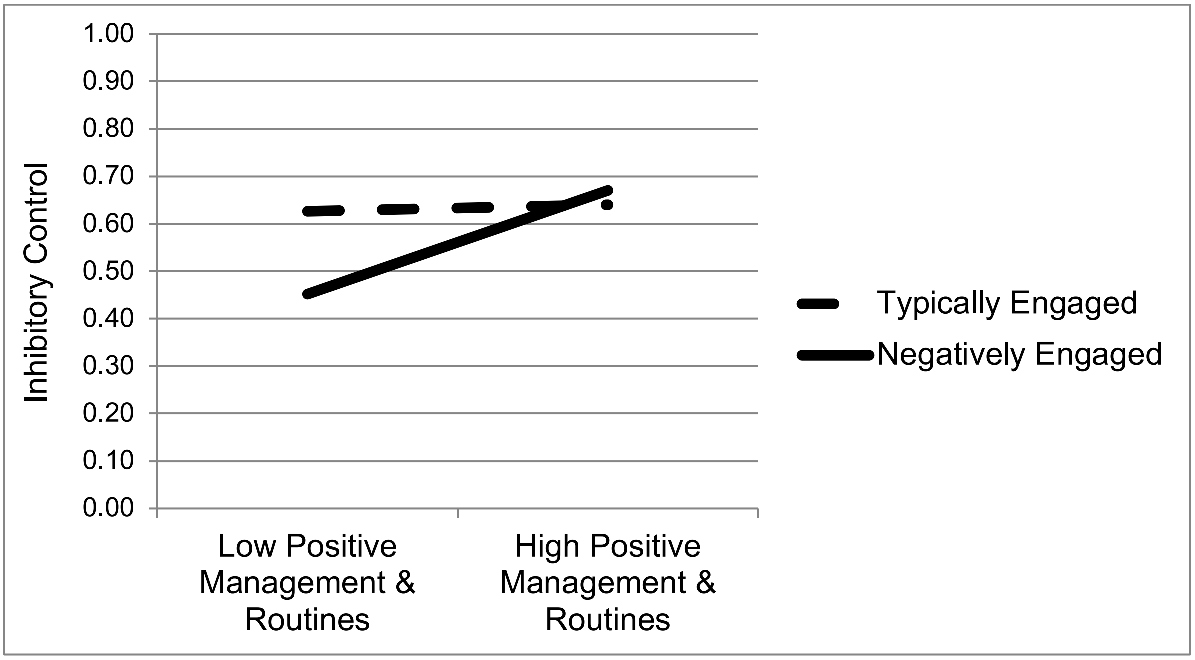

Figure 4. Graphical representation of the interaction effect between engagement profile membership and one standard deviation above (High) and below (Low) Positive Management and Routines on gains in inhibitory control across the preschool year.

MANUSCRIPT TWO

Roles of Executive Functioning and Language in Developing Low-Income Preschoolers' Behavior and Emotion Regulation

Shannon E. Reilly, M.Ed., and Jason T. Downer, Ph.D.

Center for Advanced Study of Teaching and Learning, University of Virginia

Revision submitted, *Early Childhood Research Quarterly*

Author Note

Shannon Reilly, Center for Advanced Study of Teaching and Learning, University of Virginia; Jason T. Downer, Center for Advanced Study of Teaching and Learning, University of Virginia.

This work was supported by the National Institute of Child Health and Human Development of the National Institutes of Health (grant number 2R01HD051498-06A1). The opinions expressed are those of the authors and do not represent views of the funding agency. Special thanks to the participating teachers, children, and families.

Correspondence concerning this article should be addressed to Shannon Reilly, Center for Advanced Study of Teaching and Learning, University of Virginia, PO Box 400267, Charlottesville, VA 22904. Email: sr6hd@virginia.edu

Abstract

Young children's regulation of their behaviors and emotions is a foundational skill that undergirds learning, academic achievement, and social competence (Bierman & Erath, 2006; McClelland et al., 2018). Executive functioning (EF) and language are two cognitive skillsets that facilitate behavior and emotion regulation (Blair & Ursache, 2011; Cole, Armstrong, & Pemberton, 2010). What is not yet fully understood is how these two skillsets may work together to promote these regulatory skills. The present study investigated the independent and interactive contributions of EF and language skills at preschool entry to the development of behavior and emotion regulation across the year in a primarily low-income sample. Results indicated that language at preschool entry was associated with children's emotion regulation development during preschool, especially for children who entered preschool with low EF. As such, incorporating language-promoting activities into early childhood interventions designed to facilitate emotion regulation may enhance efficacy, particularly for children at risk for later emotional, academic, and behavioral difficulties due to low emerging EF skills. Unexpectedly, language was not associated with behavior regulation, and EF was not independently linked to behavior or emotion regulation in this sample.

Keywords: preschool; executive functioning; language; self-regulation; school readiness

Roles of Executive Functioning and Language in Developing Low-Income Preschoolers' Behavior and Emotion Regulation

Young children's abilities to regulate their behaviors and emotions are foundational skills that undergird learning, academic achievement, and social competence (Bierman & Erath, 2006; Eisenberg, Spinrad, & Eggum, 2010; McClelland et al., 2018; Ponitz, McClelland, Matthews, & Morrison, 2009). Conversely, difficulty regulating behaviors and emotions has been linked to subsequent academic, emotional, and behavioral difficulties (Barkley, 1989; Moffitt et al., 2011; Steinberg & Drabick, 2015; Ursache, Blair, & Raver, 2012). As such, it is an essential task of early childhood to develop skills in regulating behaviors and emotions (Bierman & Erath, 2006; Knudsen, Heckman, Cameron, and Shonkoff, 2006; McClelland et al., 2018; Phillips & Shonkoff, 2000), particularly prior to kindergarten entry. This transition represents a sensitive developmental time that sets the foundation for later school success (Rimm-Kaufman & Pianta, 2000) and is characterized by increases in demands and expectations around self-regulation (Bassok, Latham, & Rorem, 2016).

Two sets of cognitive skills that facilitate the development of behavior and emotion regulation are executive functioning (EF; Blair & Ursache, 2011; Hay, Payne, & Chadwick, 2004) and language (Cole, Armstrong, & Pemberton, 2010; Vygotsky, 1962). EF can be understood as a multidimensional skillset – including attending, mentally manipulating information, and thinking flexibly – that enables goal-directed behavior (Baggetta & Alexander, 2016; Blair, 2016). Children's language ability is comprised of both words they use to express themselves (i.e., expressive language) and those they understand (i.e., receptive language).

EF and language develop rapidly in early childhood (Blair, Zelazo, & Greenberg, 2005; Halliday, 2006) and concurrently with one another (Gooch, Thompson, Nash, Snowling, & Hulme, 2016). While there is widespread evidence indicating that these skillsets each facilitate young children's regulation separately, what has not yet been fully explored is how they might work together to promote emotion and behavior regulation. It is particularly important to examine these pathways for children from low-SES backgrounds, who tend to already have lower EF, language, and regulatory abilities than more affluent peers by kindergarten entry (Fernald, Marchman, & Weisleider, 2013; Miech, Essex, & Goldsmith, 2001; Raver, 2004; Raver, McCoy, Lowenstein, & Pess, 2013). Thus, the present study addresses this gap in the literature by investigating independent and interactive contributions of EF and language skills at preschool entry to development in behavior and emotion regulation across the year in a primarily low-income sample.

The Importance of Regulating Behaviors and Emotions in Early Childhood

There is a large body of work with primarily low-income (Raver et al., 2011) and mixed-income (Sabol & Pianta, 2012) samples indicating that children's regulatory abilities undergird the development of pre-academic and academic skills (Shonkoff, 2011). "Self-regulation" is a general umbrella term applied to the ability to control one's attention, cognitions, emotions, and behaviors to achieve a goal (Ackerman & Friedman-Krauss, 2017; Jones, Bailey, Barnes, & Partee, 2016; McClelland et al., 2018), such as interacting effectively with others, engaging in tasks, and adapting to changing demands (Vallotton & Ayoub, 2011). Two types of self-regulation, behavior regulation and emotion regulation, were explored in this study to examine early childhood regulatory processes more specifically (Jones et al., 2016).

Behavior regulation. Although there is not a single definition for behavior regulation, it has been operationalized as conscious "control of external behaviors" (Bronson, 2000, p. 3), such as "paying attention, following instructions, and inhibiting inappropriate actions" (McClelland, Cameron, Connor et al., 2007, p. 947). It involves engaging in goal-directed behavior that is adaptive in a given setting (Bronson, 2000) and includes the ability to match expectations of the setting by regulating speech and actions (Downer, Booren, Lima, Luckner, & Pianta, 2010). This skill develops rapidly in the first few years of life from controlling simple motor responses in infancy to being able to participate in learning activities appropriately in early childhood (Bronson, 2000). Behavior regulation makes children available for learning in the classroom (e.g., Bierman, Nix, Greenberg, Blair, & Domitrovich, 2008) and enables them to respond to changing demands in the context of a particular setting (Jones et al., 2016; McClelland et al., 2018). In Head Start, early development of behavior regulation has been shown to promote concurrent growth in school readiness skills in other domains (Bierman, Domitrovich et al., 2008). Longitudinally, it has been shown to facilitate subsequent academic achievement (Jones et al., 2016; Ponitz et al., 2009; Sektnan, McClelland, Acock, & Morrison, 2010; Ursache et al., 2012) and educational attainment (Duckworth & Carlson, 2013) in mixed-income samples. Conversely, having underdeveloped behavior regulation in the first decade of life has been associated with health problems and financial difficulties, as well as elevated risk for depression, substance abuse, and criminal conviction by age 32 (Moffitt et al., 2011).

Emotion regulation. Emotion regulation has been operationalized as “processes used to manage and change if, when, and how (e.g., how intensely) one experiences emotions and emotion-related motivational and physiological states, as well as how emotions are expressed” (Eisenberg, Hofer, & Vaughan, 2007, p. 288) and that “enable an individual to function adaptively in emotionally arousing situations” (Cicchetti, Ganiban, & Barnett, 1991, p. 15). Emotion regulation also develops rapidly in the early years of life (e.g., from thumb-sucking as a self-soothing mechanism in infancy to verbally expressing wants, needs, and emotions in early childhood) and improves more slowly into adulthood (Eisenberg et al., 2010; Eisenberg & Sulik, 2012). Well-developed emotion regulation in early childhood has been associated with positive school adjustment (Herndon, Bailey, Shewark, Denham, & Bassett, 2013), greater social competence with and acceptance from peers, and the development of friendships (Hay et al., 2004). This skill can be particularly helpful in dealing with conflict and intense emotions, which tend to crop up more in early childhood once children have to interact with several new peers at once, manage competing demands, and share resources and adult attention (Hay et al., 2004; Roben, Cole, & Armstrong, 2013). Cultivating emotion and behavior regulation in preschool specifically is important because of the increase in demands and expectations – particularly around self-regulation – that occurs with the transition to kindergarten (Bassok et al., 2016; Rimm-Kaufman & Pianta, 2000). Having underdeveloped emotion regulation has been shown to lead to significant academic, social, and behavioral difficulties (Bierman & Erath, 2006; Eisenberg et al., 2010; Hay et al., 2004; Steinberg & Drabick, 2015; Ursache et al., 2012).

Assessment of Regulation Skills in Preschool

Behavior regulation and emotion regulation are primarily assessed in context, typically using observations and teacher and/or parent ratings. For instance, the degree to which children can regulate their behaviors and manage expectations in the classroom has been observed (Downer et al., 2010), and teachers have reported on the degree to which children typically regulate their behaviors and emotions in the classroom on a daily basis (Hightower, 1986; Hightower & Perkins, 2010; Shields & Cicchetti, 1997).

There are strengths and challenges to each of these types of assessments. Observations provide a standardized, independent perspective in the context of children's individual experiences in real time (Eisenberg et al., 2010). However, observer ratings can be affected by factors such as observer bias and the time of day and activity setting during which ratings take place (Booren, Downer, & Vitiello, 2012; Kim et al., 2018; Mashburn, 2017; Vitiello, Booren, Downer, & Williford, 2012). Additionally, low-incidence behaviors might not be captured even over the course of multiple observation cycles (Downer et al., 2010; Kim et al., 2018). Teacher ratings can offer more generalized information about children's average behavior from an adult who knows a child and how he or she tends to interact on a daily basis (Eisenberg et al., 2010). However, these reports are typically retrospective, and they can be reflective in part of characteristics of students, teachers, and/or classrooms. For instance, in a racially and ethnically diverse, mixed-income sample of kindergarteners, more positive teacher ratings of children's competencies and behaviors were associated with children who were female, older, and had higher cognitive skills; teachers who were non-White, had fewer years of experience, and had higher self-efficacy; and classrooms with lower child-teacher ratios and shorter school days (Mashburn, Hamre, Downer, & Pianta, 2006).

Teacher biases in ratings of children's competencies have significant consequences, as a nationally representative study found that preschoolers whose teachers overestimated their abilities made more gains in kindergarten, whereas the opposite was true for those who were underestimated (Ready & Chu, 2015). This is a particularly important consideration in the present study, because children from low-income backgrounds are more likely to have their skills underestimated (Ready & Chu, 2015).

To capitalize on the strengths of both observations and teacher ratings and to counterbalance their challenges, the present study makes use of both types of measures in assessing young children's regulatory skills.

How EF Facilitates Preschoolers' Ability to Regulate in the Classroom

While there is not yet a standardized definition for EF, it has generally been described as a "set of cognitive processes" (Baggetta & Alexander, 2016, p. 24) that "coordinat[es] multiple sources of information in the service of purposeful, goal-directed behavior" (Blair, 2016, p. 102). EF is a multidimensional construct comprised of three primary components: working memory, inhibitory control, and set shifting or cognitive flexibility (Baggetta & Alexander, 2016; Blair et al., 2005; Miyake et al., 2000; Zelazo, Blair, & Willoughby, 2016). These skills evolve in early childhood (Blair et al., 2005; Garon, Bryson, & Smith, 2008) along with the development of the prefrontal cortex (Raver et al., 2013) and continue to grow throughout childhood and adolescence (Best & Miller, 2010). EF becomes more refined and complex over time, separating into its distinct components by middle childhood and adolescence (Lehto, Juujärvi, Kooistra, & Pulkkinen, 2003; Miyake et al., 2000). In early childhood, however, a large body of evidence has found that EF hangs together as a single construct (Hughes, Ensor, Wilson,

& Graham, 2009; Wiebe, Espy, & Charak, 2008; Wiebe et al., 2011; Willoughby, Blair, Wirth, Greenberg, and the Family Life Project Investigators, 2012). Taking a developmental perspective and following this early childhood EF literature, the present study conceptualizes EF as a single construct during preschool.

Recent literature has converged on the idea that regulation and EF are conceptually and methodologically distinct, albeit related constructs (Blair & Ursache, 2011; Jones et al., 2016). They are two aspects of learning-related, non-academic skills (McClelland, Cameron, Wanless et al., 2007) at different levels of complexity. Regulation is broader and more complex than executive functioning; it involves but is not completely comprised of EF (Ackerman & Friedman-Krauss, 2017; Jones et al., 2016). EF and self-regulation have often been used interchangeably in the literature, such that it is unclear which construct is being studied. This tends to occur more between behavior regulation and EF (e.g., Petersen, Bates, & Staples, 2015; Skibbe, Montroy, Bowles, & Morrison, 2018) than emotion regulation and EF, because there is more overlap in the former two constructs.

Not only are EF and regulation distinct constructs, but there is a directional relationship such that EF is a necessary precursor to behavior and emotion regulation. Children need to first develop cognitive executive functioning skills before they can fluidly integrate and apply them in dynamic contexts through well-regulated behaviors and emotions (Blair et al., 2005; Brock, Rimm-Kaufman, Nathanson, & Grimm, 2009; Hay et al., 2004; Riggs, Blair, & Greenberg, 2004). Thus, EF subserves or underlies behavior and emotion regulation, whereas regulatory skills are outward manifestations of EF (Bierman & Erath, 2006; Blair & Ursache, 2011; Hay et al., 2004; Hofmann, Schmeichel, & Baddeley, 2012; McClelland, Cameron, Wanless et al., 2007; Ponitz et al., 2009). In classrooms of schools that serve children from socioeconomically disadvantaged backgrounds, EF has been shown to predict learning-related regulatory behaviors, such as the ability to follow directions and classroom rules while engaging in activities, which then leads to more academic growth and social competence (Nesbitt, Farran, & Fuhs, 2015; Sasser, Bierman, & Heinrichs, 2015). EF has also been related to behavior regulation longitudinally during elementary school (Riggs et al., 2004). In addition, EF facilitates children's ability to regulate their emotions in the early childhood classroom (Hay et al., 2004). Thus, EF undergirds the capacity to engage in classroom tasks, meet behavioral expectations (Blair et al., 2005), "approach learning tasks more effectively and efficiently" (Bierman, Nix et al., 2008, p. 837), and adaptively modulate the experience and expression of emotions (Cicchetti et al., 1991; Eisenberg et al., 2007).

Independent and Interactive Contributions of EF and Language

Language is another foundational skill that underpins children's abilities to regulate their behaviors and emotions (Cole et al., 2010; Roben et al., 2013). As language skills evolve in early childhood (Halliday, 2006), they enable children to use language to govern their behaviors (Vygotsky, 1962), as well as to communicate wants, needs, and feelings (Bierman & Erath, 2006; Cole et al., 2010). Vygotsky's (1962) conceptualization of self-talk is central to this claim. Namely, as children mature in the first couple years of life, they begin to govern their behavior using private speech derived from caregivers' external regulation of their behavior in infancy and toddlerhood (e.g., telling a toddler to "wait" or to "stay close and hold hands" near a busy street). Children eventually internalize this speech to plan and guide their actions (Roben et al., 2013), which has been linked to social competence and fewer behavior problems (Winsler, De Leon, Wallace, Carlton, & Willson-Quayle, 2003) as well as increased ability to perform and persist in challenging tasks (Winsler, Manfra, & Diaz, 2007) in mixed-SES samples.

Both expressive and receptive language can facilitate children's regulation in the early childhood classroom. For example, expressive language allows children to express emotions verbally rather than act out behaviorally, and receptive language enables them to comprehend the rules and expectations of the classroom, as well as to understand peers' expressions of wants and needs (Bierman & Erath, 2006; Cole et al., 2010). In a primarily low-income sample, young children with strong language skills and rapid language growth prior to age 4 were found to demonstrate better ability to regulate their anger in an emotionally frustrating task than peers with lower and more slowly developing language (Roben et al., 2013). This was in part due to employing effective regulatory strategies, such as seeking support and distracting themselves. Conversely, having low language skills in early childhood has been related to difficulty regulating behaviors, such as displaying inattention and hyperactivity (Petersen et al., 2013), and emotions, such as increased frequency and intensity of anger (Hay et al., 2004; Roben et al., 2013).

Together, literature on EF and language suggests that both skillsets contribute to children's ability to regulate their behaviors and emotions in early childhood. However, these have largely been two separate bodies of literature that have not been brought into concert with one another. An exception is a study by Gooch and colleagues (2016), which found that when EF and language at age 4 were included in models predicting behavior regulation at ages 5 and 6 in a sample of children at risk for reading difficulty, EF promoted regulation but language did not. However, there remains a dearth of research examining the independent contributions of these two cognitive skillsets to emotion and behavior regulation in a typically developing, socioeconomically disadvantaged early childhood sample. Moreover, characteristics and resources influencing skill development, such as the accumulation of skills with which children enter preschool, generally interact to promote further skill growth rather than working additively (Bronfenbrenner & Morris, 2006). Specifically, children may need certain levels of EF and language skills to be able to integrate these abilities effectively to regulate behaviors and emotions (Cole et al., 2010). As such, what is not yet known is how combinations of EF and language skills at preschool entry facilitate the development of emotion regulation and behavior regulation over the course of the preschool year for children from low-income backgrounds. Understanding these interactive effects can both help to identify children at higher risk of poor regulation in preschool classrooms due to low incoming EF and/or language skills and inform targeted interventions that promote these children's regulatory abilities in tandem with facilitating relevant cognitive skillsets.

The Present Study

The goal of the present study is to build on previous work linking executive functioning and language skills to children's ability to regulate themselves in the preschool classroom. It is important to examine these relationships specifically for children from low-income backgrounds because SES disparities in EF (Jones et al., 2016; Raver et al., 2013), language (Durham, Farkas, Hammer, Tomblin, & Catts, 2007; Fernald et al., 2013; Huttenlocher, Waterfall, Vasilyeva, Vevea, & Hedges, 2010), behavior regulation (Miech et al., 2001), and emotion regulation (Keiley, Bates, Dodge,

& Petit, 2000; Raver, 2004) begin early on and have negative implications for later development (Masten & Cicchetti, 2010). As such, the present study investigated the following research questions in a primarily low-income sample.

(1) To what extent do children's executive functioning and language skills at preschool entry independently facilitate the development of their emotion and behavior regulation over the course of the preschool year?

(2) To what extent does the relationship between children's executive functioning skills and gains in regulation depend on language skills? That is, do language skills at preschool entry act as a buffer for regulation development in children who enter preschool with low EF skills?

Based on prior literature, it is hypothesized that higher EF skills at preschool entry will be related to gains in both emotion and behavior regulation during preschool. It is expected that initial language skills will also promote emotion and behavior regulation. Given the stronger empirical evidence for the link between language and emotion regulation, it is hypothesized that language will be more strongly linked with emotion regulation than with behavior regulation. Finally, we expect that language will be especially protective for the ability of children entering preschool with low EF skills to regulate their behaviors and emotions, because language will help them to better govern their behaviors through self-talk (Vygotsky, 1962) and use their words in intense emotional situations to make their wants and needs known (Bierman & Erath, 2006; Cole et al., 2010; Hay et al., 2004).

Method

Sample

Data for the present study were collected as part of an observational study of children's experiences from preschool through kindergarten. Data were collected from fall 2016 to spring 2017. The study was conducted in 15 public preschools and Head Start centers across 2 sites in the southeastern United States. Teachers were eligible for participation if they served children who would matriculate into kindergarten during the 2017-18 school year. In inclusion classrooms, the general education teacher was selected for participation. Fifty teachers from those who consented were randomly selected to participate in the study; one child transitioned to a different study classroom during the year, so a total of 51 teachers participated. Children in participating classrooms were eligible for the study. Up to 8 consented children per classroom were randomly selected to participate after blocking by gender. When fewer than 8 children were consented, all consented children in the classroom were chosen. A total of 380 preschoolers participated in the project and comprised the analytic sample for the present study.

Teachers were 98.04% female with a mean age of 44.59 years (*SD* = 10.43, range = 23 - 63); 72.55% were White/Caucasian, 23.53% were Black/African American, and 3.92% were Hispanic/Latino. Teachers had on average 16.22 years of experience at their current facility (*SD* = 8.46, range: 1 - 32 years); 54.90% had a master's degree or higher. Approximately 94 percent of classrooms were in public schools, and 6% were housed in Head Start centers. Class sizes ranged from 16 to 21 children, with a mean of 18 (*SD* = .83). On average, 49.70% of children per class were female, 3% had limited English proficiency (LEP; *SD* = 4.61%) and 6.12% had an IEP or IFSP (*SD* = 11.32%). Approximately 83% of children in a given classroom qualified for free or reduced-price lunch (*SD* = 26.33%).

Across these 50 classrooms, 380 eligible and consented children participated in the present study. Table 1 provides descriptive information for the sample. Approximately 50% of children were male; 53.93% percent were Black/African American, 24.39% were White/Caucasian, 9.76% were Hispanic/Latino, and 11.92% were of other racial or ethnic backgrounds (e.g., Asian, American Indian or Alaska Native) or multiple races. English was not the primary language for 2.70% of children. At the beginning of preschool, children had a mean age of 52.50 months (*SD* = 3.72), and their mothers had on average 13.40 years of education (*SD* = 1.74, range: 8 - 20). Families had an average income-to-needs ratio of 1.46 (*SD* = 1.08, range = 0.07 – 4.99), meaning that the average family in the study had an annual income at 146% of the federal poverty level for their household size (e.g., $ 35,478 for a family of 4; U.S. Health and Human Services Department, 2016). Two-thirds of the sample (75.30%) was considered low-income (i.e., income-to-needs ratio less than 2).

Procedure

Recruitment. Program administrators in 2 urban areas in the southeastern United States were contacted to participate in the study, and administrators and teachers were invited to attend study recruitment sessions. Interested and eligible teachers gave informed consent, completed personal and classroom demographic surveys, and allowed data collectors to observe their classrooms. Parents or guardians of children in participating classrooms were given a letter explaining the study, an informed consent, and a demographic survey. Of consented children, up to 8 from each classroom were randomly selected to participate in the study.

Observation training and protocol. Data collectors attended a two-day training session for the *Individualized Classroom Assessment Scoring System* (inCLASS; Downer et al., 2010), an observational measure of individual children's engagement. Training involved a detailed review of all measure dimensions and a process of watching, coding, and discussing five training videos. To achieve reliability and become certified on each measure, data collectors coded five videos independently and scored within one point of a master code on 80% of dimensions prior to entering classrooms. Across the year, they continued to code reliably (i.e., within one of the master code) on 88.34% of inCLASS dimensions across five recalibration/drift segments.

For live classroom observations, data collectors alternated observation and coding cycles for the inCLASS and a classroom-level measure of teacher-child interactions across two days. Data collectors observed individual children for 10 minutes using the standardized inCLASS protocol and coded for 5 minutes. Approximately six observation cycles were collected for each child per data collection window; cycles within a window typically occurred across one or two days. During data collection, 20 percent of field observations were double coded. Interrater reliability, calculated using intraclass correlation coefficients, was acceptable for the inCLASS Behavior Control dimension across the year (fall ICC = .80; spring ICC = .79).

Data collection. Parents and teachers completed demographic surveys in the fall of preschool. Trained data collectors administered direct child assessments in the fall, winter, and spring. Another set of trained data collectors observed children's engagement in the classroom in the fall, winter, and spring. During each window, six 10-minute cycles were collected for each child across one to two days. Teachers completed ratings of children's behavior in the fall and spring.

Measures

Executive functioning. This construct was assessed using a computer-administered battery and two interactive tasks to measure multiple aspects of EF. *EF Touch* is a computer-administered battery of executive functioning tasks that has been shown to have good reliability and validity in an early childhood sample (Willoughby, Blair et al., 2012; Willoughby, Pek, Blair, & Family Life Project Investigators, 2013). Three subtests from the battery were used. On the *Animal Go/No-Go* (Pig) task, which measures inhibitory control, farm animals flashed quickly on the screen and children were asked to touch all of the animals except for the pig. *Pick the Picture* (PtP) is a task of working memory during which children were asked to consistently choose pictures from a set that they have not chosen before, holding in mind those they had already picked. On *Something's the Same* (StS), which assesses set shifting, the screen displayed two pictures that were similar on one dimension (e.g., color, size), and then added another picture that was the same as one of the first pictures in a different way. Children were asked to choose which of the first two pictures was similar to the new picture (e.g., shift to think about similarity on a different dimension). Proportion correct was calculated for each of the three subtests.

Head-Toes-Knees-Shoulders (HTKS) is a measure of EF that requires inhibitory control, working memory, and attention focusing (Ponitz et al., 2009). On this task, children were asked to learn simple commands (i.e., "touch your head," "touch your toes") then do the opposite of what the assessor said (e.g., touch their toes when asked to touch their head). An advanced trial incorporating the same task with knees and shoulders was given to children who did not reach a ceiling on the first set of items. Children received two points for each correct response and one point for a self-corrected response; scores ranged from 0 to 60 across 30 trials. HTKS has been shown to have adequate reliability and concurrent and predictive validity in a preschool sample (Ponitz et al., 2009).

Inhibitory control was also measured with the Pencil Tap subtest of the *Preschool Self-Regulation Assessment* (PSRA; Smith-Donald, Raver, Hayes, & Richardson, 2007). In this assessment, children were asked to tap their pencil once when the examiner tapped twice and twice when the examiner tapped once. Scores represent percentage of correct responses. This test shows acceptable concurrent and construct validity (Smith-Donald et al., 2007).

In fall of preschool, EF measures (EF Touch proportion correct for each subtest, HTKS raw score, and Pencil Tap proportion correct) were significantly correlated with one another (see Table 2). Based on a confirmatory factor analysis, the five EF measures fit very well as a single factor (CFI = .97, TLI = .95, RMSEA = .05, SRMR = .03), and each of the subtests had significant standardized factor loadings (EF Touch Pig = .46, EF Touch PtP = .42, EF Touch StS = .41, HTKS = .64, Pencil Tap = .68). Therefore, to represent an overall measure of EF in this study, scores from these separate measures were standardized within the sample and averaged to create a composite EF score.

Language. Expressive vocabulary was measured by the Picture Vocabulary subtest of the *Woodcock-Johnson Test of Achievement – Third Edition* (WJ-III PV); this measure had good internal consistency reliability in the original sample (McGrew &

Woodcock, 2001). Children were asked to name objects in a series of pictures. Receptive vocabulary was measured using the *Peabody Picture Vocabulary Test – Fourth Edition* (Dunn & Dunn, 2007), which has been shown to have good reliability and validity in early childhood (Community-University Partnership for the Study of Children, Youth, and Families, 2011). Children were asked to point to one of four pictures corresponding to an orally presented word. Standard scores of both assessments were used. In this sample, the WJ-III PV and the PPVT-4 were significantly correlated at .50 in fall of preschool (see Table 2). The two measures were standardized on the sample and averaged to create a composite language score.

Behavior regulation. This construct was assessed using a teacher rating and an observational measure. The *Teacher Child Rating Scale* (TCRS 2.1; Hightower & Perkins, 2010) is a 32-item teacher report of children's behaviors in the classroom. The Behavior Control subscale used in this study includes 8 items, such as "accepts imposed limits," "tolerates frustration," and "disruptive in class" (reverse-coded). In fall and spring, teachers answered items using a 5-point Likert scale ("strongly disagree" to "strongly agree"). This measure has been shown to have good internal consistency reliability ($\alpha = .90$) and test-retest reliability over seven months ($r = .70$) in children from preschool to eighth grade (Hightower & Perkins, 2010). In this sample, Behavior Control had good internal consistency reliability (fall $\alpha = .87$; spring $\alpha = .88$). The measure has differentiated children who were and were not previously identified as being at risk for problem behaviors and school difficulties, demonstrating its content and criterion-related validity (Hightower & Perkins, 2010).

The *Individualized Classroom Assessment Scoring System* (inCLASS; Downer et al., 2010) is an observational measure of individual children's engagement with teachers, peers, and tasks in the classroom. The measure has been shown to have good inter-rater and internal consistency reliability (Vitiello et al., 2012; Williford, Maier, Downer, Pianta, & Howes, 2013), construct and criterion-related validity (Downer et al., 2010), and predictive validity (Williford, Whittaker, Vitiello, & Downer, 2013) in early childhood. It has been shown to have good concurrent validity with children's teacher-reported self-awareness, student-teacher closeness and conflict, self-regulation, and engagement in a primarily African American, low-income sample of kindergarteners (Kim et al., 2018). The inCLASS has been shown to have similar measurement properties across demographic groups (i.e., strong invariance for poverty status and ethnicity, and configural invariance for gender; Bohlmann et al., 2019). Only one dimension (Behavior Control) of the Task domain was used for the purposes of this study, given its relevance to the research questions. This assessed the degree to which children match expectations in the classroom (e.g., with their volume and movement) and demonstrate patience and physical awareness. The dimension is scored on a 1-7 scale from "low" to "high," where 1-2 represent low quality, 3-5 represent mid-range quality, and 6-7 represent high quality. Children's scores were averaged across 6 cycles each in fall and spring.

Emotion regulation. The *Emotion Regulation Checklist* (Shields & Cicchetti, 1997) is a 24-item teacher-report measure. The Emotion Regulation subscale used in this study includes 8 items, such as "is a cheerful child," "responds positively to neutral or friendly overtures by adults or peers," and "displays appropriate negative emotions." In fall and spring, teachers answered items using a 4-point Likert scale ("rarely/never" to "almost always"). In this sample, Emotion Regulation had adequate internal consistency reliability (fall alpha = .80, spring alpha = .82). This measure's construct validity is supported by the fact that it adequately differentiated between well regulated and dysregulated children ages 6 to 12 (Shields & Cicchetti, 1997). It has demonstrated predictive validity in that it has been associated with higher academic achievement and productivity in the classroom in community-based samples of racially and economically diverse 4- and 5-year-olds (Graziano, Reavis, Keane, & Calkins, 2007; Howse, Calkins, Anastopoulos, Keane, & Shelton, 2003).

Covariates. In fall of preschool, families provided information about child and family characteristics. Annual income and household size were used to calculate the income-to-needs ratio, which is a family's income relative to the poverty level for their household size. Preschool teachers provided information about themselves in fall of preschool. The following variables were included as covariates in analyses: child's age (in months) and dummy codes for race/ethnicity (White/Caucasian, Hispanic/Latino, and other race, with Black/African American omitted as reference category), sex (male = 1), and primary language (non-English = 1); continuous variables for years of maternal education and family income-to-needs ratio; and teacher years of experience (continuous) and education level (master's = 1, compared to less than a master's).

Data Analysis

Missing data were as follows: 0% child sex and age, 2.89% child race/ethnicity, 2.37% child primary language, 3.16% years of maternal education, 11.58% income-to-needs ratio; 2.11% years of teacher experience, 0% teacher education level; 2.11% fall EF Touch, 1.84% fall HTKS, 1.58% fall Pencil Tap, 0.79% fall WJ-III Picture

Vocabulary, 7.89% fall PPVT-4, 1.32% fall inCLASS, 6.58% spring inCLASS, 0.79% fall Emotion Regulation Checklist, 6.05% spring ERC, 1.05% fall Teacher Child Rating Scale, and 6.05% spring TCRS. Multilevel multiple imputation was conducted to retain cases missing data on study variables; 10 data sets were imputed using Blimp 1.1 (Keller & Enders, 2017). Two-level imputation accounted for clustering of students within classrooms.

Intraclass correlation coefficients (ICCs) from unconditional regression models for each outcome were examined to determine the need to control for nestedness of children within classrooms (ICC > .10; Raudenbush & Bryk, 2002). ICCs for the three outcomes (.15 to .39) indicated the need for mixed effects multilevel (two-level) regression clustered by preschool classroom. All analyses were run on the 10 multiply imputed data sets using "mi estimate" and "xtmixed" commands. Models were fit using maximum likelihood, using the specification "mle."

To examine the independent contributions of fall executive functioning skills and language skills to the development of behavior and emotion regulation over the preschool year, three mixed effects multilevel regression models were fit with each of the three outcomes regressed on fall executive functioning and fall language abilities. All models included child, family, and teacher covariates, as well as fall controls of relevant outcomes. To examine interactive contributions (i.e., moderation), an interaction term (EF*language) was created. In the final three models, the three outcomes were regressed separately on this interaction term, EF and language predictors, and covariates.

Results

Table 3 provides results of mixed effects multilevel regression models. Base models include covariates and main effects of executive functioning and language skills on each of the three spring outcomes separately. Final models include all aforementioned variables in addition to an interaction term (EF*language).

In base models, one significant association emerged. Children with higher language skills at preschool entry were reported by teachers as having higher emotion regulation by spring of the preschool year (B = 1.10, $p = .018$), even when controlling for children's fall emotion regulation and executive functioning skills. Executive functioning was not significantly related to spring emotion regulation (B = -0.29, $p = .58$). Neither executive functioning skills nor language skills in the fall were associated with either observed (EF B = -0.01, $p = .89$; language B = 0.01, $p = .85$) or teacher-reported behavior regulation (EF B = -0.07, $p = .85$; language B = 0.02, $p = .94$).

In final models, the EF*language interaction was significantly associated with the development of teacher-reported emotion regulation (B = -1.02, $p = .011$; see Figure 1). Specifically, when children entered preschool with high EF skills, their development of emotion regulation over the course of the year was reported to be similar regardless of their language ability. Conversely, when children entered preschool with low EF skills, their emotion regulation development over the year was dependent on their language skills. Children with low EF skills but high language skills at preschool entry were rated as having significantly more emotion regulation development over the year than children with both low EF and low language skills in the fall. That is, entering preschool with high language skills was a protective factor for children with relatively low concurrent EF skills to develop their ability to regulate their emotions in the classroom. The main effect association between language and emotion regulation was maintained in this final model

(B = 1.01, *p* = .03). EF*language was not significantly related to either observed (B = -0.02, *p* = .68) or teacher-reported behavior regulation (B = -0.25, *p* = .39).

Discussion

The present study sought to better understand the independent and interactive contributions of preschoolers' executive functioning and language skills at preschool entry to their development of emotion regulation and behavior regulation during preschool. Understanding what facilitates these regulatory skills is especially important for children from low-income backgrounds, as they tend to enter school with lower skills in these domains than more affluent peers (Durham et al., 2007; Keiley et al., 2000; Miech et al., 2001; Raver et al., 2013). Consistent with our hypothesis, language was associated with gains in children's ability to regulate their emotions across the year, especially for children at risk due to low EF skills. Specifically, in this study of a primarily low-income, racially and ethnically diverse sample of preschoolers, language was both a promotive and a protective factor for the development of emotion regulation in preschool. Unexpectedly, EF at preschool entry was not associated with regulatory outcomes in this study, and fall language ability was not significantly related to the development of behavior regulation. Findings are discussed below in the context of relevant literature, and limitations and directions for future research are offered.

Early Language May be Promotive and Protective for Preschoolers' Emotion Regulation

Consistent with reviews of prior literature (Bierman & Erath, 2006; Cole et al., 2010) and research with predominantly White children from low- to mid-SES backgrounds (Roben et al., 2013), this study found that language was related to children's development of emotion regulation in early childhood. Specifically, the expressive and receptive language skillset with which children enter preschool was associated with gains in their emotion regulation over the course of the year, as rated by their teacher. However, findings did not provide evidence that language in preschool enables children to regulate their behaviors with regard to matching expectations in the classroom. The finding that language was associated with emotion regulation development in preschool is particularly important, given that there tends to be increased conflict and intense emotions during this year because children are learning to engage effectively with many new peers and adults in an unfamiliar context (Hay et al., 2004; Roben et al., 2013).

A possible practical implication of this main finding is that facilitating children’s language development in the classroom may indirectly promote children’s emotion regulation. Children in classrooms where there are higher levels of observed instructional support have been found to have more gains in language development in both racially and ethnically diverse, low-income (Gosse, McGinty, Mashburn, Hoffman, & Pianta, 2014) and mixed-income samples (Mashburn et al., 2006). Thus, early childhood teachers can promote children’s language skills by using instructionally supportive strategies, including providing children verbal feedback, frequently modeling rich language, and using open-ended and follow-up questions to engage children’s development of concepts (La Paro, Pianta, & Stuhlman, 2004). Future research might examine how the relationship between children’s language skills and their development of emotion regulation depends on the frequency and quality with which teachers provide emotional and instructional support, both generally and specifically to help children understand, reflect on, think critically about, and regulate their emotions (Bierman, Domitrovich et al., 2008; Hamre et al., 2013). For instance, preschool teachers' use of language to further emotion-related skills might include encouraging children to recognize, understand, label, express, and regulate their emotions, as well as asking them to monitor and reflect on what they could have done instead or what they might do next time in an emotionally tense situation (Nathanson, Rivers, Flynn, & Brackett, 2016). There is growing recognition around the notion that preschool is a key setting to begin developing these foundational emotional skills, but more work is needed in this area (Hoffman, Ivcevic, & Brackett, 2018). Investigating the role of classroom processes in the relationship between language and emotion regulation would provide a more comprehensive picture of factors influencing children's skill development by incorporating context in conjunction with children's personal characteristics and resources (Bronfenbrenner & Morris, 2006).

The second main finding of this study was that not only is language directly related to the positive development of emotion regulation during preschool, but it also may be protective specifically for children who enter preschool with low executive functioning skills. This provides further empirical support for the theory that foundational abilities do not develop independently of one another in early childhood but rather are developmentally intertwined and facilitate one another (Bierman, Torres, Domitrovich, Welsh, & Gest, 2009; Bohlmann, Maier, & Palacios, 2015; Nix, Bierman, Domitrovich, & Gill, 2013). Practically, this suggests that language might be a meaningful target of interventions specifically for young children at risk of emotional and behavioral difficulties (Salmon, O'Kearney, Reese, & Fortune, 2016), a vulnerable subpopulation that includes children with low EF skills (Craig et al., 2016; Rosenthal et al., 2013; Zelazo et al., 2016). As such, interventions that have a dual focus of developing academic skills (e.g., language) and social-emotional learning might be protective and preventative for preschoolers identified as having low EF skills (Bierman, Nix, al., 2008) by providing this at-risk population the tools needed to learn how to better regulate their emotions before academic, social, emotional, and/or behavioral problems develop.

Existing interventions that explicitly integrate language-learning and social competence (e.g., EF, regulation) goals to serve the needs of low-income children include Tools of the Mind (Bodrova & Leong, 2006) and the Head Start Research-based, Developmentally Informed (REDI) intervention (Bierman, Domitrovich et al., 2008). Tools of the Mind – which seeks to facilitate self-regulation, literacy, and math through a Vygotskian approach focused on play and social interactions – has been found to be effective in boosting the executive functioning and, to a lesser degree, language skills of low-income, racially, ethnically, and linguistically diverse children (Barnett et al., 2008; Bodrova & Leong, 2006). REDI seeks to enhance children's social-emotional competence and language and literacy skills by (1) helping teachers use positive behavioral management strategies and improve their use of language in the classroom, and (2) integrating dialogic reading and activities from a social-emotional skills curriculum (Promoting Alternative Thinking Strategies; PATHS; Domitrovich, Cortes, & Greenberg, 2007) into daily classroom routines (Bierman, Domitrovich et al., 2008). This intervention has been shown to effectively facilitate skills across multiple domains for children from low-income backgrounds in the short (Nix et al., 2013) and long term (Nix et al., 2016).

Understanding Unexpected, Non-significant Findings

A final key point to address is the non-significant findings that were contrary to hypotheses. Although language was consistently associated with children's emotion regulation, it was not linked to development in either observed or teacher-reported behavior regulation. It was expected that language would be more strongly related to emotion regulation because of ample studies with findings on this association (e.g., Cole et al., 2010; Roben et al., 2013); however, there was some conceptual (Vygotsky, 1962) and empirical evidence (Winsler et al., 2007) for the relationship between language and behavior regulation.

It is possible that in order to promote behavior regulation – operationalized in this study as matching expectations in the classroom, not disturbing or disrupting others' space and activities, and demonstrating patience when encountering frustrations – language needs to be specifically focused on the tasks at hand. For example, Winsler and colleagues (2003) found in a predominantly White, mixed-income sample that when young children's speech was irrelevant to the task at hand, they engaged in fewer goal-directed behaviors and were rated as having more behavior problems. In other words, a child talking aloud to himself about a task in which he is engaged (e.g., when completing a puzzle, saying "this one goes here, and this one has a pointy side") is conceptually more likely to keep his attention on the task and facilitate his ability to manage expectations in the classroom than if he is talking to himself about what he is going to have for snack or what his peers across the room are doing. The present study examined general expressive and receptive language skills but was not able to include measures of language in the context of certain tasks. It is also plausible that language is less necessary for task-related behavior regulation as it was operationalized in this study (i.e., controlling behavior to match expectations) than for more verbal, interaction-driven emotion regulation. However, these explanations are speculative and more research is needed to better understand the relationship between language and behavior regulation given this unexpected, non-significant finding.

EF, the other key cognitive predictor, was not significantly related to any regulatory outcome, which was surprising given myriad studies linking EF and regulation in early childhood (Hay et al., 2004; Jones et al., 2016) in both relatively high-SES (Riggs, et al., 2004) and low-SES (Nesbitt et al., 2015; Sasser et al., 2015) samples. It is possible that there was not enough variation in the measures of observed (M = 6.16 and SD = .76 on a 1-7 scale) and teacher-reported (M = 30.78 and SD = 7.35 on an 8-40 scale) behavior regulation, and/or that there was not much variance left to be accounted for after fall regulation scores and fall language were included in models. The negative skew and relatively low variability in the behavior regulation measures admittedly represent a measurement limitation that is an important consideration when interpreting results of this study. However, this does not explain the lack of association between EF and emotion regulation, particularly because there were interactive contributions of language *and* EF to emotion regulation development. Multiple explanations were considered as to why this association was not found.

In this study, EF was conceptualized as an overarching composite measure based on a plethora of prior work indicating that it is best understood as a single construct in early childhood (Wiebe et al., 2008; Wiebe et al., 2011; Willoughby et al., 2012; Zelazo et al., 2016) that differentiates into distinct but related subcomponents with age (Lehto et al., 2003; Miyake et al., 2000). It is possible, however, that certain aspects of EF, such as inhibitory control, could be more salient for emotion regulation than others, such as attention. Future work might explore this possibility by conducting more nuanced examinations of associations between aspects of EF and emotion regulation. It could be that EF is necessary but not sufficient for children to regulate their behaviors and emotions (Ackerman & Friedman-Krauss, 2017; Jones et al., 2016). Other emerging skills, including but not limited to language, may be required for children to effectively employ their developing EF skills to regulate their behaviors and emotions in the context of the classroom. A final consideration is that a teacher-report questionnaire was used as the sole measure of emotion regulation. Future work might incorporate multi-informant assessments of this construct.

Despite the fact that they were contrary to hypotheses, these non-significant associations between EF and behavior and emotion regulation add to the literature because they indicate that relationships among these constructs remain mixed and require further investigation (Jones et al., 2016). Additional research is needed to better understand linkages between EF and children's abilities to regulate their behaviors and emotions in the early childhood classroom.

Limitations and Future Directions

The present study is observational and correlational; thus, causal claims cannot be made about the developmental pathways between children's EF and language at preschool entry and gains in behavior and emotion regulation across the year. This study also did not test longitudinal, bidirectional relationships between language and executive functioning in preschool, as they relate to the development of behavior and emotion regulation. Future work should continue to examine bidirectional relationships between these two constructs to better understand which is a leading factor of the other in early childhood, given mixed results across various studies with predominantly White, mixed-income (Hughes et al., 2009) and racially, ethnically, linguistically, and financially diverse (Weiland, Barata, & Yoshikawa, 2014) samples.

Multi-informant measures of behavior regulation were used in this study; however, only teacher report of emotion regulation was available. In future work, it would be useful to capture emotion regulation in multiple ways, such as observation (Roben et al., 2013) or physiological measures (Eisenberg & Sulik, 2012), to investigate whether associations found in this study are maintained across multi-informant measures of emotion regulation (Eisenberg et al., 2010). In this study, language was operationalized as children's general expressive and receptive vocabularies. It would be beneficial for future work to examine language more specifically as it relates to emotions (e.g., emotion identification, labeling, scaling, self-talk) and tasks in the classroom. This would both help to unpack the importance of language in emotion regulation development found in this study, and provide insight into whether language needs to be more task-relevant to enhance behavior regulation in a given context (Winsler et al., 2003).

Finally, this study focused exclusively on how characteristics with which children enter preschool (i.e., language, EF) are associated with their development of regulation. Future research should expand on this by examining the role of classroom characteristics, such as the quality of teachers' instructional support, and match between teacher and child characteristics (e.g., sex, race/ethnicity) in facilitating the development of these skills.

Conclusion

Despite its limitations, the present study adds to the current literature with the finding that language skills at preschool entry are promotive *and* protective for gains in children's emotion regulation in the preschool classroom. This is a key time period for development of emotion regulation because conflict can occur frequently in the classroom as children learn to engage with new peers who are also inexperienced conversational partners (Hay et al., 2004). Moreover, having well-developed emotion regulation prior to the transition to kindergarten is important, because the shift to elementary school is typically accompanied by increases in demands and expectations on children's regulatory abilities (Bassok et al., 2016; Rimm-Kaufman & Pianta, 2000).

A main contribution of the present study was the finding that language acts as a protective factor for children with low EF in learning to regulate their emotions. The unexpected finding of a lack of association between EF and regulation speaks to the need for further research on these related but distinct skills in early childhood. Practically, results suggest that incorporating language-building activities into early childhood interventions designed to facilitate self-regulation may enhance efficacy, particularly for children at risk for later emotional, academic, and behavioral difficulties due to low emerging EF skills.

References

Ackerman, D. J., & Friedman-Krauss, A. H. (2017). Preschoolers' executive function: Importance, contributors, research needs and assessment options. *ETS Research Report Series*, *2017*, 1-24. doi: 10.1002/ets2.12148.

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

Barkley, R. A. (1989). The problem of stimulus control and rule-governed behavior in attention deficit disorder with hyperactivity. In L. M. Bloomingdale, & J. M. Swanson (Eds.), *Current concepts and emerging trends in attentional and behavioral disorders of childhood* (pp. 203-233). Elmsford, NY: Pergamon Press, Inc.

Barnett, W. S., Jung, K., Yarosz, D. J., Thomas, J., Hornbeck, A., Stechuk, R., & Burns, S. (2008). Educational effects of the Tools of the Mind curriculum: A randomized trial. *Early Childhood Research Quarterly*, *23*, 299-313. doi: 10.1016/j.ecresq.2008.03.001.

Bassok, D., Latham, S., & Rorem, A. (2016). Is kindergarten the new first grade?. *AERA Open*, *2*, 2332858415616358. doi: 10.1177/2332858415616358.

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

Bierman, K. L., Domitrovich, C. E., Nix, R. L., Gest, S. D., Welsh, J. A., Greenberg, M. T., ... & Gill, S. (2008). Promoting academic and social-emotional school readiness: The Head Start REDI program. *Child Development*, *79*, 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. A. Phillips (Eds.), *Blackwell handbook of early childhood development* (pp. 595-615). New York, NY: Guilford Press.

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

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*, 305-323. doi: 10.1111/j.1467-9507.2008.00490.x.

Blair, C. (2016). Executive function and early childhood education. *Current Opinion in Behavioral Sciences*, *10*, 102-107. doi: 10.1016/j.cobeha.2016.05.009.

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*, (Vol. 2, pp. 300-320). New York, NY: Guilford.

Blair, C., Zelazo, P. D., & Greenberg, M. T. (2005). The measurement of executive function in early childhood. *Developmental Neuropsychology, 28,* 561-571. doi: 10.4324/9780203764244.

Bodrova, E., & Leong, D. J. (2006). *Tools of the Mind*. Upper Saddle River, NJ: Pearson

Merrill Prentice Hall.

Bohlmann, N. L., Downer, J. T., Williford, A. P., Maier, M. F., Booren, L. M., & Howes, C. (2019). Observing children's engagement: Examining factorial validity of the inCLASS across demographic groups. *Journal of Applied Developmental Psychology*, *60*, 166-176. doi: 10.1016/j.appdev.2018.08.007.

Bohlmann, N. L., Maier, M. F., & Palacios, N. (2015). Bidirectionality in self-regulation and expressive vocabulary: Comparisons between monolingual and dual language learners in preschool. *Child Development*, *86*, 1094-1111. doi: 10.1111/cdev.12375.

Booren, L. M., Downer, J. T., & Vitiello, V. E. (2012). Observations of children's interactions with teachers, peers, and tasks across preschool classroom activity settings. *Early Education & Development*, *23*, 517-538. doi: 10.1080/10409289.2010.548767.

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*, 337-349. doi: 10.1016/j.ecresq.2009.06.001.

Bronfenbrenner, U. & Morris, P. A. (2006). The bioecological model of human development. In R. M. Lerner (Ed.). *Theoretical models of human development: Handbook of child psychology* (6th ed., Vol. 1). Hoboken, NJ: Wiley.

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

Cicchetti D., Ganiban J., & Barnett D. (1991). The development of emotion regulation. In

K. Dodge, & J. Garber (Eds.), *Contributions from the study of high-risk populations to understanding the development of emotion regulation* (pp. 15-48). New York, NY: Cambridge University Press.

Cole, P. M., Armstrong, L. M., & Pemberton, C. K. (2010). The role of language in the development of emotion regulation. In S. D. Calkins & M. A. Bell (Eds.), *Human brain development: Child development at the intersection of emotion and cognition* (pp. 59-77). Washington, DC: American Psychological Association.

Community-University Partnership for the Study of Children, Youth, and Families (2011). *Review of the Peabody Picture Vocabulary Test, Fourth Edition (PPVT-4).* Retrieved from: https://cloudfront.ualberta.ca/-/media/ualberta/faculties-and-programs/centres-institutes/community-university-partnership/resources/tools---assessment/ppvt-4may-2012.pdf.

Craig, F., Margari, F., Legrottaglie, A. R., Palumbi, R., De Giambattista, C., & Margari, L. (2016). A review of executive function deficits in autism spectrum disorder and attention-deficit/hyperactivity disorder. *Neuropsychiatric Disease and Treatment*, *12*, 1191-1202. doi: 10.2147/NDT.S104620.

Domitrovich, C. E., Cortes, R. C., & Greenberg, M. T. (2007). Improving young children's social and emotional competence: A randomized trial of the preschool "PATHS" curriculum. *The Journal of Primary Prevention*, *28*, 67-91. doi: 10.1007/s10935-007-0081-0.

Downer, J. T., Booren, L. M., Lima, O. K., Luckner, A. E., & Pianta, R. C. (2010). The Individualized Classroom Assessment Scoring System (inCLASS): Preliminary reliability and validity of a system for observing preschoolers' competence in classroom interactions. *Early Childhood Research Quarterly*, *25*, 1-16. doi: 10.1016/j.ecresq.2009.08.004.

Duckworth, A. L., & Carlson, S. M. (2013). Self-regulation and school success. In B. W. Sokol, F. M. E. Grouzet, & U. Muller. (Eds.), *Self-regulation and autonomy: Social and developmental dimensions of human conduct* (pp. 208-230). New York, NY: Cambridge University Press.

Dunn, L. M., & Dunn, D. M. (2007). *PPVT-4 manual.* Bloomington, MN: NCS Pearson.

Durham, R. E., Farkas, G., Hammer, C. S., Tomblin, J. B., & Catts, H. W. (2007). Kindergarten oral language skill: A key variable in the intergenerational transmission of socioeconomic status. *Research in Social Stratification and Mobility*, *25*, 294-305. doi: 10.1016/j.rssm.2007.03.001.

Eisenberg, N., Hofer, C., & Vaughan, J. (2007). Effortful control and its socioemotional consequences. In J. J. Gross (Ed.), *Handbook of Emotion Regulation* (Vol. 2, pp. 287-306). New York, NY: Guilford.

Eisenberg, N., Spinrad, T. L., & Eggum, N. D. (2010). Emotion-related self-regulation and its relation to children's maladjustment. *Annual Review of Clinical Psychology*, *6*, 495-525. doi: 10.1146/annurev.clinpsy.121208.131208.

Eisenberg, N., & Sulik, M. J. (2012). Emotion-related self-regulation in children. *Teaching of Psychology*, *39*, 77-83. doi: 10.1177/0098628311430172.

Fernald, A., Marchman, V. A., & Weisleder, A. (2013). SES differences in language processing skill and vocabulary are evident at 18 months. *Developmental Science*, *16*, 234-248. doi: 10.1111/desc.12019.

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

Gooch, D., Thompson, P., Nash, H. M., Snowling, M. J., & Hulme, C. (2016). The development of executive function and language skills in the early school years. *Journal of Child Psychology and Psychiatry*, *57*, 180-187. doi: 10.1111/jcpp.12458

Gosse, C. S., McGinty, A. S., Mashburn, A. J., Hoffman, L. M., & Pianta, R. C. (2014). The role of relational and instructional classroom supports in the language development of at-risk preschoolers. *Early Education & Development*, *25*, 110-133. doi: 10.1080/10409289.2013.778567.

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*, 3-19. doi: 10.1016/j.jsp.2006.09.002.

Halliday, M. A. K. (2006). *Language of early childhood* (Vol. 4). New York, NY: Continuum.

Hamre, B. K., Pianta, R. C., Downer, J. T., DeCoster, J., Mashburn, A. J., Jones, S. M., ... & Brackett, M. A. (2013). Teaching through interactions: Testing a developmental framework of teacher effectiveness in over 4,000 classrooms. *The Elementary School Journal*, *113*, 461-487. doi: 10.1086/669616.

Hay, D. F., Payne, A., & Chadwick, A. (2004). Peer relations in childhood. *Journal of Child Psychology and Psychiatry, 45*, 84-108. doi: 10.1046/j.0021-9630.2003.00308.x.

Herndon, K. J., Bailey, C. S., Shewark, E. A., Denham, S. A., & Bassett, H. H. (2013).

Preschoolers' emotion expression and regulation: Relations with school adjustment. *The Journal of Genetic Psychology*, *174*, 642-663. doi: 10.1080/00221325.2012.759525.

Hightower, A. D. (1986). The Teacher–Child Rating Scale: A brief objective measure of elementary children's school problem behaviors and competencies. *School Psychology Review 15*, 393-409.

Hightower, A. D., & Perkins, P. E. (2010). *T-CRS 2.1 Teacher Child Rating Scale: Examiner's manual*. Rochester, NY: Children's Institute.

Hoffmann, J. D., Ivcevic, Z., & Brackett, M. A. (2018). Building emotionally intelligent schools: From preschool to high school and beyond. In K. V. Keefer, J. D. A. Parker, & D. H. Saklofske (Eds.), *Emotional intelligence in education: Integrating research with practice* (pp. 173-198). Cham, Switzerland: Springer.

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

Howse, R. B., Calkins, S. D., Anastopoulos, A. D., Keane, S. P., & Shelton, T. L. (2003). Regulatory contributors to children's kindergarten achievement. *Early Education and Development*, *14*, 101-120. doi: 10.1207/s15566935eed1401_7.

Hughes, C., Ensor, R., Wilson, A., & Graham, A. (2009). Tracking executive function across the transition to school: A latent variable approach. *Developmental Neuropsychology*, *35*, 20-36. doi: 10.1080/87565640903325691.

Huttenlocher, J., Waterfall, H., Vasilyeva, M., Vevea, J., & Hedges, L. V. (2010). Sources of variability in children's language growth. *Cognitive Psychology*, *61*, 343-365. doi: 10.1016/j.cogpsych.2010.08.002.

Jones, S. M., Bailey, R., Barnes, S. P., & Partee, A. (2016). *Executive function mapping project: Untangling the terms and skills related to executive function and self-regulation in early childhood* (No. 2016-88). Cambridge, MA: U.S. Department of Health and Human Services, Office of Planning, Research, and Evaluation.

Keiley, M. K., Bates, J. E., Dodge, K. A., & Pettit, G. S. (2000). A cross-domain growth analysis: Externalizing and internalizing behaviors during 8 years of childhood. *Journal of Abnormal Child Psychology*, *28*, 161-179. doi: 10.1023/A:1005122814723.

Keller, B. T., & Enders, C. K. (2017). Blimp Software Manual (Version Beta 6.7). *Los Angeles, CA*.

Kim, H., Cameron, C. E., Kelly, C. A., West, H., Mashburn, A. J., & Grissmer, D. W. (2018). Using an individualized observational measure to understand children's interactions in underserved kindergarten classrooms. *Journal of Psychoeducational Assessment*, Advance online publication. doi: 0734282918819579.

Knudsen, E. I., Heckman, J. J., Cameron, J. L., & Shonkoff, J. P. (2006). Economic, neurobiological, and behavior perspectives on building America's future workforce. *Proceedings of the National Academy of Sciences*, *103*, 10155-10162. doi: 10.1073/pnas.0600888103.

La Paro, K. M., Pianta, R. C., & Stuhlman, M. (2004). The classroom assessment scoring system: Findings from the prekindergarten year. *The Elementary School Journal*, *104*, 409-426. doi: 10.1086/499760.

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

Mashburn, A. J. (2017). Evaluating the validity of classroom observations in the Head Start Designation Renewal System. *Educational Psychologist, 52*, 38-49. doi: 10.1080/00461520.2016.1207539.

Mashburn, A. J., Hamre, B. K., Downer, J. T., & Pianta, R. C. (2006). Teacher and classroom characteristics associated with teachers' ratings of prekindergartners' relationships and behaviors. *Journal of Psychoeducational Assessment*, *24*, 367-380. doi: 10.1177/0734282906290594.

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

McClelland, M. M., Cameron, C. E., Connor, C. M., Farris, C. L., Jewkes, A. M., & Morrison, F. J. (2007). Links between behavioral regulation and preschoolers' literacy, vocabulary, and math skills. *Developmental Psychology*, *43*, 947-959. doi: 10.1037/0012-1649.43.4.947.

McClelland, M. M., Cameron, C. E., Wanless, S. B., & Murray, A. (2007). Executive function, behavioral self-regulation, and social-emotional competence: Links to school readiness. In O. N. Saracho, & B. Spodek (Eds.), *Contemporary perspectives on social learning in early childhood education* (Vol. 1, pp. 83-107). Charlotte, NC: Information Age Publishing.

McClelland, M., Geldhof, J., Morrison, F., Gestsdóttir, S., Cameron, C., Bowers, E., ... &

Grammer, J. (2018). Self-regulation. In N. Halfon, C. B. Forrest, R. M. Lerner, & E. M. Faustman (Eds.), *Handbook of life course health development* (pp. 275-298). Cham, Switzerland: Springer International.

McGrew, K. S., & Woodcock, R. W. (2001). *Woodcock-Johnson III technical manual*. Rolling Meadows, IL: Riverside Publishing.

Miech, R., Essex, M. J., & Goldsmith, H. H. (2001). Socioeconomic status and the adjustment to school: The role of self-regulation during early childhood. *Sociology of Education, 74*, 102-120. doi: 10.2307/2673165.

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*, 49-100. doi: 10.1006/cogp.1999.0734.

Moffitt, T. E., Arseneault, L., Belsky, D., Dickson, N., Hancox, R. J., Harrington, H., ... & Sears, M. R. (2011). A gradient of childhood self-control predicts health, wealth, and public safety. *Proceedings of the National Academy of Sciences*, *108*, 2693-2698. doi: 10.1073/pnas.1010076108.

Nathanson, L., Rivers, S. E., Flynn, L. M., & Brackett, M. A. (2016). Creating emotionally intelligent schools with RULER. *Emotion Review*, *8*, 305-310. doi: 10.1177/1754073916650495.

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*, 865-878. doi: 10.1037/dev0000021.

Nix, R. L., Bierman, K. L., Domitrovich, C. E., & Gill, S. (2013). Promoting children's social emotional skills in preschool can enhance academic and behavioral functioning in kindergarten: Findings from Head Start REDI. *Early Education and Development, 24*, 1000-1019. doi: 10.1080/10409289.2013.825565.

Nix, R. L., Bierman, K. L., Heinrichs, B. S., Gest, S. D., Welsh, J. A., & Domitrovich, C. E. (2016). The randomized controlled trial of Head Start REDI: Sustained effects on developmental trajectories of social–emotional functioning. *Journal of Consulting and Clinical Psychology*, *84*, 310-322. doi: 10.1037/a0039937.

Petersen, I. T., Bates, J. E., D'Onofrio, B. M., Coyne, C. A., Lansford, J. E., Dodge, K. A., ... & Van Hulle, C. A. (2013). Language ability predicts the development of behavior problems in children. *Journal of Abnormal Psychology*, *122*, 542-557. doi: 10.1037/a0031963.

Petersen, I. T., Bates, J. E., & Staples, A. D. (2015). The role of language ability and self-regulation in the development of inattentive–hyperactive behavior problems. *Development and Psychopathology*, *27*, 221-237. doi: 10.1017/S0954579414000698.

Phillips, D. A., & Shonkoff, J. P. (Eds.). (2000). *From neurons to neighborhoods: The science of early childhood development*. Washington, DC: National Academies Press.

Ponitz, C. 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*, 605-619. doi: 10.1037/a0015365.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and* *data analysis methods* (2nd ed., Vol. 1). Thousand Oaks, CA: Sage.

Raver, C. C. (2004). Placing emotional self-regulation in sociocultural and socioeconomic contexts. *Child Development*, *75*, 346-353. doi: 10.1111/j.1467-8624.2004.00676.x.

Raver, C. C., Jones, S. M., 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. doi: 10.1111/j.1467-8624.2010.01561.x.

Raver, C. 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*, 394-408. doi: 10.1111/desc.12027.

Ready, D. D., & Chu, E. M. (2015). Sociodemographic inequality in early literacy development:The role of teacher perceptual accuracy. *Early Education and Development*, *26*, 970-987. doi: 10.1080/10409289.2015.1004516.

Riggs, N. R., Blair, C. B., & Greenberg, M. T. (2004). Concurrent and 2-year longitudinal relations between executive function and the behavior of 1st and 2nd grade children. *Child Neuropsychology*, *9*, 267-276. doi: 10.1076/chin.9.4.267.23513.

Rimm-Kaufman, S. E., & Pianta, R. C. (2000). An ecological perspective on the transition to kindergarten: A theoretical framework to guide empirical research. *Journal of Applied Developmental Psychology*, *21*, 491-511. doi: 10.1016/S0193-3973(00)00051-4.

Roben, C. K., Cole, P. M., & Armstrong, L. M. (2013). Longitudinal relations among language skills, anger expression, and regulatory strategies in early childhood. *Child Development*, *84*, 891-905. doi: 10.1111/cdev.12027.

Rosenthal, M., Wallace, G. L., Lawson, R., Wills, M. C., Dixon, E., Yerys, B. E., & Kenworthy, L. (2013). Impairments in real-world executive function increase from childhood to adolescence in autism spectrum disorders. *Neuropsychology*, *27*, 13-18. doi: 10.1037/a0031299.

Sabol, T. J., & Pianta, R. C. (2012). Patterns of school readiness forecast achievement and socioemotional development at the end of elementary school. *Child Development*, *83*, 282-299. doi: 10.1111/j.1467-8624.2011.01678.x.

Salmon, K., O'Kearney, R., Reese, E., & Fortune, C. A. (2016). The role of language skill in child psychopathology: implications for intervention in the early years. *Clinical Child and Family Psychology Review*, *19*, 352-367. doi: 10.1007/s10567-016-0214-1.

Sasser, T. R., Bierman, K. L., & Heinrichs, B. (2015). Executive functioning and school adjustment: The mediational role of pre-kindergarten learning-related behaviors. *Early Childhood Research Quarterly*, *30*, 70-79. doi: 10.1016/j.ecresq.2014.09.001.

Sektnan, M., McClelland, M. M., Acock, A., & Morrison, F. J. (2010). Relations between early family risk, children's behavioral regulation, and academic achievement. *Early Childhood Research Quarterly*, *25*, 464-479. doi: 10.1016/j.ecresq.2010.02.005.

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*, 906-916. doi: 10.1037/0012-1649.33.6.906.

Shonkoff, J. P. (2011). Protecting brains, not simply stimulating minds. *Science*, *333*, 982-983. doi: 10.1126/science.1206014.

Skibbe, L. E., Montroy, J. J., Bowles, R. P., & Morrison, F. J. (2018). Self-regulation and the development of literacy and language achievement from preschool through second grade. *Early Childhood Research Quarterly, 46,* 240-251. doi: 10.1016/j.ecresq.2018.02.005.

Smith-Donald, R., Raver, C. C., Hayes, T., & Richardson, B. (2007). Preliminary construct and concurrent validity of the Preschool Self-regulation Assessment (PSRA) for field-based research. *Early Childhood Research Quarterly*, *22*, 173-187. doi: 10.1016/j.ecresq.2007.01.002.

Steinberg, E. A., & Drabick, D. A. (2015). A developmental psychopathology perspective on ADHD and comorbid conditions: the role of emotion regulation. *Child Psychiatry & Human Development*, *46*, 951-966. doi: 10.1007/s10578-015-0534-2.

Ursache, A., Blair, C., & Raver, C. C. (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. doi: 10.1111/j.1750-8606.2011.00209.x.

U.S. Health and Human Services Department. (2016). Annual update of the HHS poverty guidelines. Retrieved from https://www.federalregister.gov/documents/2016/01/25/2016-01450/annual-update-of-the-hhs-poverty-guidelines.

Vallotton, C., & Ayoub, C. (2011). Use your words: The role of language in the development of toddlers' self-regulation. *Early Childhood Research Quarterly*, *26*, 169-181. doi: 10.1016/j.ecresq.2010.09.002.

Vitiello, V. E., Booren, L. M., Downer, J. T., & Williford, A. P. (2012). Variation in children's classroom engagement throughout a day in preschool: Relations to classroom and child factors. *Early Childhood Research Quarterly*, *27*, 210-220. doi:10.1016/j.ecresq.2011.08.005.

Vygotsky, L. S. (1962). *Language and thought.* Ontario, Canada: Massachusetts Institute of Technology Press.

Weiland, C., Barata, M. C., & Yoshikawa, H. (2014). The co-occurring development of executive function skills and receptive vocabulary in preschool-aged children: A look at the direction of the developmental pathways. *Infant and Child Development*, *23*, 4-21. doi: 10.1002/icd.1829.

Wiebe, S. A., Espy, K. A., & Charak, D. (2008). Using confirmatory factor analysis to understand executive control in preschool children: I. Latent structure. *Developmental Psychology*, *44*, 575-587. doi: 10.1037/0012-1649.44.2.575.

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

Williford, A. P., Maier, M. F., Downer, J. T., Pianta, R. C., & Howes, C. (2013). Understanding how children's engagement and teachers' interactions combine to predict school readiness. *Journal of Applied Developmental Psychology*, *34*, 299-309. doi: 10.1016/j.appdev.2013.05.002.

Williford, A. P., Whittaker, J. E. V., Vitiello, V. E., & Downer, J. T. (2013). Children's engagement within the preschool classroom and their development of self regulation. *Early Education & Development*, *24*, 162-187. doi: 10.1080/10409289.2011.628270.

Willoughby, M. T., Blair, C. B., Wirth, R. J., & Greenberg, M. (2012). The measurement of executive function at age 5: Psychometric properties and relationship to academic achievement. *Psychological Assessment*, *24*, 226-239. doi: 10.1037/a0025361.

Willoughby, M. T., Pek, J., Blair, C. B., & Family Life Project Investigators. (2013). Measuring executive function in early childhood: A focus on maximal reliability and the derivation of short forms. *Psychological Assessment, 25*, 664-670. doi: 10.1037/a0031747.

Winsler, A., De Leon, J. R., Wallace, B. A., Carlton, M. P., & Willson-Quayle, A. (2003). Private speech in preschool children: Developmental stability and change, across-task consistency, and relations with classroom behaviour. *Journal of Child Language*, *30*, 583-608. doi: 10.1017/S0305000903005671.

Winsler, A., Manfra, L., & Diaz, R. M. (2007). “Should I let them talk?”: Private speech and task performance among preschool children with and without behavior problems. *Early Childhood Research Quarterly*, *22*, 215-231. doi: 10.1016/j.ecresq.2007.01.001.

Zelazo, P. D., Blair, C. B., & Willoughby, M. T. (2016). *Executive function: Implications for education (NCER 2017-2000)*. Washington, DC: National Center for Education Research, Institute of Education Sciences, U.S. Department of Education.

Table 1
Descriptive Statistics for Study Variables

	% / Mean (SD)	Range
Child characteristics		
Age at pre-K entry (months)	52.50 (3.72)	40 - 67
Male	49.74	0 - 100
African American, non-Hispanic	53.93	0 - 100
Caucasian, non-Hispanic	24.39	0 - 100
Hispanic/Latino	9.76	0 - 100
Another race or multiple race/ethnicities	11.92	0 - 100
Primary language is non-English	2.70	0 - 100
Family characteristics		
Years of maternal education	13.40 (1.74)	8 - 20
Family income-to-needs ratio	1.46 (1.08)	0.07 - 4.99
Teacher characteristics		
Master's or higher	55.79	0 - 100
Years of experience	16.56 (8.28)	1 - 32
Child skills at pre-K entry (fall)		
Language		
PPVT-4 (receptive)	93.00 (15.27)	48 - 133
WJ-III Picture Vocabulary (expressive)	99.59 (10.02)	33 - 139
Executive Functioning		
EF Touch Pig	.87 (.14)	.05 - 1
EF Touch Pick the Picture	0.69 (.11)	.13 - .97
EF Touch Something's the Same	0.66 (.13)	.30 - .96
Head-Toes-Knees-Shoulders	8.91 (13.06)	0 - 56
Pencil Tap	0.52 (.35)	0 - 1
Regulation		
Observed behavior control	6.07 (.95)	1.67 - 7
Teacher-reported behavior control	30.18 (7.15)	8 - 40
Teacher-reported emotion regulation	51.25 (8.94)	10 - 68
Child skills at end of pre-K (spring)		
Regulation		
Observed behavior control	6.16 (.76)	2.5 - 7
Teacher-reported behavior control	30.78 (7.35)	8 - 40
Teacher-reported emotion regulation	52.68 (9.82)	10 - 68
N	380	–

Note. WJ-III = Woodcock Johnson Tests of Achievement, Third Edition. PPVT-4 = Peabody Picture Vocabulary Test, Fourth Edition. Observed behavior control measured by the Behavior Control dimension of the inCLASS. Teacher-reported behavior control assessed using the Behavior Control subscale of the Teacher Child Rating Scale. Teacher-reported emotion regulation measured by the Emotion Regulation subscale of the Emotion Regulation Checklist.

Table 2

Bivariate Correlations Among Study Variables

		1	2	3	4	5	6	7	8	9	10	11	12	13	14
	Child skills at pre-K entry (fall)														
1	PPVT-4 (receptive)	-													
2	WJ-III Picture Vocabulary (expressive)	0.50***	-												
3	Language composite	0.86***	0.87***	-											
4	EF Touch Pig	0.25***	0.17**	0.23***	-										
5	EF Touch Pick the Picture	0.26***	0.04	0.17**	0.16**	-									
6	EF Touch Something's the Same	0.36***	0.25***	0.35***	0.13*	0.24***	-								
7	Head-Toes-Knees-Shoulders	0.37***	0.11*	0.27***	0.25***	0.26***	0.29***	-							
8	Pencil Tap	0.35***	0.24***	0.34***	0.35***	0.27***	0.22***	0.44***	-						
9	EF composite	0.44***	0.26***	0.40***	0.74***	0.67***	0.69***	0.38***	0.42***	-					
10	Observed behavior control	0.11*	0.11*	0.13***	0.16**	0.07	0.08	0.08	0.11*	0.13***	-				
11	Teacher-reported behavior control	0.16***	0.02	0.10	0.17**	0.16**	0.13*	0.19***	0.21***	0.22***	0.42***	-			
12	Teacher-reported emotion regulation	0.20***	0.11*	0.17***	0.15**	0.11*	0.12*	0.19***	0.19***	0.18***	0.17***	0.50***	-		
	Child skills at end of pre-K (spring)														
13	Observed behavior control	0.13*	0.07	0.11*	0.09	0.06	0.07	0.06	0.13*	0.13*	0.40***	0.32***	0.06	-	
14	Teacher-reported behavior control	0.13*	0.01	0.08	0.15*	0.09	0.10	0.17***	0.19***	0.16**	0.31***	0.77***	0.44***	0.30***	-
15	Teacher-reported emotion regulation	0.18**	0.12*	0.17**	0.06	0.01	0.07	0.15**	0.18***	0.09	0.17*	0.44***	0.68***	0.10	0.53***

Note. * $p < .05$, ** $p < .01$, *** $p < .001$. PPVT-4 = Peabody Picture Vocabulary Test - Fourth Edition. WJ-III = Woodcock Johnson Tests of Achievement - Third Edition. Observed behavior control measured by the Behavior Control dimension of the inCLASS. Teacher-reported behavior control assessed using the Behavior Control subscale of the Teacher Child Rating Scale. Teacher-reported emotion regulation measured by the Emotion Regulation subscale of the Emotion Regulation Checklist. Language composite is the mean of PPVT-4 and WJ-III Picture Vocabulary. EF composite is the mean of the three EF Touch subtests, HTKS, and Pencil Tap.

Table 3

Independent and Interactive Contributions of Fall Executive Functioning and Language Skills to Spring Regulation Abilities

	Observed Behavior Control (inCLASS)			Teacher-reported Behavior Control (TCRS)			Teacher-reported Emotion Regulation (ERC)		
	B	SE		B	SE		B	SE	
Base Models									
Fall executive functioning	-0.01	0.05		-0.07	0.37		-0.29	0.53	
Fall language	0.01	0.05		0.02	0.32		1.10	0.46	*
Fall control	0.24	0.04	***	0.81	0.03	***	0.63	0.05	***
Child age	-0.01	0.01		-0.05	0.07		-0.03	0.1	
Child is male	-0.42	0.07	***	-0.36	0.49		-0.02	0.68	
Child is Caucasian	0.06	0.09		0.32	0.65		0.36	0.95	
Child is Hispanic/Latino	0.19	0.12		0.64	0.86		2.78	1.24	*
Child is another race/ethnicity	0.29	0.12	*	1.39	0.82		2.92	1.21	*
Child's primary language is non-English	-0.13	0.22		2.89	1.57		-0.60	2.24	
Years of maternal education	-0.01	0.02		-0.16	0.16		-0.23	0.23	
Family INR	0.04	0.04		0.67	0.26	*	0.52	0.38	
Teacher years of experience	0.02	0.01	**	0.02	0.03		-0.04	0.08	
Teacher has a master's degree or higher	-0.21	0.09	*	-1.05	0.59		-2.03	1.32	
Final Models									
EF x language	-0.02	0.04		-0.25	0.29		-1.02	0.40	*
N					380				

Note. * $p < .05$, ** $p < .01$, *** $p < .001$. inCLASS = Individualized Classroom Assessment Scoring System. TCRS = Teacher Child Rating Scale. ERC = Emotion Regulation Checklist. Fall executive functioning is a composite of the average of the following standardized variables: three EF Touch subtests, Head-Toes-Knees-Shoulders, and coefficients. The base models include covariates and main effects of EF and language on each of the three outcomes separately. The final models include all variables in the base model plus an EF*language interaction term. Coefficients for main effects and covariates were similar in the base and final models.

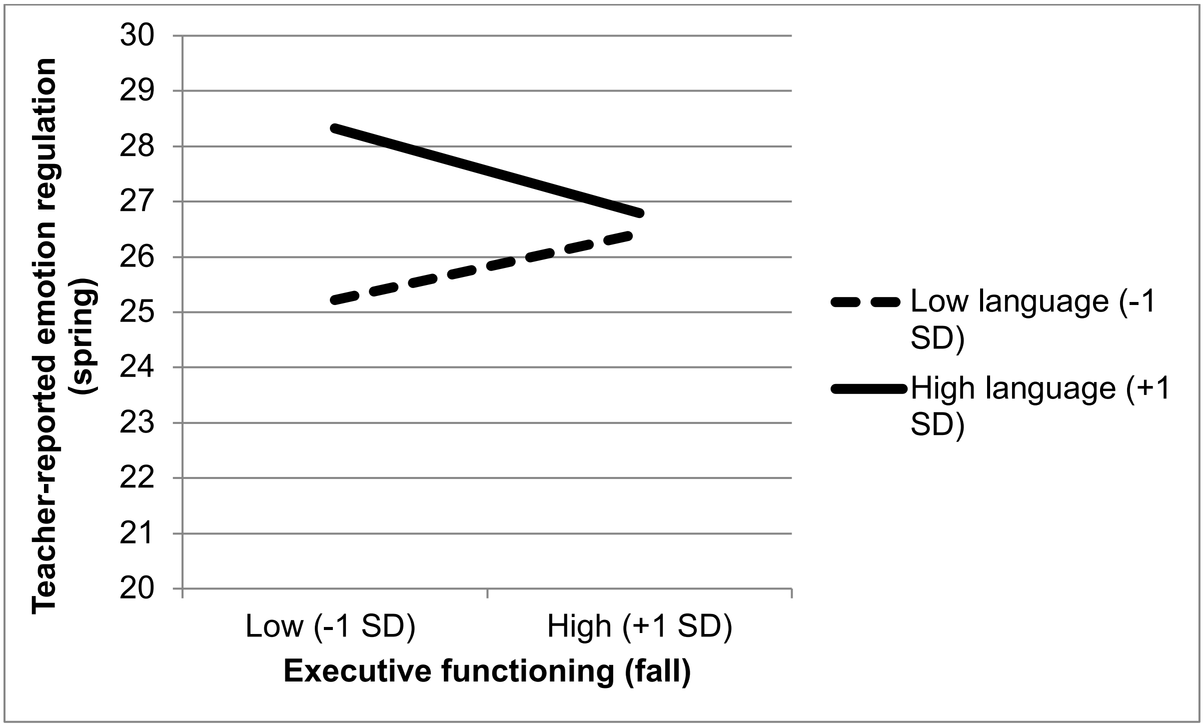

Figure 1. Graphical representation of the interaction effect between fall executive functioning and fall language on gains in teacher-reported emotion regulation across the preschool year.

MANUSCRIPT THREE

Executive Functioning from Preschool to Kindergarten: Developmental Trajectories and Demographic Differences

Shannon E. Reilly, M.Ed., and Jason T. Downer, Ph.D.

Center for Advanced Study of Teaching and Learning, University of Virginia

Author Note

Shannon Reilly, Center for Advanced Study of Teaching and Learning, University of Virginia; Jason T. Downer, Center for Advanced Study of Teaching and Learning, University of Virginia.

This work was supported by the National Institute of Child Health and Human Development of the National Institutes of Health (grant number 2R01HD051498-06A1). The opinions expressed are those of the authors and do not represent views of the funding agency. Special thanks to the participating teachers, children, and families.

Correspondence concerning this article should be addressed to Shannon Reilly, Center for Advanced Study of Teaching and Learning, University of Virginia, PO Box 400267, Charlottesville, VA 22904. Email: sr6hd@virginia.edu

Acknowledgement

The authors thank Kevin Grimm for his analytical assistance.

Abstract

Executive functioning (EF) is a key predictor of long-term success that develops rapidly in early childhood. However, its developmental trajectory across the key transition from preschool to kindergarten is not yet fully understood. Whether and how this trajectory differs based on characteristics of children and their families also remains to be characterized. In a primarily low-income, racially and ethnically diverse, urban sample, the present study found that EF development in early childhood was best characterized as nonlinear, with the majority of growth occurring during the preschool year. Further, there was intra-individual variability in EF ability at preschool entry that was in part explained by demographic differences. Specifically, boys on average began preschool with lower EF than girls, and there was a trending indication that children in poverty started with lower EF than more affluent peers. Findings were mixed as to whether children in poverty also experienced more rapid EF development than non-impoverished peers. Children's age relative to their classmates was not associated with initial EF or growth over time. Findings have implications for (1) examining EF development in early childhood with more specificity in future studies, (2) informing the timing of EF interventions in early childhood, and (3) identifying children for whom such interventions might be especially beneficial.

Key words: preschool; executive functioning; school readiness; developmental trajectories; individual differences

Executive Functioning from Preschool to Kindergarten: Developmental Trajectories and Demographic Differences

Executive functioning (EF), a multidimensional cognitive skillset that guides goal-directed behavior (Baggetta & Alexander, 2016; Blair, 2016), develops rapidly in early childhood (Garon, Bryson, & Smith, 2008) and is a key facilitator of long-term academic achievement, positive social functioning, and school success (Baggetta & Alexander, 2016; Blair, 2016; Zelazo, Blair, & Willoughby, 2016). Because foundational skills that emerge early on set the stage for cultivating more complex abilities (Masten & Cicchetti, 2010), it is essential that young children develop EF skills prior to and across the transition to kindergarten. This transition is characterized by increases in demands and expectations (Rimm-Kaufman & Pianta, 2000), which tax children's capacity to leverage underlying EF skills (Blair, Zelazo, & Greenberg, 2005; Jones, Bailey, Barnes, & Partee, 2016) to regulate themselves in a classroom setting (e.g., follow directions, sit still and pay attention, finish tasks; Bassok, Latham, & Rorem, 2016).

In the past couple of decades, there have been a plethora of studies examining EF in early childhood (e.g., Anderson & Reidy, 2012; Garon, Bryson, & Smith, 2012; Willoughby, Blair, Wirth, Greenberg, and the Family Life Project Investigators, 2012, which have enhanced understanding of what EF looks like (Friedman & Miyake, 2017), how it develops over time (Best & Miller, 2010; Hughes et al., 2009; Lee, Bull, & Ho, 2013; Miyake et al., 2000), and how it relates to other skills (Blair et al., 2005). However, many studies are cross-sectional and thus do not allow for an examination of changes in EF over time (Garon et al., 2008). More recently, longitudinal studies have begun to fill this gap (Hughes & Ensor, 2011; Hughes, Ensor, Wilson, & Graham, 2009; Willoughby, Wirth, & Blair, 2012) but remain constrained by a limited number of time points that allow for only a rough understanding of EF's trajectory during the transition to kindergarten.

Therefore, this study employs latent growth curve modeling to examine nonlinear growth across five time points during the two years of this transition period. Additionally, given the importance of EF development in early childhood for later success, along with work indicating that there are individual differences in EF development (Zelazo et al., 2016), it is important to understand factors that facilitate or hinder EF at the start of school. The present study examines whether characteristics of children and families are associated with divergent trajectories of EF development from preschool to kindergarten. Results from this study may help to inform both general and targeted interventions for EF in early childhood by identifying children who are at risk of low EF development relative to peers.

Development of Executive Functioning

While there is no standardized definition for executive functioning (EF; Baggetta & Alexander, 2016), it has generally been described as a "set of cognitive processes" (Baggetta & Alexander, 2016) that "coordinat[es] multiple sources of information in the service of purposeful, goal-directed behavior" (Blair, 2016); this skillset is engaged in "formulating goals, planning how to achieve them, and carrying out these plans effectively" (Anderson & Reidy, 2012). EF is a multidimensional construct comprised of three primary components: working memory, inhibitory control, and set shifting or cognitive flexibility (Baggetta & Alexander, 2016; Blair, Zelazo, & Greenberg, 2005; Miyake et al., 2000; Zelazo et al., 2016). *Working memory* is conceptualized as keeping information in mind and manipulating or updating it (Diamond, 2013; Garon, Bryson, & Smith, 2008; Willoughby, Wirth, & Blair, 2012; Zelazo et al., 2016); an example of this is keeping in mind and executing instructions the teacher just provided (e.g., "put on your coat and line up at the door"). *Inhibitory control* is the ability to "control one's attention, behavior, thoughts, and/or emotions to override a strong internal predisposition … and instead do what's more appropriate or needed" (Diamond, 2013); examples include "ignoring a distraction, stopping an impulsive utterance, or overcoming a highly learned response" (Zelazo et al., 2016). *Set shifting* can be understood as "changing how we think about something" (Diamond et al., 2013), such as "considering someone else's perspective on a situation or solving a mathematics problem in multiple ways" (Zelazo et al., 2016).

There is widespread evidence that EF underpins academic achievement, social competence, and school success in the short- and long-term (Baggetta & Alexander, 2016; Blair, 2016). EF is promotive of these positive outcomes because it enables children to "sit still, pay attention, remember and follow rules, and flexibly adopt new perspectives" (Zelazo et al., 2016); this in turn allows them to benefit more from learning opportunities (Morgan, Farkas, Hillemeier, Pun, & Maczuga, 2018). Specifically, having well-developed EF in preschool has been linked to growth in language, literacy, and math by the end of kindergarten (Fuhs, Nesbitt, Farran, & Dong, 2014; McClelland et al., 2014). Over time, EF skills in early childhood have been linked to academic achievement in second grade (Morgan et al., 2018), fifth grade (Sabol & Pianta, 2012), and young adulthood, as well as educational attainment by age 25 (McClelland, Acock, Piccinin, Rhea, & Stallings, 2013). Conversely, having meager EF abilities early on has been linked to a host of maladaptive outcomes, such as learning and behavioral difficulties (Zelazo et al., 2016). In fact, a longitudinal study of self-control – which can be considered the application of EF in context (Jones et al., 2016; Ponitz, McClelland, Matthews, & Morrison, 2009) – found that having low self-control in the first decade of life was associated with health problems and financial difficulties, as well as elevated risk for depression, substance abuse, and criminal conviction at age 32 (Moffitt et al., 2011). Given its role in facilitating both skill growth across domains and significant long-term outcomes, it is paramount to understand the development of EF during times of substantial growth.

EF progresses rapidly during early childhood (Blair, 2016; Garon et al., 2008) and continues to grow through childhood and adolescence (Best & Miller, 2010) along with the maturation of the prefrontal cortex (Shonkoff, 2011). Like other emerging skills, expectations for EF are developmentally graded. For example, one might expect a preschooler to be able to keep in mind and execute the rule to use "walking feet," a 10-year-old to be able to pay attention in class despite distractions, and an adolescent to appropriately plan and execute writing a term paper instead of going out with friends. As such, the construct of EF becomes more refined and complex over the first two decades of development. In fact, a large body of evidence suggests that EF hangs together as a single construct in early childhood (Hughes et al., 2009; Wiebe, Espy, & Charak, 2008; Wiebe et al., 2011; Willoughby, Blair, Wirth, & Greenberg, 2012) and even into middle childhood (Brydges, Reid, Fox, & Anderson, 2012), then becomes differentiated into the three component skills in adolescence (Baggetta & Alexander, 2016; Best & Miller, 2010; Blair & Ursache, 2011; Miyake, 2000). However, other studies have found evidence for a two-factor model disaggregating inhibitory control and working memory in early childhood (ages 3 to 6; Miller, Giesbrecht, Muller, McInerney, & Kerns, 2012; Usai, Viterbori, Traverso, & DeFranchis, 2014), and three separate but correlated factors in middle childhood (i.e., ages 8 to 13; Lehto, Juujarvi, Kooistra, & Pulkkinen, 2003). Thus, like the conceptualization and operationalization of EF (Baggetta & Alexander, 2016), the structure of EF at various ages has not yet been uniformly determined and is an area of ongoing research (Zelazo et al., 2016). However, it does appear that there are multiple, related components of EF that tend to be united early on and become increasingly refined and distinguished from one another over time.

The present study focuses on EF development in early childhood because cultivating foundational abilities early on begins a positive cascading effect on more complex, higher-level skills (Masten & Cicchetti, 2010). It is particularly important to understand EF development not just in early childhood generally but specifically during preschool and kindergarten because these years represent an important transition during which routines shift and demands and expectations increase (Rimm-Kaufman & Pianta, 2000), particularly around EF and related self-regulatory capacity (Bassok, Latham, & Rorem, 2016; Jones et al., 2016). Therefore, understanding the trajectory of EF development during this time period, as well as factors that facilitate this growth, can promote successful adjustment to elementary school.

Modeling EF during a Key Developmental Transition

EF research has been plagued by a “measurement impurity problem,” meaning that many assessments in this domain tap other abilities such as motor skills, processing speed, and language in addition to EF; they also tend to simultaneously measure multiple components of EF (Anderson & Reidy, 2012; Baggetta & Alexander, 2016; Fuhs et al., 2014; Miyake et al., 2000; Zelazo et al., 2016). This methodological roadblock makes it difficult to substantiate specific claims about EF, its development, and its relation to other skills. To address this measurement impurity issue and reduce measurement error (Willoughby, Blair et al., 2012), many studies use a latent variable approach to isolate an underlying "true" EF skill that is common across multiple tasks (Miyake & Friedman, 2012; Zelazo et al., 2016). Often, the fit of the latent EF factor is actually better than the associations among separate EF tasks, which tend to have relatively small correlations (e.g., primarily between .10 and .20; Morgan et al., 2018; Wiebe et al., 2011). This indicates that various EF tasks tap different aspects of a multidimensional EF construct, and that the overarching construct of EF is more unified than the sum of its components.

In early childhood, EF has most consistently been found to be a single latent factor (Wiebe et al., 2008, 2011; Willoughby, Blair et al., 2012; Zelazo et al., 2016), which aligns with the conceptual idea that the development of EF is relatively undifferentiated early on and becomes more distinguished with age (Lehto, Juujärvi, Kooistra, & Pulkkinen, 2003; Miyake et al., 2000). Conversely, a few studies in early childhood (Lee, Ho, & Bull, 2013; Miller et al., 2012; Usai et al., 2014) have found evidence for two EF factors that disaggregate inhibitory control from working memory and set shifting. Understanding the structure of EF in early childhood is important because it can clarify conceptual and methodological ambiguity (Jones et al., 2016), guide avenues for future research, and inform interventions about whether to focus on specific aspects of EF or the overarching construct more generally (Zelazo et al., 2016). Taking a developmental perspective and following the majority of the early childhood EF literature, the present study aims to confirm whether EF is best characterized as a unitary latent construct from preschool through kindergarten, given that significant differentiation of components has been found to occur later in development (e.g., Lehto et al., 2003; Miyake et al., 2000). In creating a latent factor of EF, it is important to determine whether this construct is psychometrically equivalent over time (i.e., that it has measurement invariance; Willoughby, Wirth et al., 2012). This has been established across several studies with different samples and different combinations of EF tasks (Hughes et al., 2009; Willoughby, Wirth et al., 2012). It will be important to replicate that measurement invariance holds in the present sample in order to compare results to findings in prior research.

Much of the early work examining the structure of EF development, particularly in early childhood, has been cross-sectional in nature (Garon et al., 2008). It is important to examine skill growth longitudinally because skills beget skills (Masten & Cicchetti, 2010), and EF is developing rapidly in early childhood (Blair, 2016; Garon et al., 2008) during the key transition from preschool to kindergarten (Rimm-Kaufman & Pianta, 2000). Although there is a dearth of prospective longitudinal studies examining EF development over time (Zelazo et al., 2016), several exceptions exist. Some earlier studies found linear growth in EF within the preschool year in a racially and ethnically diverse sample (Fuhs & Day, 2011) and between ages 4 and 6 in a low-income, primarily White sample, with a skill increase of two standard deviations over those two years (Hughes et al., 2009; Hughes & Ensor, 2011). Moreover, individual differences in linear EF growth from 4 to 6 predicted variation in academic and behavioral outcomes (Hughes & Ensor, 2011). Specifically, slow growth in EF was related to higher teacher-reported externalizing and internalizing behaviors in first grade. Children with fast EF growth from 4 to 6 reported higher academic competence at age 6 than peers with a shallower growth trajectory. Results from this study provide further evidence that these differing EF trajectories have long-term implications for children's success in school.

Following a recognized need to examine the shape of EF growth in a more nuanced fashion (Hughes et al., 2009; Hughes & Ensor, 2011), recent work has used additional time points to analyze nonlinear development in early childhood (Clark et al., 2013; Montroy et al., 2016; Wiebe, Sheffield, & Espy, 2012; Willoughby, Wirth et al., 2012). Altogether, these studies have found evidence for nonlinear development of EF from ages 3 to 7 in demographically varied samples, with accelerated development early in the trajectory that then slows over time.

Specifically, in a racially and ethnically diverse, low-income, rural sample, Willoughby, Wirth, and colleagues (2012) took an exploratory approach by comparing linear and freeloading latent growth curve models. Although the two models had comparable fit, the nonlinear (i.e., freeloading) model was statistically superior, and 60% of EF growth occurred during ages 3 to 4, with slower improvement from 4 to 5. Findings were similar in two studies of a predominantly White, relatively socioeconomically at-risk sample of children ages 3 to 5 that focused specifically on inhibitory control (Wiebe, Sheffield, & Espy, 2012) and cognitive flexibility (Clark et al., 2013). These studies found that EF growth was nonlinear – specifically, quadratic – with more EF growth from ages 3 to 4 than ages 4 to 5. In predominantly White, mixed-income samples examining the development of EF (assessed by a single measure: HTKS) from age 3 to 7, growth was found to be exponential, with faster growth in preschool and decelerated growth in early elementary school (Montroy, Bowles, Skibbe, McClelland, & Morrison, 2016). This most recent study identified that more work needs to be done to confirm the functional form of EF in the transition from preschool to elementary school with measures that capture the multidimensional nature of EF more comprehensively (Montroy et al., 2016). The present study extends this recent work by examining the nonlinear development of EF across five time points in the transition from preschool to kindergarten using multiple directly assessed measures that consider the multidimensional nature of the construct.

Demographic Differences in Early Childhood EF Development

The trajectory of EF development in early childhood is unlikely to be the same for all children, because skill development is in part dependent on characteristics and resources of children and their families (Bronfenbrenner & Morris, 2006; Rimm-Kaufman & Pianta, 2000; Yeung, Linver, & Brooks-Gunn, 2002; Zelazo et al, 2016). However, it is unknown to what extent the nuanced trajectory of EF skill development from preschool to kindergarten differs based on child and family characteristics. The present study specifically examines two child demographic factors (i.e., sex and age relative to their classmates), as well as a family characteristic (poverty status) that have been consistently associated with EF development. Understanding how EF skills at preschool entry and the rate of change in EF over time differ based on these characteristics can help to identify children at risk for relatively low EF at preschool entry and/or slow growth over time who might benefit from additional support for EF skills during this key transition period.

The majority of research finds that girls have an advantage over boys in EF in early childhood (see Hongwanishkul, Happaney, Lee, & Zelazo, 2005 for exception). Specifically, girls were found to have higher directly assessed EF skills in preschool (Clark et al., 2013; Wiebe et al., 2008, 2012), at kindergarten entry (Conway et al., 2018), and in an older sample (i.e., age 8 to 30; Kalkut, Han, Lansing, Holdnack, & Delis, 2009). In parent and teacher reports of EF in 2- to 5-year-olds, boys were rated as having more EF difficulties than girls, including inhibitory control, working memory, and planning (Isquith et al., 2004). Based on this literature, it is expected that girls will show an advantage in EF at preschool entry. In studies assessing EF growth from age 3 to age 7 in predominantly White, mixed-income samples, girls were more likely to have an "early" EF trajectory in preschool compared to boys' more protracted trajectory (Montroy et al., 2016). This provides some evidence that girls may be advantaged in rate of EF growth as well, although there is less of a consensus about this pattern. Sex differences in EF initial ability and growth over time may also be clinically relevant because they could foreshadow sex differences in rates of attention-deficit/hyperactivity disorder symptoms (ADHD; Bauermeister et al., 2007), particularly given that low EF has been related to ADHD symptoms in preschool (Thorell & Wahlstedt, 2006).

Several studies have found small but consistent positive associations between socioeconomic status (SES) and EF from early childhood through adolescence (Lawson & Farah, 2017; Lawson, Hook, & Farah, 2018). Studies using a nationally representative dataset have found SES disparities in EF skills at kindergarten entry (Conway, Waldfogel, & Wang, 2018; Little, 2017), with especially pronounced disparities for children from families with incomes below the poverty line (Conway et al., 2018). Even when studies use broader definitions of SES that include material hardship, findings have been similar. For instance, chronic exposure to poverty and financial strain from 15 months to 48 months was associated with lower EF skills at 48 months (Raver, Blair, & Willoughby, and the Family Life Project Investigators, 2013). An index of socioeconomic challenges (i.e., family income, parent education, subjective social status, and financial stress) was linearly and negatively associated with kindergarteners' directly assessed and assessor-rated EF (Finch & Obradovic, 2017). Previous work on how SES is associated with the rate of EF growth has been equivocal. When both level and rate of change in EF are examined in early childhood, family income has been associated with EF skill level both concurrently and longitudinally but has not been related to the rate of growth in EF (Hackman, Gallop, Evans, & Farah, 2015; Hughes et al., 2009). Other work has shown that socioeconomic gaps in EF tend to narrow as children age, more so than for academic skills (Little, 2017). Taken together, previous research suggests that children in poverty will begin preschool with lower EF than non-impoverished peers, but it is unclear whether poverty status will be unrelated to growth in EF over time (Hackman et al., 2009; Hughes et al., 2009), or whether children in poverty will have a faster rate of growth than more affluent peers (Little, 2017).

Finally, older children have been shown to have higher EF than younger peers (Conway et al., 2018; Hongwanishkul et al., 2005; Obradovic, Portilla, & Ballard, 2016), even when controlling for verbal ability (Carlson, 2005). What is not yet known is – over and above children's ages – how their age relative to their classmates' ages (i.e., to what extent their age deviates from the mean classroom age) is associated with their EF ability at preschool entry, as well as the rate at which it develops over time. While little work has been done examining classroom composition and executive functioning in early childhood, Ansari (2017) found that in mixed-age kindergarten classrooms, 5-year-olds who were in classrooms with preschoolers made smaller gains in EF than children in classrooms with same-age peers. In a Head Start sample, 4-year-olds demonstrated fewer academic gains during preschool when they were in mixed-age classrooms with more 3-year-olds, whereas classroom age composition was not associated with academic gains for 3-year-olds; notably, this study did not look at EF specifically (Ansari, Purtell, & Gershoff, 2016). Together, these findings may suggest that relatively older children may have lower growth in EF than peers who are relatively young for their class. However, this hypothesis is merely exploratory at this time. Despite the fact that older children on average have higher EF (e.g., Conway et al., 2018; Obradovic et al., 2016), it is unclear whether children who are *relatively* older than classmates will have higher initial EF ability, over and above their absolute age.

The Present Study

The present study aimed to replicate and extend prior research on EF in early childhood by examining whether EF is best conceptualized as a unitary latent construct, modeling the trajectory of EF, and investigating to what extent the trajectory of skill development differs based on child and family demographic characteristics. Its main contribution to the literature is that it investigates nonlinear growth in EF across five time points during a key transition point in early childhood within a low-income, racially and ethnically diverse, urban sample; past work has analyzed linear growth, examined fewer time points, and/or used a low-income rural or primarily White sample (e.g., Hughes et al., 2009; Wiebe et al., 2011; Willoughby, Wirth et al., 2012). Thus, findings from this study will provide more nuanced information for researchers about EF development from preschool to kindergarten. It may also inform practice by illuminating time periods during which EF is developing more rapidly and thus might be most malleable to intervention, as well as by identifying children who might be most at risk for low EF and/or slow EF development and thus who might benefit most from targeted intervention to support EF skills.

Based on prior research, it was hypothesized that a unitary construct of EF would best fit the data and that this construct would be psychometrically equivalent across the five time points (i.e., that measurement invariance would be established). Growth in EF was expected to be nonlinear, with more growth occurring in preschool than in kindergarten (Montroy et al., 2016; Wiebe et al., 2012; Willoughby, Wirth et al., 2012); however, it was hypothesized that a linear model would also adequately fit the data (Hughes et al., 2009; Hughes & Ensor, 2011; Wiebe et al., 2012; Willoughby, Wirth et al., 2012). We anticipated demographic differences, such that boys and children in poverty would have relatively lower EF than girls and more affluent peers at preschool entry, and that girls would have more rapid EF development over time (Montroy et al., 2016). Children in poverty may have a faster rate of growth than more affluent peers (Little, 2017), but based on prior work it was unclear *a priori* whether poverty status would be associated with EF growth. Finally, we hypothesized that being older relative to classmates on average would be associated with a slower rate of EF growth, following limited early childhood work on classroom composition (Ansari, 2017; Ansari et al., 2016).

Method

Sample

Data for the present study were collected as part of the Understanding the Power of Preschool for Kindergarten Success (P2K) project, an observational study of children's experiences from preschool through kindergarten. Data were collected from fall 2016 to spring 2018 and comprised the full sample of the first cohort of the P2K study. The study was conducted in 15 public preschools and Head Start centers across 2 sites in the southeastern United States; in the second year, children matriculated to 122 kindergarten classrooms in 35 schools. Preschool teachers were eligible for participation if they served children who would matriculate into kindergarten during the 2017-18 school year; in inclusion classrooms, the general education teacher was selected for participation. Fifty preschool teachers from those who consented were randomly selected to participate in the first year of the study; one child transitioned to a different study classroom during the year, so a total of 51 preschool teachers participated. Children in participating classrooms were eligible for the study. Up to 8 consented children per classroom were randomly selected to participate, blocked by gender. When fewer than 8 children were consented, all consented children in the classroom were chosen. A total of 380 children participated in the study and were followed into their kindergarten classrooms when possible. Consent was obtained from kindergarten teachers of study children. In spring of the kindergarten year, there was planned missingness in 50% of direct assessment data. Children were randomly selected to participate in the EF direct assessment prior to this data collection window.

Preschool teachers were 98.04% female with a mean age of 44.59 years (*SD* = 10.43, range = 23 - 63); 72.55% were Caucasian, 23.53% were African American, and 3.92% were Hispanic or Latino. Teachers had an average of 16.22 years of experience at their current facility (*SD* = 8.46, range: 1 - 32 years); 54.90% had a master's degree or higher. Approximately 94 percent of preschool classrooms were in public schools and 6% were housed in Head Start centers. Class sizes ranged from 16 to 21 children, with a mean of 18 (*SD* = .83). On average there were 7.70 study children in preschool classrooms (*SD* = .92, range = 1-9) and 3.68 in kindergarten classrooms (*SD* = 2.13, range = 1-10). On average, 49.70% of children per class were female, 3% had limited English proficiency (LEP; *SD* = 4.61%) and 6.12% had an IEP or IFSP (*SD* = 11.32%). Approximately 83% of children in a given classroom qualified for free or reduced-price lunch (*SD* = 26.33%).

Across these 50 preschool classrooms, 380 eligible and consented children participated in the present study. Table 1 provides descriptive demographic information for the sample. Approximately 50% of children were male; 53.93% percent were Black/African American, 24.39% were White/European American, 9.76% were Hispanic/Latino, and 11.92% were of other racial or ethnic backgrounds (e.g., Asian, American Indian or Alaska Native) or multiple races. English was not the primary language for 2.70% of children. At the beginning of preschool, children had a mean age of 52.50 months (*SD* = 3.72); they were 72.82 months (*SD* = 3.58) by spring of kindergarten. Descriptive information for children's age relative to their peers at the beginning of preschool (i.e., deviation in months from the mean classroom age) was as follows: *M* = 0, *SD* = 3.90, range = -15.13 to 16.38. Mothers had on average 13.40 years of education (*SD* = 1.74, range: 8 - 20). Families had an average income-to-needs ratio of 1.46 (*SD* = 1.08, range = 0.07 – 4.99), meaning that the average family in the study had an annual income at 146% of the federal poverty level for their household size (e.g., $

35,478 for a family of 4; U.S. Health and Human Services Department, 2016). Almost half of the sample (41.67%) was considered to be in poverty (i.e., income-to-needs ratio less than 1).

Procedure

Recruitment. Preschool program administrators in 2 urban areas in the southeastern United States were contacted to participate in the study, and administrators and teachers were invited to attend study recruitment sessions. Interested and eligible teachers gave informed consent, completed personal and classroom demographic surveys, and allowed data collectors to observe their classrooms. Parents or guardians of children in participating classrooms were given a letter explaining the study, an informed consent, and a demographic survey. Of consented children, up to 8 from each classroom were randomly selected to participate in the study (blocked by gender). When children matriculated to kindergarten, their teachers were asked to consent to participate in the study. Kindergarten teachers who provided informed consent completed personal and classroom demographic surveys, and allowed data collectors to observe their classrooms.

Data collection. Parents and teachers completed demographic surveys in the fall of preschool and kindergarten; teachers also completed surveys and child ratings in the spring of the relevant academic year. Trained data collectors administered direct child assessments in the fall, winter, and spring during preschool and in fall and spring during kindergarten.

Measures

Executive functioning. This construct was assessed using a computer-administered battery and two interactive tasks to measure multiple aspects of EF. *EF*

Touch is a computer-administered battery of executive functioning tasks that has been shown to have good reliability and validity in an early childhood sample (Willoughby, Blair, et al., 2012; Willoughby, Pek, Blair, & FLP Investigators, 2013). Three subtests from the battery were used in preschool. On the *Animal Go/No-Go* (Pig) task, which measures inhibitory control, farm animals flash quickly on the screen and children are asked to touch all of the animals except for the pig. *Pick the Picture* (PtP) is a task of working memory during which children are asked to consistently choose pictures from a set that they have not chosen before, holding in mind those they have already picked. On *Something's the Same* (StS), which assesses set shifting, the screen displays two pictures that are similar on one dimension (e.g., color, size), and then adds another picture that is the same as one of the first pictures in a different way. Children are asked to choose which of the first two pictures is similar to the new picture (e.g., shift to think about similarity on a different dimension).

In kindergarten, a fourth EF Touch subtest measuring inhibitory control, *Spatial Conflict* (Arrow*s*), was added because of ceiling effects on the preschool inhibitory control (Pig) task. In previous studies (Willoughby, Wirth et al., 2012), Arrows has been added at follow-up time points to reflect the developmentally graded nature of EF. In this task, children were first asked to push the button aligned with the direction in which an arrow pointed. Then, in a second trial, they were asked to push the button in the opposite direction of the arrow. This subtest is comprised of two subscales: Arrows Congruent (trials in which the button and the arrow direction align) and Arrows Switch (trials in which they are discrepant); these were included as separate variables in analyses. All other EF Touch subtests remained the same in the kindergarten year. Proportion correct was calculated for each of the subtests at each time point (three subtests in preschool, five subtests in kindergarten).

Head-Toes-Knees-Shoulders (HTKS) is a measure of executive functioning that requires inhibitory control, working memory, and attention focusing (Ponitz, McClelland, Matthews, & Morrison, 2009). On this task, children are asked to learn simple commands (i.e., “touch your head,” “touch your toes”) then do the opposite of what the assessor said (e.g., touch their toes when asked to touch their head). An advanced trial incorporating the same task with knees and shoulders was given to children who did not reach a ceiling on the first set of items. Children received two points for each correct response and one point for a self-corrected response; scores ranged from 0 to 60 across 30 trials. HTKS has been shown to have adequate reliability and concurrent and predictive validity in a preschool sample (Ponitz et al., 2009).

Inhibitory control was also measured by the Pencil Tap subtest of the *Preschool Self-Regulation Assessment* (PSRA; Smith-Donald, Raver, Hayes, & Richardson, 2007). In this assessment, children were asked to tap their pencil once when the examiner tapped twice and twice when the examiner tapped once. Scores represent percentage of correct responses. This test has shown acceptable concurrent and construct validity (Smith-Donald et al., 2007).

These measures (EF Touch subtests, HTKS, and Pencil Tap) were included in a CFA to determine whether EF was best characterized by a single latent factor in these data (see Analytic Plan for further details). Table 2 displays descriptive information on EF measures across time points.

Child and family characteristics. In fall of preschool, families provided information about child and family characteristics. Children's age in months at preschool entry was used to calculate how their age deviated from the mean age of other children in their classroom. Annual income and household size were used to calculate the income-to-needs ratio, which is a family's income relative to the federal poverty level for their household size. Child sex (male = 1), poverty status at preschool entry (in poverty = 1), and children's deviation from their classroom's mean age (continuous) were included in models as variables of interest. Child age (continuous) and child race/ethnicity (dummy codes for White/European American, Hispanic/Latino, and other race, with Black/African American omitted as the reference category) were also included in models as time-invariant covariates. See the Appendix for a bivariate correlation matrix of all study variables.

Data Analytic Plan

Data Preparation

Clustering. Intraclass correlation coefficients (ICCs) from unconditional regression models and design effects (Mass & Hox, 2004) were examined to determine the need to control for nestedness of children within preschool and/or kindergarten classrooms. The design effect takes into account the ICC and the average cluster (in this case, classroom) size (design effect = 1+[average cluster size-1]*ICC). This second step of assessing nestedness was important for this study because of the small cluster sizes in kindergarten (mean number of study children in a kindergarten classroom = 3.68, *SD* = 2.13). ICCs greater than .10 indicate that a given type of nestedness should be considered in clustering (Raudenbush & Bryk, 2002). Design effects greater than or equal to 2 indicate that multilevel models are necessary; if they are less than 2, it suggests that not using multilevel models is unlikely to significantly bias results (Maas & Hox, 2004).

ICCs ranged from 0 to .17, with 4 out of 29 ICCs that were above the .10 cutoff (1 in preschool, 3 in kindergarten; Raudenbush & Bryk, 2002). However, all design effects (ranging from 1 to 1.81, with average cluster sizes of 7.7 for pre-K and 3.68 for kindergarten) were less than the cutoff of 2 (Maas & Hox, 2004). These findings indicate that not using multilevel modeling is unlikely to significantly distort results. However, to provide a more conservative, robust estimate of results given the dependency of data when children are nested in classrooms, two-level multiple imputation was conducted and all analyses used TYPE=COMPLEX and maximum likelihood estimation with robust standard errors to cluster standard errors and test statistics by preschool classroom.

Missing data. Missing demographic data were as follows: 0% child sex and age, 2.89% child race/ethnicity, and 11.58% income-to-needs ratio. Table 2 displays the percentage of missing data on EF measures at each time point. There were no significant demographic differences between children who did and did not attrite prior to kindergarten spring, or between children who were and were not randomly selected to participate in the kindergarten spring data collection window.

Patterns of missingness for study variables were assessed using Little's test of missing completely at random (MCAR) to determine whether data were MCAR, missing at random (MAR), or not missing at random (NMAR; Li, 2013; Little, 1988). The large percentage of data missing at the kindergarten spring data collection window can be considered MCAR because children were randomly selected to participate at this time point. However, due to attrition, some data may not be MCAR. Little's test of MCAR indicated this was the case, $\chi^2 = 1947.69$ (1419), $p = 0$. As a second step, covariate-dependent missingness was assessed to determine whether key study variables (i.e., components of EF at each time point) were significantly predicted by covariates. This was the case, $\chi^2 = 2742.74$ (7848), $p = 1.0$, which provides evidence that the effect of missingness on results will be minimal (Wiebe et al., 2011) and that data can be considered MAR (Li, 2013). As such, multiple imputation was conducted to retain cases missing data on study variables (von Hippel, 2007), making the full analytic sample 380. Multilevel (i.e., two-level) multiple imputation accounted for clustering of students within preschool classrooms. Ten data sets were imputed using Blimp 1.1 (Keller & Enders, 2017).

Outlier analyses. Because CFA is sensitive to outliers, univariate and multivariate outlier analyses were conducted (Brydges, Reid, Fox, & Anderson, 2012). Univariate outliers were defined as data points greater than 3 standard deviations from the mean (i.e., $z \leq -3.29$ or $z \geq 3.29$; Brydges et al., 2012). Sixty-four univariate outliers were detected across 29 EF variables. Square root and logarithmic transformations did not normalize outliers, so the variables were left untransformed. Maximum likelihood estimation with robust standard errors (MLR) were used in analyses to account sources of unmodeled heterogeneity, including outliers, non-normality, and non-independence of observations (Hox, Maas, & Brinkhuis, 2010; Maas & Hox, 2004; Maydeu-Olivares, 2017). No multivariate outliers were identified using the computationally powerful "bacon" command in Stata Version 14.2 (Billor, Hadi, & Velleman, 2000; Weber, 2010).

Confirmatory Factor Analysis

A CFA was conducted in Mplus Version 7.4 (Muthén & Muthén, 1998-2015) on the HTKS, Pencil Tap, and EF Touch subtests to determine whether a single unitary factor of EF best fit the data. The hypothesized one-factor solution (in which all measures load onto a single latent factor) was compared to a two-factor solution that separated inhibitory control from working memory and set shifting (following Miller et al., 2012, and Usai et al., 2014). In the two-factor model, Pencil Tap and Pig subtests were regressed on the Inhibitory Control factor, and Pick the Picture and Something's the Same subtests loaded on the Working Memory/Set Shifting factor. HTKS, which assesses aspects of all three EF components (Ponitz et al., 2009), was cross-loaded on both factors. For the two kindergarten time points, the additional EF Touch subscales, Arrows Congruent and Arrows Switch, were regressed on the single latent factor in the unitary model and on the Inhibitory Control factor in the two-factor model. A three-factor model was not compared because (1) most prior literature indicates that EF is a unitary factor in early childhood, with a small subset finding two factors, and (2) there were not enough measures specific to each of the three components to enable an examination of each separately.

In CFA models, the first factor loading (HTKS) was constrained to 1 and the variance of the latent factor was constrained to 1 for identification purposes. For the two kindergarten time points, the unique covariance between the two Arrows subscales (Congruent and Switch) was modeled. In the two-factor model, the latent factors were allowed to correlate with one another, following prior methodology (Fuhs & Day, 2011; Willoughby, Wirth et al., 2017) and consideration that in older samples when components are separable, they are best fit as correlated with one another (e.g., Lehto et al., 2010; Miyake et al., 2000).

Conducting a CFA both minimized measurement error in single measures and enabled comparison of multiple structural models to determine the best fit to the data (Fuhs & Day, 2011). An important consideration for comparing these CFAs and for subsequent tests of measurement invariance is that an analytic sample of 380 is small but sufficient for these types of analyses, given that *N*s of 100 to 200 are considered minimum sample sizes for CFAs (Brown, 2014). Model fit was assessed using Bentler's comparative fit index (CFI), the root-mean-square error of approximation (RMSEA), the standardized root mean square residual (SRMR), and χ^2 tests. Because the latter is sensitive to small sample sizes and relatively few degrees of freedom (Kenny, Kaniskan, & McCoach, 2015; Perry, Nicholls, Clough, & Crust, 2015), the Satorra-Bentler χ^2 difference test was used to account for the small sample, non-normality of data, and complex analytical models (Satorra & Bentler, 2001). Models with CFI greater than or equal to .95 *and* RMSEA less than or equal to .05 are considered to exhibit "good fit," whereas models meeting just one of these criteria or models with CFI greater than or equal to .90 and RMSEA less than .08 are considered to have "adequate fit" (Hu & Bentler, 1999; Fuhs & Day, 2011; Willoughby, Wirth et al., 2012). SRMR less than .08 is indicative of good fit (Chen, Sousa, & West, 2005). Notably, RMSEA is affected by non-normality (Maydeu-Olivares, 2017) and small sample sizes such that it tends to "overreject true models" (Chen, 2007; Hu & Bentler, 1998); as such, in looking across fit statistics, it is logical in this case to prioritize those that adjust for these issues – namely, the CFI and the Satorra-Bentler χ^2 test (Asparouhov, Muthen, & Muthen, 2006; Satorra &

Bentler, 2001; Tabachnick and Fidell, 2007). Models with smaller nonsignificant χ^2 values indicate better fit (Fuhs & Day, 2011; Wiebe et al., 2011). If two models were not significantly different, the simpler model was chosen for parsimony (Fuhs & Day, 2011).

Measurement Invariance across Time

Once the best-fitting model (one- versus two-factor) was chosen, a CFA of that model was conducted for each of the five time points to determine whether measurement invariance of the EF construct over time is established. Establishing invariance indicates that the EF construct is psychometrically equivalent at each time point, which is a prerequisite before moving forward with latent growth curve modeling (Fuhs & Day, 2011; Willoughby, Wirth et al., 2012). Testing measurement invariance involves fitting a set of increasingly restrictive models (Fuhs & Day, 2011; Schmitt et al., 2017) and examining change in goodness-of-fit indices (Cheung & Rensvold, 2002).

In the first, least restrictive model (configural invariance), factor loadings, intercepts, and residual variances were allowed to vary freely. Following prior work, the means of the latent variables were fixed to 0 and the variances were fixed to 1 at all time points (Schmitt, Geldhof, Purpura, Duncan, & McClelland, 2017; Willoughby, Wirth et al., 2012). The model was identified by fixing the loadings of the HTKS measure to 1 (Chen, Sousa, & West, 2005). No additional constraints were imposed on any model parameters. The latent factors were allowed to covary across time points, as was each indicator across time points. Within each kindergarten time point, the unique covariance between the two Arrows subscales (Congruent and Switch) was modeled. Next, scalar (weak) invariance was tested by equating factor loadings across time points. Then, both factor loadings and intercepts were constrained to be equal across time points to assess scalar (strong) invariance. Finally, to test strict measurement invariance, factor loadings, intercepts, *and* residual variances were equated across time points.

Overall model goodness-of-fit was assessed using the CFI, RMSEA, and SRMR. Determining whether these increasingly strict types of measurement invariance held was assessed with change in model fit using the Satorra-Bentler χ^2 difference ($\Delta\ \chi^2$) test. Decreases in CFI greater than .01, changes in RMSEA greater than .015, and significant $\Delta\ \chi^2$ tests indicated significantly worse model fit (Chen, 2007; Cheung & Rensvold, 2002). When the $\Delta\ \chi^2$ test statistic was not significant, the more restrictive model was retained (Fuhs & Day, 2011). If only partial support for measurement invariance is established, overall goodness-of-fit should be considered, and it can still be acceptable to proceed to latent growth curve models with caution (Hughes et al., 2009; Willoughby et al., 2012).

Key Study Questions

Once the latent unitary structure and measurement invariance were tested, key study questions were addressed by conducting longitudinal latent growth curve modeling (LGCM; Duncan & Duncan, 2004; Singer, Willett, & Willett, 2003) across five time points (fall, winter, and spring of preschool and fall and spring of kindergarten) in Mplus Version 7.4. Given the small sample size, there were not enough degrees of freedom to estimate the number of freed parameters needed for latent growth curve models comprised of latent EF factors (Brown, 2014). Thus, to be able to answer key study questions, EF factor scores were calculated in Mplus and included in LGCMs as observed variables. Relatedly, even with observed factor scores there were not enough degrees of freedom to estimate parameters for the quadratic model (Clark et al., 2013; Wiebe et al., 2012). Thus, only the linear (Hughes et al., 2009; Hughes & Ensor, 2011) and freeloading (Willoughby et al., 2012) unconditional LGCMs were compared.

All LGCMs used TYPE=COMPLEX and maximum likelihood estimation with robust standard errors to account for nestedness of children within preschool classrooms. For model identification purposes, the intercepts of the EF factors were fixed to 0, and the variances of the latent intercept and slope were fixed to 1. To make the growth parameters interpretable, the time points were centered at the preschool entry time point and fixed at 0, .28, .54, .98, and 1.52 to represent the average time in years between preschool entry and each subsequent assessment. This accounted for the fact that time intervals between each assessment were not identical.

First, unconditional models (i.e., with no covariates) were run to assess the relative fit of the linear and freeloading functional forms to the data. The linear model included a linear slope term and the time point parameters were specified by the average time in years between assessments as discussed above. In the freeloading model, the first time point (preschool fall) parameter was set to 0 and the last (kindergarten spring) was set to 1. The three intermediate time points were freed to estimate the cumulative proportion of change that occurs between each time point (with 100% of the change occurring between preschool fall and kindergarten spring; Bollen & Curran, 2006). The best-fitting baseline (i.e., unconditional) growth model was determined through a series of nested model comparisons. Overall and relative fit indices were similar to those used in CFA comparisons, with the addition of the following: the Akaike Information Criterion (AIC; Akaike, 2011), Bayesian Information Criterion (BIC; Schwarz, 1978), and sample-sized adjusted BIC (ABIC; Burnham & Anderson, 2004). Relatively lower

AIC, BIC, and ABIC values signify better fit. The CFI was not used to assess relative fit across LGCMs given that this can result in incorrect comparisons across models (Widaman & Thompson, 2003).

Once the best-fitting baseline model (i.e., linear or freeloading) was selected, that model was extended to include child and family characteristics as predictors of the average (i.e., mean) baseline EF at preschool entry (i.e., intercept) and rate of growth in EF across the five time points (i.e., slope), as well as intra-individual variability (i.e., variance) in initial ability and growth rate (Hughes et al., 2009). Specifically, we assessed how children's sex (male = 1), poverty status (in poverty = 1), and deviation from the mean age of their classmates in months (continuous) were associated with the mean and variance of the intercept and slope. These predictive LGCMs controlled for the following covariates: child age and race/ethnicity (dummy codes for White/European American, Hispanic/Latino, and other races, with Black/African American omitted as the reference category).

To obtain more nuanced information about intra-individual variability in the trajectory of EF development (i.e., initial ability and growth rate), multiple group analyses were also performed in Mplus to assess whether and how the trajectory of EF development (as determined by the best-fitting functional form) differed for boys versus girls and for children in poverty versus not in poverty. To determine whether the mean intercepts and slopes were statistically significantly different between each of the groups in these two analyses, we constrained the intercept and then the slope in follow-up analyses to assess whether this resulted in a significant reduction in fit (i.e., significant Δ χ^2 test *p*-value).

Results

Confirmatory Factor Analyses

Table 3 displays results of CFAs comparing one- and two-factor models of the EF construct across the five time points. The fit of the one-factor model was at least acceptable at all time points (i.e., CFI ≥ .95 or RMSEA/SRMR ≤ .05, or CFI ≥ .90 and RMSEA/SRMR ≤ .08). The two-factor model demonstrated acceptable fit in kindergarten fall (CFI = .945, RMSEA = .063, SRMR = .03) and good fit (i.e., CFI ≥ .95 and RMSEA/SRMR ≤ .05) at all other time points. Although the two-factor model statistically fit the data better than the one-factor model in preschool fall, $\Delta\ \chi^2$ (2) = 7.36, p = 0.025, and winter, $\Delta\ \chi^2$ (2) = 13.69, p = .001, there were no significant differences between the two models for the latter three time points: preschool spring $\Delta\ \chi^2$ (2) = 5.61, p = .06, kindergarten fall $\Delta\ \chi^2$ (2) = 0.75, p = .69, and kindergarten spring $\Delta\ \chi^2$ (2) = 3.85, p = .15.

Given this statistical equivalence for the majority of the time points and the acceptable overall model fit of the one-factor model at all time points, conceptually based on prior literature (e.g., Hughes et al., 2009; Willoughby et al., 2012), and for parsimony (Fuhs & Day, 2011), the unitary model of EF at each time point was retained in all subsequent analyses.

Measurement Invariance across Time

The configural invariance model fit the unitary EF factor adequately, χ^2 (313) 461.38, p = < .001, CFI = .946, RMSEA = .036, SRMR = .064. Constraining the factor loadings (metric invariance) resulted in significantly worse fit, $\Delta\ \chi^2$ (18) = 99.31, p < .001, CFI = .907, RMSEA = .045, SRMR = .18, specifically because Pencil Tap was found to be invariant across time. Thus, increasingly restrictive models had poor, significantly worse fit: scalar invariance $\Delta\ \chi^2$ (18) = 80.58, $p < .001$, CFI = .881, RMSEA = .05, SRMR = .113; residual invariance $\Delta\ \chi^2$ (17) = 69.30, $p < .001$, CFI = .845, RMSEA = .056, SRMR = .316.

Freeing the factor loading – and subsequently the intercept and residual variance – for Pencil Tap resulted in improved fit that provided evidence for adequate partial scalar invariance: partial metric invariance, $\Delta\ \chi^2$ (14) = 37.18, $p < .001$, CFI = .936, RMSEA = .038, SRMR = .111; partial scalar invariance, $\Delta\ \chi^2$ (14) = 46.05, $p < .001$,; CFI = .923, RMSEA = .041, SRMR = .105; partial residual invariance, $\Delta\ \chi^2$ (14) = 44.51, $p < .001$, CFI = .901, RMSEA = .045, SRMR = .324. Although all of the $\Delta\ \chi^2$ tests were significant (indicating worsening fit), based on other metrics (i.e., decrease in CFI ≤ .01, change in RMSEA ≤ .015; Cheung & Rensvold, 2002) there was evidence for adequate partial scalar invariance.

However, because Pencil Tap was not invariant across time, it was excluded from subsequent LGCM analyses. Table 4 presents measurement invariance statistics for the unitary EF factor sans Pencil Tap, comprised of HTKS and EF Touch Pig, Pick the Picture, Something's the Same, and Arrows Congruent and Switch. Despite significant $\Delta\ \chi^2$ tests comparing increasingly restrictive models, there was evidence for scalar invariance based on the small change in RMSEA (< .015; Cheung & Rensvold, 2002). Moreover, the overall fit of the scalar invariance model was adequate (CFI = .942, RMSEA = .036). Additionally, based on prior literature, it is acceptable to move forward cautiously to LGCMs with adequate scalar invariance (Willoughby et al., 2012). Figure 1 displays a graphical representation of the final unitary latent EF factors across the five time points, including standardized factor loadings and residual variances, as well as correlations among the latent factors.

Latent Growth Curve Models

Table 5 presents results of these unconditional models. The freeloading model fit the data adequately, χ^2 (7) = 56.16, $p < .001$, SRMR = .077, although the RMSEA was unacceptably high (RMSEA = .135). This may have been due in part to the small sample size and/or non-normality (Chen, 2007; Hu & Bentler, 1998; Maydeu-Olivares, 2017). The freeloading model fit the data better than the linear model, based on lower AIC, BIC, and ABIC and a significant χ^2 difference test, $\Delta\ \chi^2$ (3) = 30.32, $p < .001$. This indicated that the significant EF growth from preschool to kindergarten found in this study (mean slope = 2.66, $p < .001$) was best characterized as nonlinear.

Figure 2 presents a graphical representation of the unstandardized freeloading slope estimates across time. This shows that 29.40% of EF growth occurred between preschool fall and winter, 18.50% between preschool winter and spring, 27.5% between preschool spring and kindergarten fall, and 24.6% between kindergarten fall and spring. In other words, over 75% of EF growth occurred before kindergarten entry, which as hypothesized indicates more rapid EF growth in preschool with a relative slowing toward the end of kindergarten. There was a negative correlation between the intercept and slope (r = -.32, $p < .001$), suggesting that starting preschool with lower EF was associated with faster EF growth from preschool to kindergarten. Findings also indicated that there was significant intra-individual variability in initial EF ability at preschool entry (intercept variance = .33, $p < .001$) and the rate of EF growth over time (slope variance = .32, $p <$

.001). This variability was explored further by adding demographic predictors and covariates to the freeloading LGCM.

Figure 3 displays a graphical depiction of this extended LGCM with standardized parameter estimates for intercepts and slopes. Compared to girls, boys began preschool with significantly lower EF ability (estimate = -.25, $p < .001$). Although not statistically significant, there was a trending association between poverty status and intercept suggesting that children in poverty may begin preschool with lower EF ability than non-impoverished peers (estimate = -.13, $p = .05$). There was no significant association between initial EF ability and children's age relative to the average age of their classmates. None of the three variables of interest was significantly associated with the slope of EF. However, a non-hypothesized finding was that compared to African American children, European American children had a faster rate of EF growth from preschool to kindergarten (estimate = .25, $p < .001$). No other covariates were significantly related to the intercept or slope. Overall, the model accounted for 17% of the variability in initial EF ability and 12% of the variability in rate of EF growth over time.

Multiple Group Analyses

Table 6 presents parameter estimates for multiple group analyses by child sex and by poverty status. Table 7 displays fit statistics for unconstrained MGA models compared to those that constrain the intercept and slope. Similar to LGCMs, MGA models tested associations between EF and demographic characteristics, using a slightly different method that provided more nuance regarding the nature of demographic differences in EF. Specifically, whereas the LGCM results showed whether there were significant differences based on demographic characteristics, MGA models allowed for direct comparisons of intercept and slope differences across groups to be able to characterize those differences.

Poverty status. When parameter estimates were separated by poverty status, it appeared descriptively as though children in poverty began preschool with lower EF ability (estimate = -1.51) than non-impoverished peers (estimate = -0.94). However, similar to the predictive LGCM results, this was only a marginally significant finding as evidenced by the χ^2 difference test, $\Delta\chi^2$ (1) = 2.83, p = .09. Unlike the predictive LGCM, the MGA model fit significantly worse when the slopes were constrained to be equal across groups, $\Delta\ \chi^2$ (1) = 8.28, p = .004. This provides some evidence – although inconsistently so – that children who are not in poverty may develop EF more quickly (estimate = 2.94) than their impoverished peers (estimate = 2.42) from preschool to kindergarten.

Child sex. Similar to the predictive LGCM results, when parameter estimates were separated by child sex it appeared as though boys began preschool with lower EF ability (estimate = -0.86) than girls (estimate = -1.50) but that their rate of growth was almost identical (female estimate = 2.67, male estimate = 2.64). This interpretation was supported by the fact that the model fit decreased significantly when the intercept was constrained to be equal for boys and girls, $\Delta\ \chi^2$ (1) = 15.97, $p < .001$, but not when the slope was constrained to be equal across groups, $\Delta\ \chi^2$ (1) = 1.94, p = .16.

When covariates were included in multiple group analysis models, these differences became more nuanced. Girls in poverty had lower initial EF than their non-impoverished female peers (estimate = -0.199, p = .024), but poverty was not associated with EF growth for girls. For boys, poverty status was not associated with initial EF but boys in poverty had significantly shallower growth trajectories than more affluent male peers (estimate = -0.27, p = .02). These differential poverty-EF associations across sexes could account for the marginally significant findings for EF intercept and the inconsistent findings for EF slope. There was also a differential finding by race/ethnicity across sexes: whereas European American girls had marginally steeper EF trajectories than African American female peers (estimate = 0.239, p = .051), European American boys had significantly more growth than African American male peers (estimate = 0.264, p = .038). Importantly, although this finding for girls was marginally rather than statistically significant, point estimates were very similar (.24 for girls compared to .26 for boys), indicating that this finding should be interpreted – albeit cautiously – across both sexes.

Discussion

EF is a foundational ability that develops rapidly in early childhood (Garon et al., 2008) and has been linked to school success in the short- and long-term (Baggetta & Alexander, 2016; Blair, 2016; Zelazo et al., 2016), presumably in part because it allows them to benefit more from learning opportunities (Morgan et al., 2018). Examining how EF develops during the transition from preschool to kindergarten – and how this might look dissimilar for different children – has important implications for research and practice. This includes further unpacking the construct of EF and how it unfolds over time, as well as identifying potentially advantageous timing for EF interventions and which children might differentially benefit from them. As such, the present study leveraged data across five time points in a primarily low-income, racially and ethnically diverse sample of children to investigate the latent construct of EF, the trajectory of its growth through kindergarten, and demographic differences in EF development.

As hypothesized, both a unitary and a two-factor EF construct fit the data well. The one-factor model was selected because data indicated that the one-factor model provided a superior fit at the majority of time points, based on prior literature, and for parsimony. This one-factor construct of EF demonstrated adequate scalar invariance, although model fit was not as high as has been found in previous studies; potential reasons for this are explored below. Consistent with expectations, children demonstrated significant, non-linear growth in EF from preschool to kindergarten, with approximately 75% of growth occurring by kindergarten entry. There was significant intra-individual variability in children's EF at preschool entry and their growth over time, which was in part explained by demographic differences. As hypothesized, boys – and children in poverty, to a marginally significant degree – entered preschool with lower EF than girls and more affluent peers. In LGCM analyses, child sex, poverty status, and relative age were not associated with the rate of EF growth; a non-hypothesized finding was that African American children tended to have a lower rate of growth than European American peers. Results of multiple group analyses corroborated LGCM findings for child sex and indicated that children in poverty may have a faster rate of growth than non-impoverished peers. An unexpected, null finding was that children's age relative to their classmates was not associated with their initial EF ability or growth in EF over time.

Adequate Support for EF as a Unitary Factor across Time

Notably, CFA results indicated that both the one- and two-factor models of EF fit the data well and for the majority of time points were statistically indistinguishable. The one-factor model was selected in this study because of (1) its overall good model fit across time points, (2) the fact that it was not statistically distinct from the two-factor model (and thus, the preferred model) at the majority of time points, (3) relatively high correlations between the factors in the two-factor model, (4) consistency with the majority of prior research, and (5) for parsimony. As such, the present study supports prior research that has found EF to be a unitary construct (Wiebe et al., 2008, 2011; Willoughby, Blair et al., 2012; Zelazo et al., 2016). However, the fact that both the unitary and bipartite constructs of EF fit the data well does not discount prior work that focused on a two-factor model of EF (Lee et al., 2013; Miller et al., 2012; Usai et al., 2014) and indicates that more longitudinal EF measurement work is needed, especially across this key early childhood transition period. Following calls from prior work, it will continue to be important to understand the developmentally graded structure of EF to further clarify conceptual and methodological ambiguity (Baggetta & Alexander, 2016; Jones et al., 2016; Zelazo et al., 2016).

Although findings were generally consistent with hypotheses, it was unexpected that when the two-factor model did fit significantly better, it was during preschool fall and winter rather than later in development, when the EF construct has been hypothesized and found to become more differentiated (Lee et al., 2013). This finding calls into question whether the significant difference between the one- and two-factor model early in preschool and the subsequent superiority of the unitary construct through kindergarten reflected actual developmental changes in the construct of EF or whether it was primarily related to limitations in measurement. Previous work has suggested that we may not yet have the tools to be able to precisely assess and adequately differentiate between EF subcomponents at different ages (Jones et al., 2016). Relatedly, in this study Pencil Tap was not found to be invariant over time, indicating that this task may be assessing a different construct at different time points between preschool and kindergarten. It is possible that early in preschool this measure may load more on certain skills such as motor control and auditory comprehension, and that it is more strongly based on inhibitory control later in preschool and in kindergarten. This possibility is partially supported by the fact that there was a bimodal distribution for Pencil Tap in the fall of preschool, such that the two most common performances were close to 0% correct and close to 100% correct, whereas at all subsequent time points, the distribution was unimodal and negatively skewed. The implications of this for future research are that perhaps Pencil Tap may be a more appropriate tool for assessing inhibitory control when children on average have adequate motor control, attention, and comprehension. Future work should continue to explore age effects on Pencil Tap performance, given that it is a commonly used early childhood measure.

Prior to interpreting LGCM results, it is essential to recognize that the fit of measurement invariance and LGCMs was generally adequate, but not strong. Moreover, it is important to consider potential conceptual and technical reasons for less than ideal fit. Research on the construct of EF has been plagued by a "measurement impurity problem" given its complexity, such that measures purporting to assess EF also tend to assess other cognitive functions (e.g., motor skills, processing speed, and language) and often assess multiple subcomponents of EF simultaneously (Anderson & Reidy, 2012; Baggetta & Alexander, 2016; Fuhs et al., 2014; Miyake et al., 2000; Zelazo et al., 2016). Notably, the makeup of how these various skills load onto task performance may change over time, such as was discussed above for Pencil Tap. Ceiling effects may be another potential reason why Pencil Tap did not demonstrate measurement invariance (kindergarten spring *M* = .91, *SD* = .17, possible range = 0 – 1). Another measure, EF Touch Pig, also demonstrated ceiling effects by spring of kindergarten (*M* = .96, *SD* = .05, possible range = 0 - 1). It is arguable that rather than a measurement limitation, these ceiling effects may be reflective of an underlying developmental change in the construct of EF, such that it needs to be assessed with different, more complex tasks over time. This study introduced two additional inhibitory control subscales during the kindergarten time points to account for this and make the construct more developmentally graded. An argument against this hypothesis, however, is that EF Touch has been found to have good measurement invariance over time (Willoughby et al., 2012). Notably, however, the present study used a smaller number of EF Touch subtests and added the HTKS (the other additional measure, Pencil Tap, was excluded because the measure varied across time). For this reason and because of their demographically distinct samples, measurement invariance statistics are not directly comparable across the two studies.

Another measurement consideration was that many of the measures tended to be skewed and/or kurtotic (see Table 2). Despite the fact that MLR estimation was used to adjust for non-normality, fit statistics such as the χ^2 difference test and the RMSEA tend to be affected by non-normality such that they are more likely to overreject a true model in these cases; these fit statistics are also affected by sample size (Chen et al., 2005; Cheung & Rensvold, 2002; Kenny, Kaniskan, & McCoach, 2015; Maydeu-Olivares, 2017). Thus, it was important to consider other metrics of model fit when examining measurement invariance. Based on change in CFI, RMSEA, and SRMR, there was evidence for adequate scalar invariance, particularly once Pencil Tap was removed from analyses. Achieving adequate scalar invariance was necessary to be able to interpret significant growth in EF in LGCMs as actual developmental change over time, rather than as merely reflective of measurement artifact (Chen et al., 2005). More globally, it is possible that using observed factor scores rather than latent EF factors at each time point introduced additional measurement error into the model that impacted fit; this was a limitation of the sample size in the context of the parameters that needed to be estimated to answer key research questions.

The Trajectory of EF Development from Preschool to Kindergarten

Consistent with hypotheses, there was significant, nonlinear growth in EF, with steeper development in preschool than in kindergarten. This aligns with prior work examining the functional form of EF in early childhood (Hughes et al., 2009; Hughes & Ensor, 2011; Wiebe et al., 2012; Willoughby, Wirth et al., 2012). Specifically, Willoughby and colleagues (2012) followed children from ages 3 to 5 and found that 60 percent of growth in EF occurred between ages 3 and 4; two additional studies found similarly higher growth rates from 3 to 4 than 4 to 5 (Clark et al., 2013; Wiebe et al., 2012). In this study, which followed children from 4 to 6, 75 percent of growth occurred in the first year (i.e., between ages 4 and 5). Together, this indicates a trend that children make relatively more gains earlier in their development (when compared to a year later). This is consistent with the notion that EF develops rapidly in early childhood (Blair, 2016; Garon et al., 2008) and suggests a sensitive period for EF development during that time.

Although growth was found to be nonlinear, the present study was unable to make more specific claims about the functional form of EF from preschool to kindergarten, because the small sample size prohibited modeling of quadratic or cubic forms, which required more parameters than the available degrees of freedom. Previous work has found EF growth in early childhood to be quadratic (Clark et al., 2013; Wiebe et al., 2012) or exponential (Montroy et al., 2016), both with accelerating growth early on and decelerating growth later in childhood. Given the relative percentage of growth found in preschool and kindergarten in the present study, it is likely that these functional forms would provide a good fit to these data as well. Notably, this slower growth toward the end of kindergarten could represent a failure of current measures to adequately capture continued growth as children's EF abilities develop. Conversely, it may be representative of a sensitive period in early childhood where growth is more rapid earlier on in children's development (Garon et al., 2008).

If the latter is accurate, one practical implication of this relatively slower growth in kindergarten is that it may indicate a potentially beneficial time for intervention aimed at promoting children's EF, particularly given the shifts in routines and expectations and increases in demands that occur during the transition to kindergarten (Bassok et al., 2016; Jones et al., 2016; Rimm-Kaufman & Pianta, 2000). Although findings from prior studies (Clark et al., 2013; Wiebe et al., 2012; Willoughby et al., 2012) indicate that relatively more growth occurred between ages 3 and 4 than 4 and 5, the kindergarten year may still be a beneficial time and setting for intervention given that 79 percent of 3- to 5-year-old children were enrolled in full-day kindergarten programs as of 2017 (National Center for Education Statistics, 2019), compared to 42 percent of three-year-olds and 66 percent of 4-year-olds enrolled in preprimary education in 2016 (McFarland et al., 2018). Conversely, this slowed growth may suggest that children in general are less sensitive to EF-focused intervention in kindergarten than they are in preschool. Future work examining the relative efficacy of EF interventions in preschool and kindergarten contexts could further elucidate whether either year may provide a relatively more beneficial time period to bolster the development of this foundational skillset.

Notably, almost 28 percent of EF growth occurred over the summer between preschool spring and kindergarten fall, when it is less likely that children would have been engaged in learning opportunities. It is important to note that data were not collected at the tail end or very beginning of the school year to allow for transition time; thus, some of this observed learning could have occurred in the last few weeks of preschool and/or the first few weeks of kindergarten and skills that were actually gained over the summer may be overstated (Alexander, Entwisle, & Olson, 2007). However, this may suggest that EF is less subject to "summer learning loss" (Stewart, Watson, & Campbell, 2018, p. 517) – which disproportionately affects children from low-income backgrounds who may have less access to learning opportunities outside of school than more affluent peers – than academic skills, perhaps because EF is less dependent on formal instruction. It would be useful to have information about children's summer activities and/or to compare the nonlinear growth of EF and academic skills to be able to test this hypothesis.

Demographic Differences in the Trajectory of EF Development

In this study, there was significant intra-individual variability in children's EF at preschool entry and their rate of skill growth over time that was partly explained by demographic characteristics. Consistent with prior research (Conway et al., 2018; Little, 2017; Raver et al., 2013), children in poverty began preschool with lower EF than non-impoverished peers. This may be due to the fact that low-SES families – and especially families in poverty – have been shown to have increased stress (Bradley & Corwyn, 2002), which negatively affects children's skill development (Evans & Kim, 2013; Thompson, 2014) through physiological mechanisms by disrupting their cortisol reactivity (Blair, Granger, & Peters Razza, 2005) and/or psychosocial pathways through less responsive parenting practices (Gershoff, Aber, Raver, & Lennon, 2007; Lucassen et al., 2015). Similar to extant research on poverty and the trajectory of EF development in early childhood, findings were inconsistent across two model specifications, although they trended in the same direction. In LGCM analyses there was a negative, albeit non-significant association between poverty status and EF growth, whereas in multiple group analyses, children in poverty had slower rates of growth than non-impoverished peers. In the latter analysis, the small sample size and non-normally distributed data may have led the χ^2 difference test to overreject the model (i.e., find that the model fit significantly worse when the slopes were constrained to be equal across groups). Indeed, the χ^2 difference test statistic was relatively small (8.28), albeit statistically significant (p = .004). Thus, this association between poverty status and rate of EF growth should be interpreted with caution; even if it is a statistically significant difference, it may not be practically or clinically significant. Findings suggest that more work is needed to understand how being in poverty affects the way that children's EF develops over time.

With more certainty and in partial support of hypotheses, this study found that boys began preschool with lower EF than girls but that their rate of growth over time was similar. This sex-related gap has been well characterized in academic outcomes (Matthews, Ponitz, & Morrison, 2009; Pomerantz, Altermatt, & Saxon, 2002). Findings from this study may explain one way through which this gap occurs, because EF has been shown to predict academic abilities such as language, literacy, and math (Blair, 2016;

Fuhs et al., 2014; McClelland et al., 2014; Morgan et al., 2018). Although not a main focus of the study, one non-hypothesized finding with the covariate of race/ethnicity indicated that European American children tended to have faster rates of EF growth than African American peers, particularly for boys. This has long-term implications, as literacy gaps for African American boys in fifth grade have been explained in part by their kindergarten learning-related skills that are associated with EF (e.g., attentiveness, task persistence, organization; Matthews, Kizzie, Rowley, & Cortina, 2010).

Based on prior research, it may be that biases and discrimination inside and outside the classroom could be compromising the EF of African American children in particular. For instance, teachers tend to rate young African American students as having lower frustration tolerance and task orientation than European American peers (Sbarra & Pianta, 2001) and favor European American students in terms of interactions, positive expectations, and referrals for services (Tenenbaum & Ruck, 2007). Notably, when African American students and their kindergarten teachers have closer and less conflictual relationships, the students tend to have more positive social-emotional outcomes in elementary and middle school (Hamre & Pianta, 2001; Iruka, Burchinal, & Cai, 2010). Later in life, exposure to racism has been shown to compromise older African American students' ability to demonstrate EF skills (Bair & Steele, 2010). Together with findings from this study, this points to the need to provide additional, culturally sensitive support for the EF development of African American children, particularly boys.

Although this study found significant demographic differences in how children's EF develops in early childhood, it is important to note that overall, the model accounted for only 17 percent of the variability in EF at preschool entry and 12 percent of the variability in the rate of EF growth over time. Thus, there are many other factors besides demographic characteristics that were not included in this study that could potentially affect the development of children's EF. Given the relationship between family stress and EF (Evans & Kim, 2013; Lucassen et al., 2007) – particularly for children in poverty (Raver et al., 2013), who comprised almost half of this study's sample – examining the role of family factors such as parenting practices could provide additional insight into how children develop this important foundational skillset (Bernier, Carlson, & Whipple, 2010). Future research might also examine classroom factors that may be associated with children's EF, including the ways in which children engage with the teachers, peers, and tasks in their classrooms (Williford, Whittaker, Vitiello, & Downer, 2013).

Limitations and Future Directions

It is essential to interpret the contributions of this study's finding within the context of a number of limitations. Although the sample was racially and ethnically diverse, children were primarily from low-income households (and approximately half were in poverty). It is important to understand the experiences of this at-risk demographic, but this limits generalizability to socioeconomically disadvantaged children. Future research might leverage a mixed-income sample to understand the trajectory of EF development and its demographic correlates more broadly.

Another threat to generalizability was the fact that over half of the sample (roughly 59%) was missing data in the spring of kindergarten because of planned missingness. Given that these data were conceptualized as being missing completely at random and there was rich information about children who were not assessed at that time point, data could be multiply imputed. However, because more than half of these data were imputed, results can only offer suggestions about what EF looks like at the end of kindergarten and about the trajectory of nonlinear growth from preschool entry to the end of kindergarten. This is an important limitation in the context of extant EF literature, because there is not yet consensus about nonlinear growth in EF within the preschool and kindergarten years. However, it was possible to obtain a nuanced understanding of the trajectory of EF development during preschool because of the three data collection windows in that year. For instance, there appeared to be more growth in preschool from fall to winter than from winter to spring. In future studies, it would be helpful to have two consecutive years with fall, winter, and spring data for all children to assess whether the within-year growth trajectory looks similar in kindergarten.

The strength of having five time points – thus, enabling the possibility of examining quadratic and cubic functional forms of EF development – was ultimately limited by the small analytic sample size ($N = 380$) whose degrees of freedom were not sufficient to estimate the parameters needed for these more specific nonlinear forms (Brown, 2014). Future studies should leverage larger datasets to examine EF development across at least five time points to characterize its functional form more specifically. Another measurement limitation was that there were not enough distinct measures of the three conceptual subcomponents of EF (i.e., inhibitory control, working memory, and set shifting) to be able to examine whether a three-factor model of EF fit the data well. Although this would be unexpected in early childhood, it would be beneficial going forward to compare a three-factor model to the unitary and bipartite models, particularly given continued debate in the field about the construct of EF in early childhood.

Finally, the less-than-ideal (although ultimately adequate) model fit for CFAs and LGCMs requires that interpretations be made with caution. Specifically, these findings may overstate the extent to which EF is definitively a unitary construct in early childhood, particularly given that the bipartite model also fit the data well. It may also overstate the specificity of the nonlinear trajectory of EF development from preschool to kindergarten (e.g., accelerated and then decelerated growth), particularly given that more specific nonlinear functional forms were not tested. In future studies, fit might improve by testing these more specific forms (e.g., quadratic, exponential) that were unable to be tested in this study given sample size restrictions, or by using latent factors of EF in LGCMs rather than observed factor scores.

Conclusion

Despite its limitations, the present study makes multiple contributions to the literature by examining demographic differences in EF development across five time points in a primarily low-income, ethnically and racially diverse sample. Specifically, findings provide further evidence that a unitary construct of EF develops nonlinearly in early childhood, as is hypothesized in the majority of prior work (e.g., Hughes & Ensor, 2009; Willoughby et al., 2012). Further demographic differences found in developmental trajectories of EF may offer one potential explanation for achievement gaps based on income, sex, and race/ethnicity that have significant implications for children's development over time. Practically, findings point to kindergarten as a potentially beneficial time for intervention to continually bolster EF skills, given decelerated growth found during this time period, and suggest that children in poverty, boys, African American children, and particularly African American boys may benefit from additional opportunities to develop this skillset. Finally, the present study addresses the "measurement impurity problem" of EF research in the context of its findings and limitations, and suggests ways future research can further elucidate this key foundational school readiness skill.

References

Akaike, H. (2011). Akaike's information criterion. In M. Lovric (Ed.), *International Encyclopedia of Statistical Science* (pp. 25-25). Berlin, Germany: Springer-Verlag Berlin Heidelberg.

Alexander, K. L., Entwisle, D. R., & Olson, L. S. (2007). Lasting consequences of the summer learning gap. *American Sociological Review*, *72*, 167-180.

Ansari, A. (2017). Multigrade kindergarten classrooms and children's academic achievement, executive function, and socioemotional development. *Infant and Child Development*, *26*, e2036. doi: 10.1002/icd.2036.

Ansari, A., & Purtell, K. M. (2018). What happens next? Delivering on the promise of preschool. *Early Childhood Research Quarterly*. Advance online publication. doi: 10.1016/j.ecresq.2018.02.015.

Ansari, A., Purtell, K., & Gershoff, E. (2016). Classroom age composition and the school readiness of 3-and 4-year-olds in the Head Start program. *Psychological Science*, *27*, 53-63. doi: 10.1177/0956797615610882.

Asparouhov, T., Muthén, B., & Muthén, B. O. (2006). Robust chi square difference testing with mean and variance adjusted test statistics. *Matrix*, *1*, 1-6.

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

Bair, A. N., & Steele, J. R. (2010). Examining the consequences of exposure to racism for the executive functioning of Black students. *Journal of Experimental Social Psychology*, *46*, 127-132. doi: 10.1016/j.jesp.2009.08.016.

Bassok, D., Latham, S., & Rorem, A. (2016). Is kindergarten the new first grade?. *AERA Open*, *2*, 2332858415616358. doi: 10.1177/2332858415616358.

Bauermeister, J. J., Shrout, P. E., Chávez, L., Rubio-Stipec, M., Ramírez, R., Padilla, L., ... & Canino, G. (2007). ADHD and gender: are risks and sequela of ADHD the same for boys and girls?. *Journal of Child Psychology and Psychiatry*, *48*, 831-839. doi: 10.1111/j.1469-7610.2007.01750.x.

Bernier, A., Carlson, S. M., & Whipple, N. (2010). From external regulation to self-regulation: Early parenting precursors of young children's executive functioning. *Child Development*, *81*, 326-339. doi: 10.1111/j.1467-8624.2009.01397.x

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

Blair, C. (2016). Executive function and early childhood education. *Current Opinion in Behavioral Sciences*, *10*, 102-107. doi: 10.1016/j.cobeha.2016.05.009.

Blair, C., Granger, D., & Peters Razza, R. (2005). Cortisol reactivity is positively related to executive function in preschool children attending Head Start. *Child Development*, *76*, 554-567. doi: 10.1111/j.1467-8624.2005.00863.x

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*, (Vol. 2, pp. 300-320). New York, NY: Guilford Press.

Blair, C., Zelazo, P. D., & Greenberg, M. T. (2005). The measurement of executive function in early childhood. *Developmental Neuropsychology, 28,* 561-571. doi: 10.4324/9780203764244.

Billor, N., Hadi, A. S., & Velleman, P. F. (2000). BACON: blocked adaptive computationally efficient outlier nominators. *Computational Statistics & Data Analysis*, *34*, 279-298. doi: 10.1016/S0167-9473(99)00101-2.

Bollen, K. A., & Curran, P. J. (2006). *Latent curve models: A structural equation perspective* (Vol. 467). Hoboken, NJ: John Wiley & Sons.

Bradley, R. H., & Corwyn, R. F. (2002). Socioeconomic status and child development. *Annual Review of Psychology*, *53*, 371-399. doi: 10.1146/annurev.psych.53.100901.135233

Brown, T. A. (2014). *Confirmatory factor analysis for applied research* (2nd ed.). New York, NY: Guilford Publications.

Brydges, C. R., Reid, C. L., Fox, A. M., & Anderson, M. (2012). A unitary executive function predicts intelligence in children. *Intelligence*, *40*, 458-469. doi:10.1016/j.intell.2012.05.006.

Burnham, K. P., & Anderson, D. R. (2004). Multimodel inference: Understanding AIC and BIC in model selection. *Sociological Methods and Research*, 33, 261–304. doi: 10.1177/0049124104268644.

Carlson, S. M. (2005). Developmentally sensitive measures of executive function in preschool children. *Developmental Neuropsychology*, *28*, 595-616. doi: 10.1207/s15326942dn2802_3.

Chen, F. F. (2007) Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, *14*, 464-

504. doi: 10.1080/10705510701301834.

Chen, F. F., Sousa, K. H., & West, S. G. (2005). Teacher's corner: Testing measurement invariance of second-order factor models. *Structural Equation Modeling*, *12*, 471-492. doi: 10.1207/s15328007sem1203_7.

Clark, S. L., & Muthén, B. (2009). Relating latent class analysis results to variables not included in the analysis. Manuscript submitted for publication. Retrieved from: www.statmodel.com/download/relatinglca.pdf.

Clark, C.A.C., Sheffield, T.D., Chevalier, N., Nelson, J.M., Wiebe, S.A., and Espy, K.A. (2013). Charting early trajectories of executive control with the shape school. *Developmental Psychology, 49,* 1481-1493. doi: 10.1037/a0030578.

Conway, A., Waldfogel, J., & Wang, Y. (2018). Parent education and income gradients in children's executive functions at kindergarten entry. *Children and Youth Services Review, 91*, 329-337. doi: 10.1016/j.childyouth.2018.06.009.

Cook, R. D., & Weisberg, S. (1982). *Residuals and influence in regression*. New York: Chapman and Hall.

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

Duncan, T. E., & Duncan, S. C. (2004). An introduction to latent growth curve modeling. *Behavior Therapy*, *35*, 333-363. doi: 10.1016/S0005-7894(04)80042-X.

Felt, J. M., Depaoli, S., & Tiemensma, J. (2017). Latent growth curve models for biomarkers of the stress response. *Frontiers in Neuroscience, 11*, 315-332. doi: 10.3389/fnins.2017.00315.

Finch, J. E., & Obradović, J. (2017). Unique effects of socioeconomic and emotional parental challenges on children's executive functions. *Journal of Applied Developmental Psychology*, *52*, 126-137. doi:10.1016/j.appdev.2017.07.004.

Friedman, N. P., & Miyake, A. (2017). Unity and diversity of executive functions: Individual differences as a window on cognitive structure. *Cortex*, *86*, 186-204.

Fuhs, M. W., & Day, J. D. (2011). Verbal ability and executive functioning development in preschoolers at Head Start. *Developmental Psychology*, *47*, 404-416. doi: 10.1037%2Fa0021065.

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*, 1698-1709. doi: 10.1037/a0036633.

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

Gershoff, E. T., Aber, J. L., Raver, C. C., & Lennon, M. C. (2007). Income is not enough: Incorporating material hardship into models of income associations with parenting and child development. *Child Development*, *78*, 70-95. doi: 10.1111/j.1467-8624.2007.00986.x.

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

Hamre, B. K., & Pianta, R. C. (2001). Early teacher–child relationships and the trajectory of children's school outcomes through eighth grade. *Child Development*, *72*, 625-

638. doi: 10.1111/1467-8624.00301.

Hix-Small, H., Duncan, T., Duncan, S., & Okut, H. (2004). A multivariate associative finite growth mixture modeling approach examining adolescent alcohol and marijuana use. *Journal of Psychopathology and Behavioral Assessment*, 26, 255–270. doi: 10.1023/B:JOBA.0000045341.56296.fa.

Hongwanishkul, D., Happaney, K. R., Lee, W. S., & Zelazo, P. D. (2005). Assessment of hot and cool executive function in young children: Age-related changes and individual differences. *Developmental Neuropsychology*, *28*, 617-644. doi: 10.1207/s15326942dn2802_4.

Hox, J. J., Maas, C. J., & Brinkhuis, M. J. (2010). The effect of estimation method and sample size in multilevel structural equation modeling. *Statistica Neerlandica*, *64*, 157-170. doi: 10.1111/j.1467-9574.2009.00445.x.

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*, 1-55. doi: 10.1080/10705519909540118.

Hughes, C., & Ensor, R. (2011). Individual differences in growth in executive function across the transition to school predict externalizing and internalizing behaviors and self-perceived academic success at 6 years of age. *Journal of Experimental Child Psychology*, *108*, 663-676. doi: 10.1016/j.jecp.2010.06.005.

Hughes, C., Ensor, R., Wilson, A., & Graham, A. (2009). Tracking executive function across the transition to school: A latent variable approach. *Developmental Neuropsychology*, *35*, 20-36. doi: 10.1080/87565640903325691.

Iruka, I. U., Burchinal, M., & Cai, K. (2010). Long-term effect of early relationships for African American children's academic and social development: An examination from kindergarten to fifth grade. *Journal of Black Psychology*, *36*, 144-171. doi: 10.1177/0095798409353760.

Isquith, P. K., Gioia, G. A., & Espy, K. A. (2004). Executive function in preschool children: Examination through everyday behavior. *Developmental Neuropsychology*, *26*, 403-422. doi: 10.1207/s15326942dn2601_3.

Jones, S. M., Bailey, R., Barnes, S. P., & Partee, A. (2016). *Executive function mapping project: Untangling the terms and skills related to executive function and self regulation in early childhood* (No. 2016-88). Cambridge, MA: U.S. Department of Health and Human Services, Office of Planning, Research, and Evaluation.

Kalkut, E. L., Han, S. D., Lansing, A. E., Holdnack, J. A., & Delis, D. C. (2009). Development of set-shifting ability from late childhood through early adulthood. *Archives of Clinical Neuropsychology*, *24*, 565-574. doi: 10.1093/arclin/acp048.

Kenny, D. A., Kaniskan, B., & McCoach, D. B. (2015). The performance of RMSEA in models with small degrees of freedom. *Sociological Methods & Research*, *44*, 486-507. doi: 10.1177/0049124114543236.

Lawson, G. M., & Farah, M. J. (2017). Executive function as a mediator between SES and academic achievement throughout childhood. *International Journal of Behavioral Development*, *41*, 94-104. doi: 10.1177/0165025415603489.

Lawson, G. M., Hook, C. J., & Farah, M. J. (2018). A meta-analysis of the relationship between socioeconomic status and executive function performance among children. *Developmental Science*, *21*, e12529. doi: 10.1111/desc.12529.

Lee, K., Bull, R., & Ho, R. M. (2013). Developmental changes in executive functioning. *Child Development*, *84*, 1933-1953. doi: 10.1111/cdev.12096.

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

Li, C. (2013). Little's test of missing completely at random. *Stata Journal*, *13*, 795-809. doi: 10.1177/1536867X1301300407.

Lipsey, M. W., Farran, D. C., & Durkin, K. (2018). Effects of the Tennessee Prekindergarten Program on children's achievement and behavior through third grade. *Early Childhood Research Quarterly*, 45, 155-176. doi: 10.1016/j.ecresq.2018.03.005.

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*, 103-109. doi: 10.3102/0013189X17698700.

Little, R. J. (1988). A test of missing completely at random for multivariate data with missing values. *Journal of the American Statistical Association*, *83*, 1198-1202.

Lo, Y., Mendell, N., & Rubin, D. (2001). Testing the number of components in a normal mixture. *Biometrika*, 88, 767–778. doi: 10.1093/biomet/88.3.767.

Maas, C. J., & Hox, J. J. (2004). Robustness issues in multilevel regression analysis. *Statistica Neerlandica*, *58*, 127-137. doi: 10.1046/j.0039-0402.2003.00252.x.

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*, 757-771. doi: 10.1037/a0019616.

Matthews, J. S., Ponitz, C. C., & Morrison, F. J. (2009). Early gender differences in self-regulation and academic achievement. *Journal of Educational Psychology*, *101*, 689-704. doi: 10.1037/a0014240.

Maydeu-Olivares, A. (2017). Maximum likelihood estimation of structural equation models for continuous data: Standard errors and goodness of fit. *Structural Equation Modeling: A Multidisciplinary Journal*, *24*, 383-394. doi: 10.1080/10705511.2016.1269606.

McFarland, J., Hussar, B., Wang, X., Zhang, J., Wang, K., Rathbun, A., ... & Mann, F. B. (2018). *The condition of education 2018 (NCES 2018-144).* Washington, DC: National Center for Education Statistics, U.S. Department of Education. Retrieved from: https://nces.ed.gov/pubs2018/2018144.pdf.

Miller, M. R., Giesbrecht, G. F., Müller, U., McInerney, R. J., & Kerns, K. A. (2012). A latent variable approach to determining the structure of executive function in preschool children. *Journal of Cognition and Development*, *13*, 395-423. doi: 10.1080/15248372.2011.585478.

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*, 8-14. doi: 10.1177/0963721411429458.

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*, 49-100. doi: 10.1006/cogp.1999.0734.

Moffitt, T. E., Arseneault, L., Belsky, D., Dickson, N., Hancox, R. J., Harrington, H., ... & Sears, M. R. (2011). A gradient of childhood self-control predicts health, wealth, and public safety. *Proceedings of the National Academy of Sciences*, *108*, 2693-2698. doi: 10.1073/pnas.1010076108.

Montroy, J. J., Bowles, R. P., Skibbe, L. E., McClelland, M. M., & Morrison, F. J. (2016). The development of self-regulation across early childhood. *Developmental Psychology*, *52*, 1744-1762. doi: 10.1037/dev0000159.

Morgan, P. L., Farkas, G., Hillemeier, M. M., Pun, W. H., & Maczuga, S. (2018). Kindergarten children's executive functions predict their second-grade academic achievement and behavior. *Child Development*. Advance online publication. doi: 10.1111/cdev.13095.

Muthén, B. O. (2004). Latent variable analysis: Growth mixture modeling and related techniques for longitudinal data. In D. Kaplan (Ed.), *Handbook of quantitative methodology for the social sciences* (pp. 345–368). Newbury Park, CA: Sage.

Nylund, K. L., Asparouhov, T., & Muthén, B. O. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: A monte carlo simulation study. *Structural Equation Modeling-A Multidisciplinary Journal*, 14, 535–569. doi: 10.1080/10705510701575396.

Obradović, J., Portilla, X. A., & Ballard, P. J. (2016). Biological sensitivity to family income: differential effects on early executive functioning. *Child Development*, *87*, 374-384. doi: 10.1111/cdev.12475.

Perry, J. L., Nicholls, A. R., Clough, P. J., & Crust, L. (2015). Assessing model fit: Caveats and recommendations for confirmatory factor analysis and exploratory structural equation modeling. *Measurement in Physical Education and Exercise Science*, *19*, 12-21. doi: 10.1080/1091367X.2014.952370.

Pomerantz, E. M., Altermatt, E. R., & Saxon, J. L. (2002). Making the grade but feeling distressed: gender differences in academic performance and internal distress. *Journal of Educational Psychology*, *94*, 396-404. doi: 10.1037/0022-0663.94.2.396.

Ponitz, C. 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*, 605-619. doi: 10.1037/a0015365.

Raver, C. C., Blair, C., & Willoughby, M. (2013). Poverty as a predictor of 4-year-olds' executive function: New perspectives on models of differential susceptibility. *Developmental Psychology*, *49*, 292-304. doi: 10.1037/a0028343.

Sabol, T. J., & Pianta, R. C. (2012). Patterns of school readiness forecast achievement and socioemotional development at the end of elementary school. *Child Development*, *83*, 282-299. doi: 10.1111/j.1467-8624.2011.01678.x.

Satorra, A., & Bentler, P. M. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika*, *66*, 507-514. doi: 10.2139/ssrn.199064.

Sbarra, D. A., & Pianta, R. C. (2001). Teacher ratings of behavior among African American and Caucasian children during the first two years of school. *Psychology in the Schools*, *38*, 229-238. doi: 10.1002/pits.1013.

Schmitt, S. A., Geldhof, G. J., Purpura, D. J., Duncan, R., & McClelland, M. M. (2017). Examining the relations between executive function, math, and literacy during the transition to kindergarten: A multi-analytic approach. *Journal of Educational Psychology*, *109*, 1120-1140. doi: 10.1037/edu0000193.

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, 6, 461–464.

Shonkoff, J. P. (2011). Protecting brains, not simply stimulating minds. *Science*, *333*, 982-983. doi: 10.1126/science.1206014.

Singer, J. D., Willett, J. B., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. New York, NY: Oxford University Press.

Tabachnick, B. G., Fidell, L. S., & Ullman, J. B. (2007). *Using multivariate statistics* (Vol. 5). Boston, MA: Pearson.

Tenenbaum, H. R., & Ruck, M. D. (2007). Are teachers' expectations different for racial minority than for European American students? A meta-analysis. *Journal of Educational Psychology*, *99*, 253-273. doi: 10.1037/0022-0663.99.2.253.

Thompson, R. A. (2014). Stress and child development. *The Future of Children*, 41-59. doi: 10.1353/foc.2014.0004.

Thorell, L. B., & Wåhlstedt, C. (2006). Executive functioning deficits in relation to symptoms of ADHD and/or ODD in preschool children. *Infant and Child Development*, *15*, 503-518. doi: 10.1002/icd.475.

U.S. Health and Human Services Department. (2016). Annual update of the HHS poverty guidelines. Retrieved from:

https://www.federalregister.gov/documents/2016/01/25/2016-01450/annual-update-of-the-hhs-poverty-guidelines.

Usai, M. C., Viterbori, P., Traverso, L., & De Franchis, V. (2014). Latent structure of executive function in five-and six-year-old children: A longitudinal study. *European Journal of Developmental Psychology*, *11*, 447-462. doi: 10.1080/17405629.2013.840578.

Weber, S. (2010). Bacon: An effective way to detect outliers in multivariate data using Stata (and Mata). *The Stata Journal*, *10*, 331-338. doi: 10.1177/1536867X1001000302.

Widaman, K. F., & Thompson, J. S. (2003). On specifying the null model for incremental fit indices in structural equation modeling. *Psychological Methods*, *8*, 16-37. doi: 10.1037/1082-989X.8.1.16.

Wiebe, S. A., Espy, K. A., & Charak, D. (2008). Using confirmatory factor analysis to understand executive control in preschool children: I. Latent structure. *Developmental Psychology*, *44*, 575-587. doi: 10.1037/0012-1649.44.2.575.

Wiebe, S. A., Sheffield, T. D., & Espy, K. A. (2012). Separating the fish from the sharks: A longitudinal study of preschool response inhibition. *Child Development*, *83*, 1245-1261. doi: 10.1111/j.1467-8624.2012.01765.x.

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

Williford, A. P., Maier, M. F., Downer, J. T., Pianta, R. C., & Howes, C. (2013). Understanding how children's engagement and teachers' interactions combine to predict school readiness. *Journal of Applied Developmental Psychology*, *34*, 299-309. doi: 10.1016/j.appdev.2013.05.002.

Willoughby, M. T., Blair, C. B., Wirth, R. J., Greenberg, M., & the Family Life Project Investigators (2012). The measurement of executive function at age 5: Psychometric properties and relationship to academic achievement. *Psychological Assessment*, *24*, 226-239. doi: 10.1037/a0025361.

Willoughby, M. T., Pek, J., & Blair, C. B., & Family Life Project. (2013). Measuring executive function in early childhood: A focus on maximal reliability and the derivation of short forms. *Psychological Assessment,* 25, 664–670. doi:10.1037/a0031747.

Willoughby, M. T., Wirth, R. J., & Blair, C. B. (2012). Executive function in early childhood: Longitudinal measurement invariance and developmental change. *Psychological Assessment*, *24*, 418-431. doi: 10.1037/a0025779.

von Hippel, P. T. (2007). Regression with missing Ys: An improved strategy for analyzing multiply imputed data. *Sociological Methodology*, *37*, 83-117. doi: 10.1111/j.1467-9531.2007.00180.x.

Yeung, W. J., Linver, M. R., & Brooks–Gunn, J. (2002). How money matters for young children's development: Parental investment and family processes. *Child Development*, *73*, 1861-1879. doi: 10.1111/1467-8624.t01-1-00511.

Zelazo, P. D., Blair, C. B., & Willoughby, M. T. (2016). *Executive function: Implications for education (NCER 2017-2000)*. Washington, DC: National Center for Education Research, Institute of Education Sciences, U.S. Department of Education.

Table 1

Descriptive Statistics for Demographic Characteristics at Preschool Entry

	% / Mean (SD)	Range
Child characteristics		
Age at pre-K entry (months)	52.50 (3.72)	40 - 67
Male	49.74	0 - 100
African American, non-Hispanic	53.93	0 - 100
European American, non-Hispanic	24.39	0 - 100
Hispanic/Latino	9.76	0 - 100
Another race/multiple races/ethnicities	11.92	0 - 100
Deviation (months) from mean classroom age	0 (3.90)	-15.13 - 16.38
Family characteristics		
In poverty (INR < 1)	41.67	0 - 100
N	380	

Note. INR = Income-to-needs ratio.

Table 2

Descriptive Statistics and Missingness for EF Measures Across Time Points

	Mean	(SD)	Range	Skewness	Kurtosis	*N*	% Missing
HTKS							
PK Fall	8.91	13.06	0 - 56	1.48	4.29	373	1.84
PK Winter	16.69	17.23	0 - 60	0.73	2.34	365	3.95
PK Spring	21.53	19.48	0 - 60	0.34	1.75	355	6.58
KG Fall	31.24	19.62	0 - 60	-0.39	1.78	312	17.89
KG Spring	39.42	17.34	0 - 60	-1.01	3.05	154	59.47
Pencil Tap							
PK Fall	0.52	0.35	0 - 1	-0.13	1.54	374	1.58
PK Winter	0.72	0.3	0 - 1	-1.00	2.71	365	3.95
PK Spring	0.76	0.27	0 - 1	-1.39	4.01	355	6.58
KG Fall	0.87	0.21	0 - 1	-2.44	9.11	312	17.89
KG Spring	0.91	0.17	.06 - 1	-3.16	13.94	154	59.47
EF Touch Pig							
PK Fall	0.87	0.14	.05 - 1	-2.65	12.13	301**	20.79
PK Winter	0.92	0.09	.20 - 1	-2.78	17.50	327	13.95
PK Spring	0.93	0.09	.20 - 1	-3.22	20.30	342	10.00
KG Fall	0.95	0.06	.60 - 1	-2.31	9.50	309	18.68
KG Spring	0.96	0.05	.68 - 1	-2.55	10.82	153	59.47
EF Touch PtP							
PK Fall	0.69	0.11	.13 - .97	-1.03	6.10	356	6.32
PK Winter	0.73	0.1	.31 - 1	-0.84	5.10	353	7.11
PK Spring	0.73	0.1	.19 - 1	-0.63	5.00	351	7.63
KG Fall	0.76	0.1	.31 - 1	-0.7	4.57	313	17.63
KG Spring	0.79	0.09	.56 - 1	-0.23	2.50	153	59.47
EF Touch StS							
PK Fall	0.66	0.13	.30 - .96	0.27	2.61	353	7.11
PK Winter	0.72	0.15	.30 - 1	-0.26	2.45	352	7.37
PK Spring	0.77	0.14	.41 - 1	-0.45	2.33	350	7.89
KG Fall	0.83	0.13	.33 - 1	-1.03	3.67	312	17.89
KG Spring	0.85	0.11	.33 - 1	-1.74	6.68	153	59.47
EF Touch Arrows Cn.							
KG Fall	0.81	0.19	.11 - 1	-1.69	5.73	315	17.11
KG Spring	0.87	0.18	0 - 1	-2.72	11.86	154	59.47
EF Touch Arrows Sw.							
KG Fall	0.70	0.30	0 - 1	-1.05	2.84	315	17.11
KG Spring	0.79	0.27	0 - 1	-1.8	5.24	154	59.47

Notes. PK = preschool. KG = kindergarten. HTKS = Head-Toes-Knees-Shoulders. PtP = Pick the Picture. StS = Something's the Same. Cn. = Congruent. Sw. = Switch. Means for HTKS are raw scores. Means for Pencil Tap, EF Touch Pig, Pick the Picture, and Something's the Same subtests represent percentage correct. ** For EF Touch Pig, *N*s are significantly smaller and percentage missing is significantly higher because this subtest has a sample trial. If a child failed the sample, the full subtest was not administered. The increasing *N*s and decreasing percentage missing from fall to spring indicate that fewer children failed the sample trial at each successive data collection window.

Table 3

Confirmatory Factor Analysis Standardized Factor Loadings and Fit Statistics

	One-factor models						Two-factor models				
	T1	T2	T3	T4	T5		T1	T2	T3	T4	T5
Factor loadings						*IC factor loadings*					
HTKS	0.64***	0.69***	0.68***	0.70***	0.82***	HTKS	0.34**	0.28*	0.30*	0.28	0.57**
PT	0.68***	0.62***	0.64***	0.54***	0.55***	PT	0.79***	0.78***	0.73***	0.56***	0.62***
EFT Pig	0.46***	0.44***	0.47***	0.54***	0.49***	EFT Pig	0.47***	0.50***	0.52***	0.56***	0.51**
EFT PTP	0.42***	0.39***	0.51***	0.50***	0.50***	EFT ArrCn	-	-	-	0.30***	0.27**
EFT STS	0.41***	0.62***	0.74***	0.66***	0.57***	EFT ArrSw	-	-	-	0.28**	0.35**
EFT ArrCn	-	-	-	0.28***	0.24**	*WM/SS factor loadings*					
EFT ArrSw	-	-	-	0.27***	0.33**	HTKS	0.33*	0.45***	0.40**	0.427	0.32
						EFT PTP	0.50***	0.41***	0.53***	0.496***	0.56***
						EFT STS	0.50***	0.74***	0.81***	0.688***	0.71***
						IC-WM/SS corr.	0.65***	0.60***	0.78***	0.90***	0.63***
Fit Statistics						*Fit Statistics*					
CFI	0.973	0.945	0.994	0.953	0.947	CFI	1	0.999	1	0.945	0.965
RMSEA	0.054	0.082	0.03	0.054	0.059	RMSEA	0	0.017	0	0.063	0.051
SRMR	0.029	0.045	0.027	0.038	0.048	SRMR	0.011	0.023	0.01	0.036	0.042
$\chi2$	10.49	17.14**	6.63	24.85*	19.89	$\chi2$	1.96	3.30	0.46	24.79**	15.48
(df)	(5)	(5)	(5)	(13)	(13)	(df)	(3)	(3)	(3)	(11)	(11)
MLR scaling correction factor	0.93	1.15	1.27	1.2	1.22	MLR scaling correction factor	0.81	1.14	1.18	1.16	1.1
						$\chi2$ diff. from 1-factor	7.36*	13.69**	5.61	0.75	3.85
						(df)	(2)	(2)	(2)	(2)	(2)
						p	0.025	0.001	0.06	0.69	0.15

Note. * $p< .05$, ** $p< .01$, *** $p<.001$. T = Time. HTKS = Head-Toes-Knees-Shoulders. PT = Pencil Tap. EFT = EF Touch. PTP = Pick the Picture. STS = Something's the Same. Arr Cn = Arrows Congruent. Arr Sw = Arrows Switch. IC = Inhibitory control. WM/SS = Working Memory/Set Shifting. CFI = Comparative fit index. RMSEA = Root mean square error of approximation. SRMR = Standardized root mean square residual. df = degrees of freedom. MLR = Maximum likelihood estimation with robust standard errors. diff = difference.

Table 4

Fit Statistics for Measurement Invariance Across Time, Excluding Pencil Tap

	Configural	Metric	Scalar	Residual
CFI	0.973	0.963	0.942	0.907
RMSEA	0.026	0.03	0.036	0.044
SRMR	0.056	0.091	0.09	0.366
$\chi 2$	249.37**	281.40**	334.16***	413.75***
(df)	(198)	(212)	(226)	(240)
MLR scaling correction factor	0.948	0.96	0.97	1.07
$\chi 2$ diff.	-	29.87*	68.28***	44.17***
(df)		(14)	(14)	(14)
p	-	0.01	< .001	< .001

Note. * $p < .05$, ** $p < .01$, *** $p < .001$. CFI = Comparative fit index. RMSEA = Root mean square error of approximation. SRMR = Standardized root mean square residual. df = degrees of freedom. MLR = Maximum likelihood estimation with robust standard errors. diff = difference.

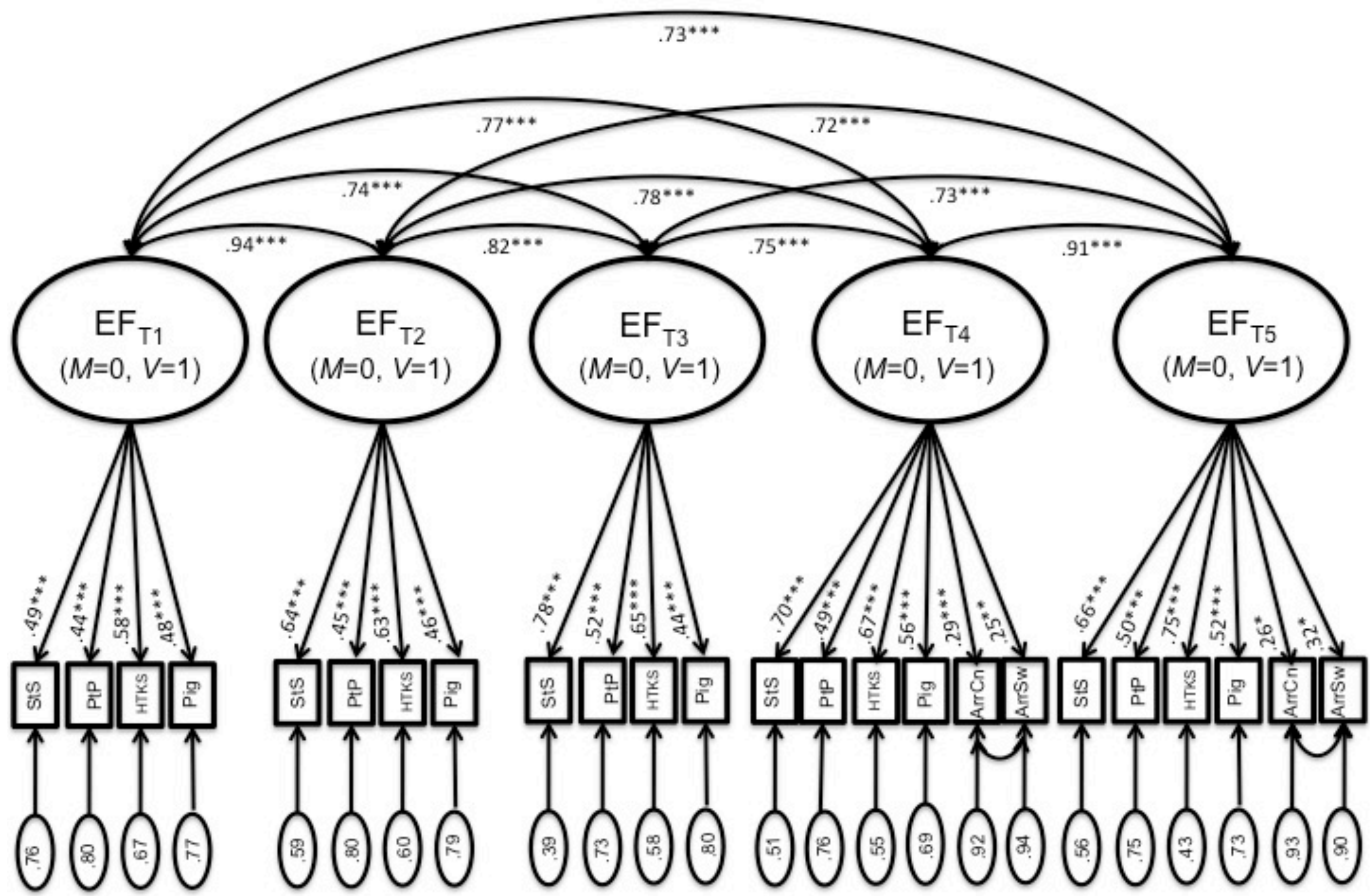

Figure 1. Graphical depiction of results for the confirmatory factor analysis of the unitary executive functioning construct across the five time points, including correlations among latent factors, standardized factor loadings for each subtest at each time point, and residual variance for each subtest at each time point. EF = Executive functioning (latent factor). * $p < .05$, ** $p < .01$, *** $p < .001$. T = Time. M = Mean. V = Variance. StS = EF Touch Something's the Same. PtP = EF Touch Pick the Picture. HTKS = Head-Toes-Knees-Shoulders. Pig = EF Touch Pig. ArrCn = EF Touch Arrows Congruent. ArrSw = EF Touch Arrows Switch.

Table 5

Parameter Estimates and Fit Statistics for Unconditional Latent Growth Curve Models

	Linear		Freeloading	
Parameter Estimates	Estimate	SE	Estimate	SE
Mean (std.)				
Intercept	-0.93***	0.10	-1.12***	0.10
Slope	2.68***	0.33	2.66***	0.29
Variance				
Intercept	0.33***	0.03	0.33***	0.03
Slope	0.13***	< .001	0.32***	0.06
Intercept-slope corr. (std.)	-0.30**	0.10	-0.32***	0.09
R^2				
T1	0.64***	0.05	0.72***	0.05
T2	0.61***	0.04	0.59***	0.03
T3	0.51***	0.04	0.51***	0.03
T4	0.60***	0.05	0.63***	0.05
T5	0.72***	0.09	0.69***	0.07
Fit Statistics				
AIC	3525.692	-	3466.011	-
BIC	3565.094	-	3517.233	-
ABIC	3533.366	-	3475.987	-
RMSEA	0.142	-	0.135	-
SRMR	0.111	-	0.077	-
		-		-
$\chi2$	86.48***	-	56.16***	-
(df)	(10)		(7)	
$\chi2$ diff.		30.32***		
(df)		(3)		
p		< .001		

Note. * $p < .05$, ** $p < .01$, *** $p < .001$. SE = Standard error. Std. = Standardized. Corr = Correlation. T = Time. AIC = Akaike Information Criterion. BIC = Bayesian Information Criterion. ABIC = Adjusted BIC. RMSEA = Root mean square error of approximation. SRMR = Standardized root mean square residual. df = degrees of freedom. diff = difference.

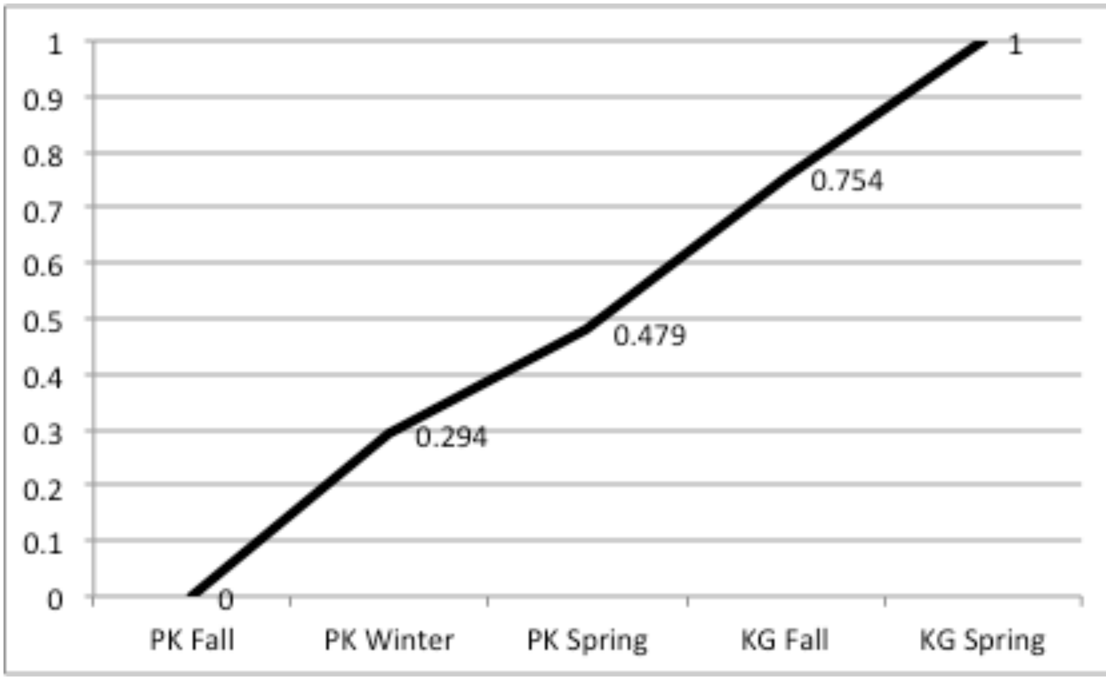

Figure 2. Graphical representation of the unstandardized freeloading latent growth curve model slope estimates across the five time points. PK = preschool. KG = kindergarten.

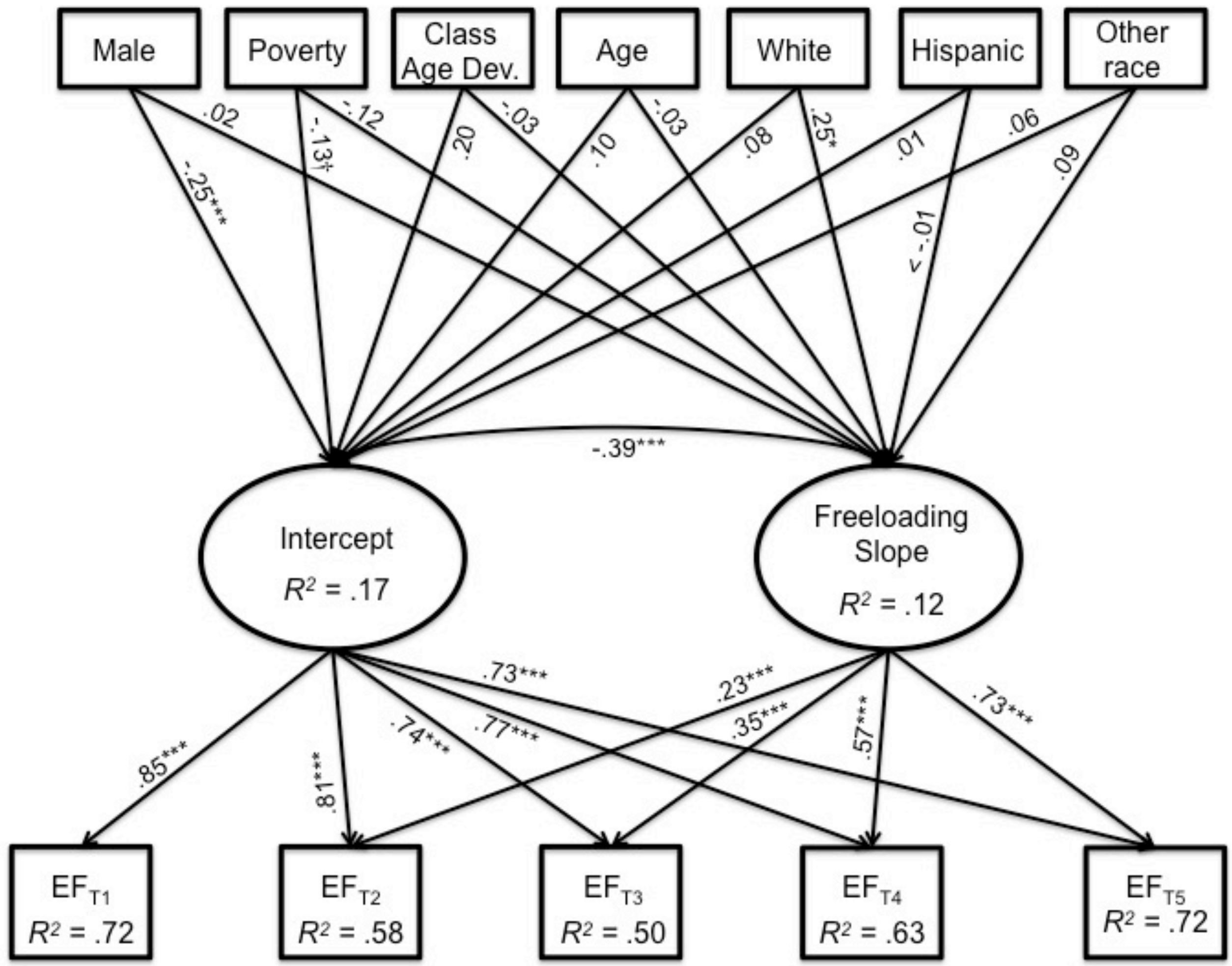

Figure 3. Graphical depiction of results for the freeloading latent growth curve model with demographic characteristics, including standardized estimates for predictors, correlation between the intercept and slope, R^2 for the intercept and slope, standardized factor loadings for the observed EF factor score on the intercept and slope at each time point, and R^2 for each for the EF factor score at each time point. * $p < .05$, ** $p < .01$, *** $p < .001$. Class age dev. = Deviation in months from the mean classroom age. EF = Executive functioning (latent factor). T = Time.

Table 6
Parameter Estimates for Multiple Group Analyses

	Female		Male		Not in poverty		In poverty	
	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE
Mean (std.)								
Intercept	-0.86***	0.12	-1.50***	0.17	-0.94***	0.112	-1.51***	0.20
Slope	2.67***	0.38	2.64***	0.38	2.94***	0.396	2.42***	0.39
Variance								
Intercept	0.35***	0.05	0.27***	0.04	0.40***	0.04	0.23***	0.04
Slope	0.31***	0.08	0.34***	0.09	0.30***	0.07	0.33***	0.09
Intercept-slope corr.	-0.36**	0.13	-0.29*	0.13	-0.52***	0.10	-0.07	0.17
R2								
T1	0.74***	0.06	0.70***	0.08	0.76***	0.05	0.64***	0.08
T2	0.63***	0.05	0.51***	0.05	0.61***	0.04	0.52***	0.05
T3	0.53***	0.05	0.46***	0.05	0.51***	0.05	0.48***	0.06
T4	0.67***	0.07	0.58***	0.07	0.58***	0.06	0.66***	0.07
T5	0.67***	0.08	0.71***	0.09	0.60***	0.10	0.77***	0.09
N	191		189		222		158	

Note. * $p < .05$, ** $p < .01$, *** $p < .001$. Std. = Standardized. Corr = Correlation. T = Time.

Table 7
Multiple Group Analysis Fit Statistics

	Child Sex		
Fit Statistic	Unconstrained	Intercepts Constrained	Slopes Constrained
AIC	3457.575	3472.237	3457.224
BIC	3548.199	3558.92	3543.908
ABIC	3475.225	3489.119	3474.106
RMSEA	0.135	0.147	0.132
SRMR	0.085	0.114	0.086
$\chi 2$	76.84***	92.81***	78.78***
(df)	(17)	(18)	(18)
$\chi 2$ diff.	-	15.97***	1.94
(df)		(1)	(1)
p	-	< .001	0.16

	Poverty Status		
Fit Statistic	Unconstrained	Intercepts Constrained	Slopes Constrained
AIC	3443.883	3445.584	3449.997
BIC	3534.507	3532.268	3536.681
ABIC	3461.532	3462.467	3466.879
RMSEA	0.141	0.139	0.145
SRMR	0.087	0.094	0.092
$\chi 2$	82.36***	85.19***	90.64***
(df)	(17)	(18)	(18)
$\chi 2$ diff.	-	2.83	8.28***
(df)		(1)	(1)
p	-	0.09	0.004

Note. * p< .05, ** p< .01, *** p<.001. AIC = Akaike Information Criterion. BIC = Bayesian Information Criterion. ABIC = Adjusted BIC. RMSEA = Root mean square error of approximation. SRMR = Standardized root mean square residual. df = degrees of freedom. diff = difference.

Appendix

Bivariate Correlations Among Study Variables

		1	2	3	4	5	6	7	8	9	10	11
	EF Variables											
1	HTKS PKF	-										
2	HTKS PKW	0.57***	-									
3	HTKS PKS	0.50***	0.63***	-								
4	HTKS KGF	0.42***	0.50***	0.53***	-							
5	HTKS KGS	0.32***	0.36***	0.39***	0.71***	-						
6	PT PKF	0.44***	0.44***	0.54***	0.43***	0.41***	-					
7	PT PKW	0.31***	0.44***	0.43***	0.45***	0.47***	0.59***	-				
8	PT PKS	0.29***	0.35***	0.45***	0.45***	0.46***	0.54***	0.69***	-			
9	PT KGF	0.21***	0.28***	0.32***	0.41***	0.48***	0.45***	0.59***	0.60***	-		
10	PT KGS	0.20*	0.18*	0.29***	0.33***	0.49***	0.25***	0.35***	0.36***	0.47***	-	
11	EFT Pig PKF	0.25***	0.27***	0.20***	0.27***	0.27**	0.35***	0.31***	0.30***	0.39***	0.30***	-
12	EFT Pig PKW	0.21***	0.24***	0.22***	0.28***	0.42***	0.33***	0.36***	0.38***	0.35***	0.20*	0.37***
13	EFT Pig PKS	0.18***	0.19***	0.29***	0.14*	0.18*	0.30***	0.26***	0.35***	0.27***	0.17*	0.28***
14	EFT Pig KGF	0.19***	0.19***	0.22***	0.32***	0.34***	0.22***	0.22***	0.30***	0.24***	0.24**	0.13*
15	EFT Pig KGS	0.09	0.09	0.17*	0.28***	0.35***	0.34***	0.28***	0.28***	0.35***	0.28***	0.19*
16	EFT PTP PKF	0.26***	0.24***	0.20***	0.27***	0.14	0.27***	0.20***	0.15**	0.10	-0.02	0.16**
17	EFT PTP PKW	0.22***	0.23***	0.18**	0.21***	0.39***	0.25***	0.22***	0.23***	0.20***	0.22**	0.22***
18	EFT PTP PKS	0.25***	0.27***	0.33***	0.19***	0.36***	0.23***	0.15**	0.31***	0.19***	0.11	0.09
19	EFT PTP KGF	0.20***	0.18**	0.23***	0.27***	0.35***	0.34***	0.34***	0.41***	0.34***	0.25**	0.13*
20	EFT PTP KGS	0.12	0.14	0.15***	0.28***	0.35***	0.24**	0.27***	0.21**	0.30***	0.17*	0.19*
21	EFT STS PKF	0.29***	0.31***	0.26***	0.22***	0.23**	0.22***	0.19***	0.14**	0.12*	0.17*	0.13*
22	EFT STS PKW	0.36***	0.46***	0.36***	0.38***	0.32***	0.35***	0.32***	0.36***	0.22***	0.16	0.24***
23	EFT STS PKS	0.37***	0.42***	0.51***	0.34***	0.44***	0.38***	0.31***	0.44***	0.31***	0.22**	0.23***
24	EFT STS KGF	0.35***	0.36***	0.40***	0.49***	0.46***	0.35***	0.28***	0.41***	0.26***	0.17*	0.20**
25	EFT STS KGS	0.23**	0.27***	0.35***	0.44***	0.49***	0.39***	0.36***	0.40***	0.37***	0.21**	0.29**
26	EFT AC KGF	0.02	0.08	0.09	0.15**	0.09	0.14*	0.18**	0.29***	0.15**	-0.03	0.09
27	EFT AC KGS	< .01	0.07	-0.03	0.12	0.20*	-0.01	0.20*	0.24**	0.24**	0.12	-0.08
28	EFT AS KGF	0.09	0.14*	0.11	0.21***	0.19*	0.13*	0.19***	0.24***	0.18**	0.09	0.07
29	EFT AS KGS	0.06	0.06	0.09	0.15	0.27***	0.13	0.22**	0.27***	0.24**	0.20*	<.01
	Demographics											
30	Child is male	-0.13*	-0.10	-0.20***	-0.15**	-0.19*	-0.16**	-0.16**	-0.11*	-0.10	-0.13	-0.15**
31	In poverty	-0.15**	-0.05	<-0.01	-0.16**	-0.18*	-0.17**	-0.13*	-0.12*	-0.21***	-0.02	-0.01
32	Age dev.	0.17**	0.18***	0.14**	0.17**	0.25**	0.25***	0.27***	0.20***	0.16**	0.09	0.12*
33	Child age	0.17***	0.24***	0.13*	0.17**	0.21**	0.22***	0.24***	0.18***	0.13*	0.05	0.13*
34	Afr. Amer.	-0.09	-0.11*	-0.08	-0.08	-0.06	-0.09	-0.08	-0.13*	-0.1	0.13	0.01
35	Eur. Amer.	0.07	0.09	-0.01	0.19***	0.27**	0.04	0.04	0.05	0.11	0.03	0.04
36	Hisp.	-0.01	<-0.01	0.03	-0.05	-0.05	0.04	0.04	0.05	0.04	0.04	-0.02
37	Anoth. race	0.05	0.05	0.10	-0.08	-0.19*	0.04	0.05	0.08	-0.04	-0.24**	-0.05

Note. * $p < .05$, ** $p < .01$, *** $p < .001$. PKF = Preschool Fall. PKW = Preschool Winter. PKS = Preschool Spring. KGF = Kindergarten Fall. KGS = Kindergarten Spring. HTKS = Head-Toes-Knees-Shoulders. EFT = EF Touch. PTP = Pick the Picture. STS = Something's the Same. AC = Arrows Congruent. AS = Arrows Switch. Age dev. = Deviation from mean class age. Afr. Amer. = Child is African American. Eur. Amer. = Child is European American. Hisp. = Child is Hispanic/Latino. Anoth. race = Child is another race/ethnicity.

Bivariate Correlations Among Study Variables, Continued

	12	13	14	15	16	17	18	19	20	21	22	23
EF Variables												
12 EFT Pig PKW	-											
13 EFT Pig PKS	0.37***	-										
14 EFT Pig KGF	0.20***	0.24***	-									
15 EFT Pig KGS	0.12	0.14	0.39***	-								
16 EFT PTP PKF	0.22***	0.03	0.24***	0.09	-							
17 EFT PTP PKW	0.14*	0.13*	0.12*	0.25**	0.29***	-						
18 EFT PTP PKS	0.19***	0.20***	0.23***	0.23**	0.16**	0.31***	-					
19 EFT PTP KGF	0.19**	0.27***	0.27***	0.19*	0.10	0.37***	0.37***	-				
20 EFT PTP KGS	0.13	0.14	0.38***	0.37***	0.14	0.36***	0.35***	0.36***	-			
21 EFT STS PKF	0.15**	0.10	0.14*	0.08	0.24***	0.23***	0.17**	0.15*	0.18*	-		
22 EFT STS PKW	0.19***	0.20***	0.19***	0.15	0.26***	0.29***	0.25***	0.30***	0.18*	0.45***	-	
23 EFT STS PKS	0.30***	0.33***	0.29***	0.25**	0.18***	0.30***	0.41***	0.31***	0.21**	0.38***	0.52***	-
24 EFT STS KGF	0.29***	0.20***	0.35***	0.15	0.13*	0.21***	0.29***	0.33***	0.28***	0.31***	0.47***	0.57***
25 EFT STS KGS	0.42***	0.30***	0.21*	0.21**	0.09	0.30***	0.29***	0.23**	0.39***	0.32***	0.45***	0.61***
26 EFT AC KGF	0.17**	0.21***	0.20***	0.30***	0.06	0.07	0.06	0.12*	0.08	-0.01	0.07	0.17**
27 EFT AC KGS	0.21*	0.10	0.07	0.24**	-0.07	0.02	0.08	<.01	0.05	0.03	-0.05	0.13
28 EFT AS KGF	0.10	0.07	0.16**	0.22**	0.03	0.01	<.01	0.14*	0.09	0.06	0.15**	0.12
29 EFT AS KGS	0.14	0.08	0.20*	0.16*	-0.12	0.16	0.13	0.14	0.11	0.08	0.03	0.20*
Demographics												
30 Child is male	-0.13*	-0.06	-0.08	-0.13	-0.14**	-0.14**	0.01	-0.08	-0.05	-0.19***	-0.05	-0.09
31 In poverty	-0.04	-0.06	-0.13*	-0.16	-0.09	-0.03	-0.12*	-0.03	-0.18*	-0.08	-0.10	-0.17**
32 Age dev.	0.15**	0.15**	0.13*	0.11	0.10	0.12*	0.07	0.20***	0.14	0.17**	0.24***	0.11*
33 Child age	0.15**	0.13*	0.11	0.12	0.11*	0.15**	0.09	0.20***	0.10	0.12*	0.24***	0.09
34 Afr. Amer.	0.05	-0.04	-0.08	-0.09	0.02	-0.08	-0.10	-0.07	-0.09	-0.07	-0.10	-0.17**
35 Eur. Amer.	0.02	-0.08	0.08	0.12	-0.01	0.06	0.01	0.05	0.17*	0.10	0.09	0.08
36 Hisp.	-0.15**	0.09	0.05	0.01	-0.06	0.02	0.04	-0.01	0.02	0.04	<.01	0.07
37 Anoth. race	0.04	0.08	-0.04	-0.03	0.02	0.04	0.10	0.05	-0.10	-0.06	0.04	0.09

Note. * $p< .05$, ** $p< .01$, *** $p<.001$. PKF = Preschool Fall. PKW = Preschool Winter. PKS = Preschool Spring. KGF = Kindergarten Fall. KGS = Kindergarten Spring. EFT = EF Touch. PTP = Pick the Picture. STS = Something's the Same. AC = Arrows Congruent. AS = Arrows Switch. Age dev. = Deviation from mean class age. Afr. Amer. = Child is African American. Eur. Amer. = Child is European American. Hisp. = Child is Hispanic/Latino. Anoth. race = Child is another race/ethnicity.

Bivariate Correlations Among Study Variables, Continued

		24	25	26	27	28	29	30	31	32	33	34	35	36
	EF Variables													
24	EFT STS KGF	-												
25	EFT STS KGS	0.59***	-											
26	EFT AC KGF	0.19***	0.19*	-										
27	EFT AC KGS	0.14	0.09	0.42***	-									
28	EFT AS KGF	0.08	0.13	0.31***	0.10	-								
29	EFT AS KGS	0.14	0.16*	0.20*	0.44***	0.20*	-							
	Demographics													
30	Child is male	-0.07	-0.10	0.04	-0.01	0.04	0.01	-						
31	In poverty	-0.23***	-0.22	-0.15*	-0.1	-0.08	-0.13	-0.11*	-					
32	Age dev.	0.17**	0.14	-0.03	-0.07	-0.02	<.01	-0.03	0.08	-				
33	Child age	0.16**	0.14	-0.03	-0.09	-0.03	-0.02	-0.01	0.09	0.91***	-			
34	Afr. Amer.	-0.18**	-0.19*	-0.02	-0.11	-0.08	-0.17*	-0.12*	0.31***	0.05	0.05	-		
35	Eur. Amer.	0.17**	0.22**	0.07	0.20*	0.09	0.15	0.05	-0.28***	-0.10	-0.04	-0.61***	-	
36	Hisp.	0.01	-0.02	-0.01	-0.22**	<-.01	0.04	0.02	-0.10	0.06	<.01	-0.35***	-0.18***	-
37	Anoth. race	0.04	0.02	-0.05	0.07	<.01	0.04	0.09	-0.01	<.01	-0.03	-0.41***	-0.21***	-0.12*

Note. * $p< .05$, ** $p< .01$, *** $p<.001$. PKF = Preschool Fall. PKW = Preschool Winter. PKS = Preschool Spring. KGF = Kindergarten Fall. KGS = Kindergarten Spring. EFT = EF Touch. STS = Something's the Same. AC = Arrows Congruent. AS = Arrows Switch. Age dev. = Deviation from mean class age. Afr. Amer. = Child is African American. Eur. Amer. = Child is European American. Hisp. = Child is Hispanic/Latino. Anoth. race = Child is another race/ethnicity.