Original Research

The power of play: investigating student success in kindergarten classrooms

Karyn A. Allee

Atlanta Graduate Teacher Education, Tift College of Education, Mercer University, 3001 Mercer University Drive, Atlanta, Georgia 30341, USA

Correspondence to: Karyn A. Allee, allee_ka@mercer.edu

Abstract: Teaching in kindergarten has shifted in recent decades, with the US lagging behind other countries that embrace play as a core pedagogical approach. While global efforts, such as the UN Convention on the Rights of the Child, and national curricula in countries like Canada, Australia, and New Zealand promote play, opinions on its role in early elementary education (K-2) remain divided in the US, and more research is needed to develop effective teaching strategies. This quasi-experimental, pilot study explored the effects of two pedagogical approaches on Title I kindergarten students' executive function (EF), receptive vocabulary, and academic achievement, hypothesizing that purposeful play would particularly benefit students from low-income backgrounds. Results showed that the play-based group made significantly greater reading gains, with links between stronger teacher-reported EF skills and higher academic progress. Although some limitations exist, the findings underscore the potential of play-based pedagogy to enhance children's educational outcomes.

Keywords: Executive function, Kindergarten, Play-based pedagogy, Academic achievement, Title I, Reading, Math

Background

While nations grapple with educational reform efforts globally, these efforts have been particularly pronounced in the US (DeLuca et al., 2020). Educational reforms, notably the Improving America's Schools Act of 1994 and the No Child Left Behind Act of 2002, have aimed to improve educational outcomes for all children by emphasizing standardized testing and accountability. These reforms have fundamentally altered teaching practices across the country from pre-kindergarten through 12th grade education (i.e., the range of publicly funded primary and secondary education in the US and Canada for children aged ~4-18 years). The unintended consequences of these efforts, however, have been a curriculum narrowing through increased content control (i.e., how curriculum responds to high-stakes assessments), pedagogic control (i.e., colloquially referred to as teaching to the test), and formal control (i.e., how high-stakes assessments drive other educational decisions), often with detrimental effects (Au, 2007).

In the US, for example, the disproportionate emphasis on mastering discrete literacy and mathematics learning (i.e., curriculum control) has come at the expense of teaching subjects like science, social studies, art, and music (Dee et al., 2013; Milner et al., 2017). This shift has resulted in more didactic instructional methods (i.e., pedagogical control), characterized by whole-group, teacher-driven, worksheet- and computer-based activities (Allee-Herndon et al., 2022; Repko-Erwin, 2017). These priorities and this narrowing have affected state and local education agencies' choices which further apply pressure to classroom teachers often resulting in what Hatch called accountability shovedown (2002). Kindergarten, which remains non-compulsory for children in 33 states as of 2023 (Education Commission of the States, 2024), is not considered a high-stakes

Received: Jun.10, 2024; Revised: Sep.21, 2024; Accepted: Oct.28, 2024; Published: Nov.01, 2024

DOI: https://doi.org/10.55976/rppe.22024127578-103

testing grade, yet has been impacted by accountability shovedown. While kindergarten students do not generally take standardized assessments that are considered to be high-stakes (i.e., require mandatory grade-level retention for poor performance), they do regularly take progress monitoring and summative standardized tests, and teachers feel the pressure to prepare children for upper grade testing. Consequently, the curriculum in many kindergartens has shifted from a play-based, holistic approach to a more didactic, academically rigorous one, with a primary focus on literacy, mathematics, and test preparation (Allee-Herndon et al., 2022; Repko-Erwin, 2017). Despite these changes, the evidence suggests the prior persistent, predictable achievement gaps (i.e., with student population subgroups such as race/ethnicity, language proficiency, socioeconomic status, or exceptional education services received) have not closed (Rothwell, 2016), or at least have not closed in any meaningful way (Center on Educational Policy, 2007). In essence, children are still being left behind, particularly those attending underfunded and under-resourced schools.

Changed expectations for school readiness and kindergarten achievement outcomes

This observable shift in the educational culture of kindergarten and early learning spaces has also changed perceptions of school readiness and academic success (Allee-Herndon & Roberts, 2019, 2021; Bailey et al., 2019; Bassok et al., 2016; Pyle et al., 2018; Repko-Erwin, 2017). School readiness, typically defined across the domains of child, family, and school (Brown & Lan, 2015; National Association for the Education of Young Children, 2021), refers to the skills, behaviors, and knowledge necessary for formal education. Historically, both families and educators valued children's ability to communicate needs and curiosity (West et al., 1993). However, families previously emphasized skills like sitting still, using tools, counting, and alphabet recognition more than teachers did (West et al., 1993), and legislative changes have since aligned teachers' priorities with those family expectations.

Kindergarten teachers now report (Bassok et al., 2016; Brown & Lan, 2015) prioritizing alphabet knowledge and the ability to hold a pencil (~33%) upon school entry and that children should know how to read upon kindergarten exit (~50%). The focus on academic standards and testing has reduced time for daily art and music lessons (~17%) and generated the reduction or removal of discovery or play centers (~20%) in favor of daily workbook use (~15%) and other types of didactic instruction aligned with assessments and accountability measures (Bassok et al., 2016; Brown & Lan, 2015). This shift reflects a rejection of understanding that school readiness is bi- or tri-directional and has emphasized almost exclusively child-focused readiness, often viewed through a deficit lens, prompting instructional changes that contribute to the academic shovedown (Brown & Lan, 2015; Hatch, 2002; Iruka et al., 2022). Consequently, whole-group instruction has increased, reducing play and children's autonomy in service of preparing for and administering assessments (Allee-Herndon et al., 2022; Bassok et al., 2016; Pyle et al., 2018; Repko-Erwin, 2017).

While direct instruction can help learners with literacy and math (de Bilde et al., 2015; Gersten & Carnine, 1984; Myers & Ankrum, 2018), there is evidence that it is not always effective (Dean & Kuhn, 2007; Taylor & Bilbrey, 2012). Discipline-based instructional frameworks like the National Council for Teaching Mathematics' High Leverage Practices (2014) and Next Generation Science Standards (National Research Council et al., 2007) advocate active, verbal, creative, and discovery-based learning and are more in line with cognitive and social constructivist theories (Piaget, 1977; Vygotsky, 1978) than a didactic instructional approach. However, despite this apparent contradiction, the most vulnerable or marginalized children in the most underfunded and under-resourced schools often receive the least of this engaging instruction (Allee-Herndon et al., 2022; Wood et al., 2022), especially as a result of this schoolification of early learning (Wood et al., 2022). This schoolification epidemic (Ring & O'Sullivan, 2018) and its resultant didactic, narrowed instruction and learning both creates and amplifies equity issues.

Intended and unintended equity impacts of educational improvement efforts

Disparities in school readiness and performance in kindergarten can stem from various sources, with poverty being a significant factor (Allee-Herndon & Roberts, 2019). Economic disparities are already evident in infancy and become even more pronounced at kindergarten age (Burchinal et al., 2011; Halle et al., 2009). Some researchers argue that by kindergarten, the achievement gap is so substantial that it may be insurmountable (Burchinal et al., 2011). Academic and cognitive disparities based on income are notable as early as kindergarten across various subjects (Bailey et al., 2019; Curran, 2017; Gilkerson et al., 2018; Mazzocco & Claessens, 2020). Given the crucial role of successful early childhood experiences in long-term outcomes (Brownell et al., 2015; Schweinhart, 2018), it is crucial to ensure that all children have a positive kindergarten experience.

One key explanation for the income-based kindergarten achievement gap is the impact of adverse childhood experiences, trauma, and chronic toxic stress (e.g., from experiencing extreme poverty and systemic disadvantage or trauma) on brain development (Blair & Raver, 2016; Madrick, 2020; Roos et al., 2019). Chronic stress can trigger a persistent fight-or-flight response and release stress hormones that affect the amygdala, reduce brain size, and delay prefrontal cortex (PFC) development in young children (Agorastos et al., 2019; Allee-Herndon & Roberts, 2019). The PFC is vital for higher-order cognitive functions, emotional regulation, and behavior, all of which are foundational for school readiness (Bailey & Jones, 2019). It encompasses executive function which is responsible for attention, impulse control, planning, goal setting, decision-making, learning, and memory (Bailey et al., 2019; Blair & Raver, 2015; Fitzpatrick et al., 2014), as well as self-regulation (Colliver et al., 2022) and approaches to learning.

PFC developmental delays are closely linked to socioeconomic status (Allee-Herndon & Roberts, 2019; Bailey & Jones, 2019; Blair & Raver, 2015; Fitzpatrick et al., 2014), correlate strongly with school readiness (Blair, 2016; Raver et al., 2011; Vitiello & Greenfield, 2017), and predict academic achievement (Coldren, 2013; Curran, 2017; Gimbert et al., 2019; Meixner et al., 2019; Morgan et al., 2019; Nesbitt et al., 2019; Skibbe et al., 2019). While the mechanisms linking PFC skills and academic success are not fully understood (Ellwood-Lowe et al., 2016), these skills are known to predict success in school, both in kindergarten and over time. Children with PFC developmental delays may seem unprepared for school behaviorally, academically, and socially, but early positive interventions can buffer these children and mitigate the effects of adversity (Shonkoff, 2011).

Effective interventions promote brain growth and reduce stress through predictable routines, responsive relationships, a sense of safety and agency, and skill practice with support (Allee-Herndon & Roberts, 2019). Play and playful learning environments can create these conditions. Movement during play helps build neural pathways and consolidate new learning (Egger et al., 2019), while language practice and vocabulary development occur through sharing, negotiating, and collaborating (Allee-Herndon et al., 2022). Play also enhances executive function skills such as working memory, cognitive flexibility, and inhibitory control through rule-following, planning, problem-solving, and cooperation (Center on the Developing Child, 2017; Moreno et al., 2017). Additionally, play fosters cognitive skills and academic concepts through discovery, inquiry, experimentation, and application (McDonald, 2018; Mraz et al., 2016; Riek, 2014).

Play aligns with constructivist learning theories (Piaget, 1977; Vygotsky, 1978) and developmentally appropriate practice (National Association for the Education of Young Children, 2021), supporting high-quality pedagogy aimed at enhancing PFC development (Shonkoff, 2011). Recognized as foundational for children's development across various domains of well-being and growth (Nesbitt et al., 2023), play can be understood along a continuum. Expanding on Pyle and Danniels' (2017) continuum of play-based learning—which ranges from free play to learning through games—Zosh et al. (2018) describe playful learning as being tied to explicit learning goals and initiated and/or directed by either children or adults. Hirsh-Pasek and colleagues have coined the term active, playful learning to articulate how learning can occur through play, with or without adult facilitation, and with varying levels of structure (Nesbitt et al., 2023, Active Playful Learning section).

This active, play-based approach, which is more likely in our current climate to incorporate purposeful or guided play (Allee-Herndon & Roberts, 2021) than free play or the completely immersive thematic play from 30 years ago, is contrary to the current US contemporary approach which requires extensive "sitting and getting." Active, playful learning, or guided or purposeful play, is designed and facilitated through the lens of learning objectives aligned with academic standards without sacrificing children's agency and interests and has been shown to generate more positive outcomes across developmental domains than direct instruction (Nesbitt et al., 2023). Children with developmental delays in PFC can appear to be the children most in need of growth opportunities such as those playful learning can provide. However, in an effort to "catch them up" to their peers, they are often asked to do the most rigid, sedentary, teacher-directed learning which evidence suggests is less effective in achieving these intended goals (although some evidence suggests this is not true, i.e., Chiatovich & Stipek, 2016). Children who struggle to remember directions, control their bodies or voices, have difficulty resolving social conflict, struggle to persist in the face of academic struggle, or display externalizing behaviors—which are all potential indicators of PFC delays—are the least prepared to focus on worksheets, computer-based instruction, and direct instruction but are expected to do this the most (Allee et al., 2023; Allee-Herndon et al., 2022). This lack of learner and learning task alignment likely contributes to the increase in externalizing behaviors and exclusionary discipline which disproportionately impacts historically marginalized and underserved learners and further removes them from the learning happening in the classroom (Allee-Herndon et al., 2019; Lee & Bierman, 2016; Razza et al., 2015; Skiba et al., 2011). Evidence suggests school absenteeism, which further removes students from learning, is also related to external factors such as poverty (Ansari & Gottfried, 2020) adding confounding factors to their disadvantage. Children who are "behind" may need authentic, engaging, developmentally appropriate, playful learning experiences the most, but shifting the US educational culture will not be easy.

The via media as a potential solution

In contrast to most of the US, other countries are integrating play with assessment, prioritizing a balanced pedagogical approach, especially in schools serving vulnerable learners such as US Title I schools with diverse racial, ethnic, linguistic, or cultural backgrounds and schools in densely urban areas (Allee et al., 2023; Allee-Herndon et al., 2022; DeLuca et al., 2020). A significant challenge in the US is the perceived dichotomy between focusing on rigorous standards and incorporating play (Bassok et al., 2016; Ranz-Smith, 2007; Repko-Erwin, 2017). This perception has increased pressure and tension in early childhood classrooms (Dealey & Stone, 2018; Nitecki & Chung, 2013; Pyle et al., 2018) as educators contend with accountability measures. Structured, didactic classrooms are often viewed as being at one end of the pedagogical spectrum, opposite to play-based learning and child-directed play (Allee-Herndon et al., 2019; Pyle & Danniels, 2017; Repko-Erwin, 2017), implying a necessary choice between the two.

However, a few US states (i.e., Connecticut, Oklahoma, and New Hampshire) are beginning to consider or have already passed legislation to support playful learning, although these efforts are still in the minority (Blinkoff et al., 2023). Rather than trying to convince policymakers to abandon the notion that play and academic rigor are mutually exclusive, it may be more effective to propose a via media approach—similar to international trends and supported by families, policy advisors, and researchers (Brown et al., 2019; Parker et al., 2022). The via media, or "middle road," emphasizes moderation and balance, integrating play into the curriculum without sacrificing academic standards. Active, playful learning (Nesbitt et al., 2023), which incorporates purposeful or guided play and a student-centered focus, represents this balanced approach where teachers create environments that scaffold and support children's learning aligned with specific goals (Allee-Herndon & Roberts, 2021; Mraz et al., 2016; Weisberg et al., 2016). Unlike free play, which allows complete child autonomy and is further removed on the continuum from didactic instruction, guided play provides structured opportunities for learning while also incorporating student agency (Allee-Herndon et al., 2019; Pyle & Danniels, 2017; Repko-Erwin, 2017; Stockard et al., 2018).

To shift towards more developmentally appropriate, stress-reducing, and child-friendly environments, which evidence suggests can support PFC development (Shonkoff, 2011) and other positive child development outcomes, we need more evidence that play supports learning and to empower teachers to apply these findings (Bishop et al., 2020; Koslouski & Stark, 2021). Combining play-based pedagogy with organized and intentional direct instruction, planned through a developmentally appropriate lens, is likely to yield better outcomes than current practices (Allee-Herndon et al., 2022; Hu et al., 2015). Adding to the growing evidence supporting playful learning has the potential to justify instructional shifts back toward this direction in the US and elsewhere. Given the lack of desired results (Center on Educational Policy, 2007; Rothwell, 2016) from the past 30 years of narrow, academically intense instruction (Au, 2007), a balanced approach combining standards-driven instruction with active, playful learning is a promising both/and starting point.

Present study

The current research aims to gather evidence as part of a larger study conducted in Central Florida. A Title I school was intentionally selected to examine the effects of different pedagogical approaches on kindergarten student outcomes, given that Title I schools serve higher percentages of economically disadvantaged students (US Department of Education, 2018). The participating educators were also chosen purposefully (Patton, 2002) to investigate the impacts of incorporating play-based learning versus a traditional didactic approach. In contrast to countries such as Canada, Australia, New Zealand, and many Northern European nations, the US kindergarten curriculum largely avoids play, viewing it as counterproductive to rigorous, academically-focused learning and accountability (Bassok et al., 2016; Ranz-Smith, 2007; Repko-Erwin, 2017). Play-based learning, also known as guided or purposeful play, is often described as offering children freedom of choice, discovery, and exploration within an adult-facilitated structure (Allee-Herndon & Roberts, 2021). Hirsh-Pasek and colleagues have expanded on this concept to conceptualize a framework for active, playful using a three-part equation that adds cultural values to the science of how children learn with critical components of what children should learn (Nesbitt et al., 2023). International readers might find their definition of play-based learning more similar to the US kindergartens of over 30 years ago, which included elements like housekeeping centers, sand and water tables, and various art centers. In this study, "play-based" is used as a shorthand to refer to classrooms in which purposeful or guided play alongside district- or state-mandated curricula to help students achieve learning goals aligned with the active, playful learning mindset. This paper also uses the term "contemporary classroom" to describe the predominantly didactic, teacher-directed environment found in schoolified US kindergartens.

The primary aim of this research is to conduct preliminary investigations into the potential benefits of playful learning as an instructional method to mitigate the negative effects of income insecurity on academic achievement in kindergarten. The study focuses on examining how small, play-based pedagogical shifts might influence student outcomes, guided by the following research questions:

1.To what extent do pedagogical differences in a contemporary classroom and a classroom prioritizing active, playful learning influence executive function, vocabulary, and reading and math academic achievement among Title I kindergarten students?

2.Are there relationships between posttest measures of vocabulary and reading and math academic achievement and teacher posttest measures of students' executive function among kindergarteners in a Title I school?

Emerging evidence (Allee-Herndon & Roberts, 2019; Hirsh-Pasek et al., 2022; Nesbitt et al., 2023) shows play, including playful learning, contributes to positive cognitive and physical development, social and emotional well-being, and academic skill development (Nesbitt et al., 2023). It was hypothesized that, even without a specific experimental condition, children in the play-based classroom would outperform their peers in the contemporary classroom on measures of academic outcomes (i.e., receptive vocabulary, reading and math achievement scores) and executive function health. It was also hypothesized, based on prior research evidence (Allee et al., 2023; Bailey & Jones, 2019, Blair & Raver, 2014; Gimbert et al., 2019; Meixner et al., 2019; Morgan et al., 2017, 2019; Nesbitt et al., 2019; Skibbe et al., 2019), that there would be relationships between children's academic outcomes and executive function. Specifically, the hypothesis was that greater executive function wellness or health would be positively correlated with children's strong academic performance.

Method

Participants

This naturalistic, quasi-experimental study received approval from the university Institutional Review Board and the local school and district administrators. The participants were 30 kindergarten students purposively recruited (Patton, 2002) from a Title I elementary school in Central Florida. Title I status is based on the percentage of students eligible for Free or Reduced-Price Lunch (FRPL; USDOE, 2018), which is frequently used as a surrogate variable for economic insecurity. Teacher A, the play-based instructor, volunteered after recruiting on social media, while Teacher B, representing the contemporary classroom, was chosen by the principal for their distinctly different pedagogical approach. Pre-kindergarten literacy skills (e.g., phonological awareness, letter recognition) were assessed in the summer before kindergarten entry, and the 20 students with the best results were assigned to the play-based classroom. The remaining students were distributed among the five other classrooms, including Teacher B's contemporary classroom. Presumably, the school was interested in providing "enrichment" for more "advanced", ability-grouped students. One could also assume that students who were considered to be less "academically at-risk" could afford to play more and/or Teacher A had a long-standing history at this school and was able to "get away" with teaching differently than the other teachers. Regardless, this school-based decision, which had been the practice for multiple years, presented a complication for the study. As such, statistical adjustments accounted for the initial differences in literacy scores to compare growth rather than raw scores which is discussed in greater detail in the Results and Discussion sections.

WAfter securing informed consent from educators, parents and family caregivers were recruited during Curriculum Night and Open House presentations as well as through classroom newsletters. Parents were informed that the study aimed to analyze instructional approaches without revealing the specific research hypotheses, and the teachers were kept blind to the research hypotheses, too, only knowing play was a variable of interest in an investigation of different instructional approaches. Children were eligible if they were in one of the two classrooms, had parental consent, and gave verbal assent to participate. Out of 39 potential participants, 31 students were included: 19 from the play-based classroom (68% FRPL) and 12 from the contemporary classroom (67% FRPL). Despite the unequal sample sizes, the students were demographically similar (Table 1).

Table 1. Participant demographics by condition

	Play-Based Kindergarten n = 19		**Contemporary Kindergarten n = 12**	
Gender	Female = 11 (57.9%)	Male = 8 (42.1%)	Female = 7 (58.3%)	Female = 7 (58.3%)
Race/Ethnicity	Asian = 1 (5.3%)		Asian = 0 (0%)	
	Hispanic = 5 (26.3%)		Hispanic = 2 (16.7%)	
	White = 10 (52.6%)		White = 9 (75.0%)	
	Black = 3 (15.8%)		Black = 1 (8.3%)	
	Yes	No	Yes	No
ESE	0 (0%)	19 (100%)	0 (0%)	12 (100%)
Gifted/ Talented	0 (0%)	19 (100%)	0 (0%)	12 (100%)
504 Plan	0 (0%)	19 (100%)	0 (0%)	12 (100%)
EL	1 (5.3%)	18 (94.7%)	1 (8.3%)	11 (91.7%)
FRPL	13 (68.4%)	6 (31.6%)	8 (66.7%)	4 (33.3%)
Age at Pretest	M = 5.64 years Range = 5.11 – 6.6 years		M = 5.52 years Range = 5.10 – 5.82 years	

Note. ESE = Exceptional Student Education. 504 Plan = Plans schools put in place to support students with disabilities by removing barriers and providing accommodations. EL = English Learners. FRPL = Free or Reduced-Price Lunch, which is often used as a proxy for students' socioeconomic status.

Classroom conditions

Both classrooms adhered to state kindergarten academic standards and used district-adopted curricula and assessments (Allee-Herndon et al., 2022). However, the instructional environments differed significantly. Regular classroom visits were conducted to observe instruction at various times of the day and days of the week using the School-Age Care Environment Rating Scale, Updated Edition (SACERS-U; Harms et al., 2013) and field notes for data collection. Formal analysis of that data is not included in this paper, but casual observation of the two classrooms paints two very distinct classroom spaces, and a summary of similarities and differences are represented in Figure 1. While the physical environments, among other factors like teacher personality, may or may not have contributed to differential student outcomes, a brief description is included below for context.

Teacher A used an active, playful learning approach in the classroom with flexible seating, colorful wall resources and anchor charts, and a variety of learning space configurations, from whole group instruction to centers and individual workstations. Student work was displayed on the walls, music and movement were part of the morning circle, and books and materials were easily accessible to the children. Teacher A incorporated 30 minutes of guided play learning stations (e.g., writing, literacy games, math games aligned to specific learning goals) and 30 minutes of free-play centers (e.g., housekeeping, blocks, puppets, games, art) each day. Students regularly engaged in choosing how and what to play, emphasizing a spirit of collaboration, creative innovation, critical thinking, confidence building, and content-knowledge development, what Golinkoff and Hirsh-Pasek (2016) coined the "6Cs" of what children learn through play-based experiences. Teacher B in the contemporary classroom had a more austere classroom space with very little adorning the walls, very little color in the classroom, a rug area with assigned seats, and desks arranged for whole group instruction or isolated learning as a behavior management tool. While the student desks were periodically rearranged, they were always configured in whole group structures. There were wall-based resources present, but they were fewer, and Teacher B was not observed referencing them often. Both teachers had 30 minutes of outdoor recess daily after lunch.

Data collection procedures

Three instruments were used in pre- and posttests to measure students' language, academic, and executive function outcomes. The Peabody Picture Vocabulary Test 4th Edition (PPVT-4; Dunn & Dunn, 2007) assessed receptive vocabulary, administered individually in a quiet room. The Behavior Rating Inventory of Executive Function 2nd Edition (BRIEF2; Gioia et al., 2015) collected data on executive function from teachers via surveys the teachers completed in September (beginning) and May (end of the school year). Completing the teacher survey took teachers 10-15 minutes per student, and they were provided classroom coverage and release time by their principal to enter the data so they did not need to complete the surveys on their own time. The i-Ready Reading and Math Diagnostic Assessments (Curriculum Associates, n.d.), required by the district measured academic achievement. An overview of the data collection timeline is presented in Figure 2 (Allee et al., 2023 and Allee-Herndon et al., 2022 for more procedural detail).

Peabody Picture Vocabulary Test 4th Edition. The PPVT-4 (Dunn & Dunn, 2007) measures receptive vocabulary and has been used in prior studies as a proxy for evaluating language and cognitive development (Allee-Herndon et al., 2022). The assessor provides a verbal cue (e.g., "Show me 'elbow'"), and the participant selects the corresponding image. The PPVT-4 has two formats; in this study, Format A was used at pretest, and Format B was used at posttest. The PPVT-4 is reliable and valid, with strong convergent validity and test-retest reliability. The PPVT-4 has convergent validity with a variety of other language and cognition instruments including the Expressive Vocabulary Test, Second Edition ($.80 < r < .84$); Clinical Evaluation of Language Fundamentals, Fourth Edition ($.67 < r < .75$); and Group Reading Assessment and Diagnostic Evaluation ($.81 < r < .91$; Dunn & Dunn, 2013). It is also reliable across administration age groups (2.6 – ≥81 years old), has alternate form ($n = 508$, $.87 < r < .93$) and test-retest reliability ($n = 340$, $.92 < r < .96$), and demonstrates consistency with kindergarten-aged children ($.94 < r < .97$; Dunn & Dunn, 2013).

Behavior Rating Inventory of Executive Function 2nd Edition. The BRIEF2 (Gioia et al., 2015) measures executive (dys)function via 63 items across nine clinical subfactors scored within three different indices and as a Global Executive Composite (GEC) score. Lower scores indicate better executive function; scores below 60 are considered to indicate a child's executive function is within normal limits for their age and gender, scores between 60-64, 65-69, and ≥70 indicate mildly, potentially clinically, and clinically elevated concern respectively (Gioia et al., 2015). It is reliable with Teacher Screening Form coefficients between .36 to .80, with strong internal consistency and test-retest reliability ($r = .90$) and T-score stability values showing little change (average T-score change on the three indices and GEC of 2.50 points). Interrater reliability scores between teachers and teachers are moderately stable ($r = .57$) as compared to parents and teachers ($r = .72$) and parents and parents ($r = .71$). The BRIEF2 has strong internal consistency (e.g., GEC Teacher [$r = .98$]), and concurrent validity with the Child Behavior Checklist, the Behavior Assessment System for Children, Second Edition, the Parent Rating Scales, the Conners Third Edition–Parent Short Form, and the ADHD-Rating Scale-IV as cited in Gioia et al. (2015).

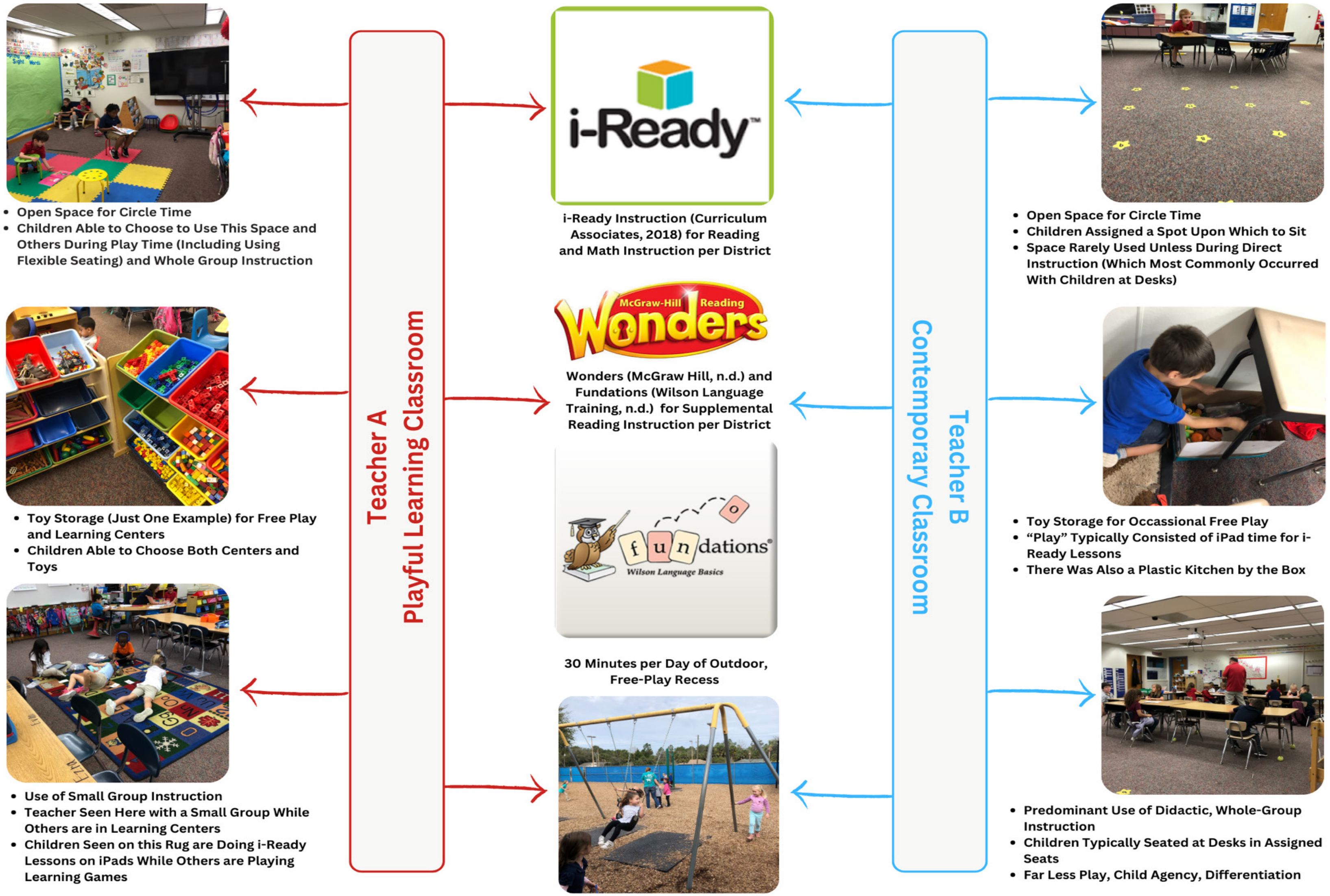

Figure 1. Illustrative overview of sample similarities and differences in classroom condition

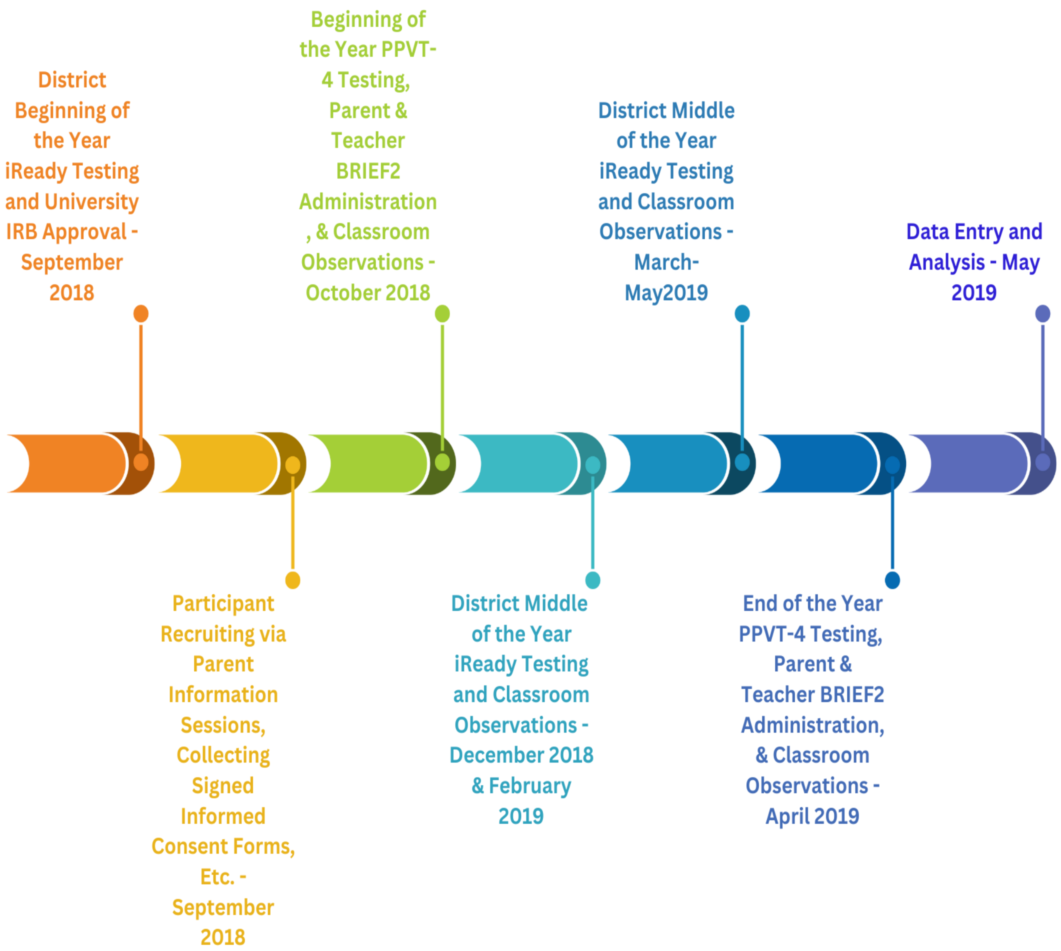

Figure 2. Study data collection timeline

i-Ready Reading and Math Diagnostic Assessment. The study district required the use of i-Ready Diagnostic Assessments for reading and math (Curriculum Associates, n.d.) three times per year for elementary students—beginning, middle, and end—even in grades not required to take high-stakes state testing like kindergarten. The district selected this assessment tool and the corresponding instructional tools because i-Ready is intended for K-12 students and is aligned to state standards, the Every Student Succeeds Act (2015) requirements, and the What Works Clearinghouse (Allee-Herndon et al., 2022). The i-Ready Diagnostic Assessments are adaptive tests, and kindergarten students at this school took them during small group instruction on i-Pads with headphones to listen to questions and prompts, as a way to support pre- and early readers. The reading test assesses phonological awareness, phonics, high-frequency words, vocabulary, and text comprehension, while the math test covers algebra, number operations, geometry, and measurement (Curriculum Associates, 2018). The American Institutes for Research (AIR; 2020) determined the i-Ready Diagnostic Assessment has an acceptable test-retest reliability (n = 120,194, rmedian = .70) and marginal reliability (n = 184,261, r = .91; AIR, 2020). They also found strong correlations to the Florida Standards Assessment (i.e., the state high-stakes assessment for children in 3rd through 10th grades) with i-Ready Reading (n = 291,000, .83 < r < .85) and i-Ready Math (n = 286,000, .87 < r < .88) indicating good predictive validity and reliability for elementary students. AIR also determined there were correlations to 1st grade Lexile scores (n = 840, rmedian = .88), though generalization to kindergarten should be done cautiously.

Design and analysis

This study employed a pretest-posttest, non-equivalent control group design to assess the impact of pedagogical differences on students' receptive vocabulary, executive function, and academic achievement. A difference-in-differences (DiD) approach was used with classroom condition and time as the independent variables, and PPVT-4, Teacher BRIEF2, and i-Ready Reading and Math Diagnostic Assessment scores as the dependent variables. Covariates included age, gender, race/ethnicity, and FRPL status. The analytic sample (n = 28) excluded cases with missing data. Additionally, Spearman's Rank Correlation Coefficient analysis examined the relationships between posttest measures of reading and math achievement and executive function, including all 31 students in this analysis.

Results

The effects of playful learning pedagogy

To assess the impact of classroom condition (play-based vs. contemporary) on developmental outcomes, a series of 26 separate Difference-in-Differences (DiD) regression analyses were conducted, controlling for baseline age, gender, race/ethnicity, and FRPL status. Each model evaluated the main effects of time (pretest vs. posttest), condition (play-based vs. contemporary), and the interaction between time and condition, along with the covariates. The overall fit of the models was assessed using F-statistics, degrees of freedom, and adjusted R^2 values. The results of the regression models are summarized in Table 2. For each outcome, the table presents the unstandardized coefficients (B), standard errors (SE), and p-values for time, condition, and interaction effects. In addition, the table includes the F-statistic and associated p-value, as well as the adjusted R^2 for each model, indicating how much of the variance in each outcome was explained by the predictors.

Main effect for classroom condition. An analysis of the main effect of classroom condition revealed significant differences across several outcomes. The classroom condition had a statistically significant negative effect on Vocabulary scores (B = -103.130, p = .003) and Task Monitoring (B = -17.996, p = .009), indicating that students in different conditions displayed notable differences in their ability to monitor tasks and their vocabulary knowledge. Students in the contemporary classroom had a mean baseline (i.e., beginning of the year or BOY) Vocabulary score that was 21.42 points higher than students in the play-based classroom (MC = 383.700, MPB = 362.280), but at posttest (i.e., end of the year or EOY), the children in the play-based classroom had increased their Vocabulary scores by a mean score of 89.889 whereas children in the contemporary classroom had only increased an average of 9.300 points. Despite the discrepancy at baseline, the rate of Vocabulary change for children in the play-based classroom was greater, and the end results were higher at posttest (MPB = 452.167, MC = 393.00). While students in both classroom conditions increased their mean Task Monitoring score slightly from pretest (MPB = 43.556, MC = 62.300) to posttest (MPB = 44.167, MC = 63.400) with a mean difference score of 0.611 and 1.100 respectively, the more interesting difference is in the scores themselves. For BRIEF2 scores of executive function, unlike PPVT-4 and i-Ready scores, lower scores indicate greater health. Children in the play-based classroom had mean Task Monitoring scores at both time points well within the typical range, while children in the contemporary classroom had mean scores crossing the threshold into mildly elevated levels of concern.

Table 2. Results of the regression models

	Main Effects B (SE)		Interaction Effect B (SE)	Overall Model		
Variables	**Condition**	**Time**	**Condition*Time**	**F(7,48)**	**Adjusted R2**	**Significant Covariates**
Receptive Vocabulary						
1. PPVT-4 Raw Scores	20.836 (11.278)	17.300 (5.715)**	-5.356 (7.128)	5.453***	.362	Race: B = -5.684 (2.760)*
Reading						
2. i-Ready Reading Overall Score†	-27.975 (22.304)	50.800 (11.302)***	50.089 (14.097)***	32.242***	.799	Age: B = 3.818 (1.301)**
3. Phonological Awareness	-2.241 (36.000)	74.600 (18.243)***	22.511 (22.752)	11.147***	.564	
4. Phonics	-25.485 (35.068)	54.500 (17.770)**	66.167 (22.164)**	21.824***	.726	Age: B = 5.692 (2.045)**
5. High Frequency Words	-2.817 (32.113)	67.500 (16.273)***	51.389 (20.296)*	24.849***		
6. Vocabulary	-103.130 (32.335)**	9.300 (16.386)	80.589 (20.437)***	10.135***	.538	Age: B = 4.961 (1.886)**
7. Comprehension: Literary Text	-12.889 (41.098)	49.200 (20.826)*	36.133 (25.975)	7.110***	.437	
8. Comprehension: Informational Text	-31.310 (38.175)	43.600 (19.345)*	46.289 (24.127)	8.868***		
Math						
9. i-Ready Math Overall Score†	1.594 (13.484)	34.600 (6.833)***	12.789 (8.522)	21.261***	.721	Age: B = 2.329 (.786)**
10. Number Sense and Operations	26.578 (16.080	45.400 (8.148)***	-4.900 (10.163)	13.783***	.619	
11. Algebra and Algebraic Thinking	-10.253 (21.331)	31.000 (10.809)**	26.111 (13.481)	11.658***	.576	
12. Measurement and Data	-5.122 (19.382)	27.600 (9.822)**	17.844 (12.250)	10.121***	.537	Age: B = 2.962 (1.130)*
13. Geometry	-9.178 (17.229)	33.200 (8.731)***	12.022 (10.889)	11.339***	.568	
Executive Function						
14. Global Executive Composite‡	-9.1117 (6.983)	5.500 (3.539)	-6.389 (4.413)	11.633***	.575	
15. Behavior Regulation Index†	-8.120 (9.711)	5.300 (4.921)	-5.967 (6.137)	5.407***	.359	
16. Inhibit	-2.593 (10.421)	6.300 (5.281)	-7.078 (6.587)	3.503**	.242	
17. Self-Monitor	-15.028 (7.655)	2.600 (3.879)	-2.822 (4.838)	9.614***	.523	
18. Emotional Regulation Index†	-2.013 (9.716)	7.000 (4.924)	-10.000 (6.141)	6.028***	.390	FRPL: B = -3.891 (1.642)*
19. Shift	-11.789 (6.567)	3.200 (3.328)	-5.422 (4.150)	14.600***	.634	

	Main Effects B (SE)		Interaction Effect B (SE)	Overall Model		
20. Emotional Control	9.480 (12.218)	8.800 (6.191)	-12.300 (7.722)	2.231*	.135	
21. Cognitive Regulation Index†	-11.988 (5.730)	3.300 (2.904)	-3.133 (3.622)	12.994***	.604	
22. Initiate	-4.522 (6.433)	2.700 (3.260)	-3.756 (4.066)	4.207***	.290	FRPL: B = -.003 (1.087)**
23. Working Memory	-11.108 (6.169)	2.700 (3.126)	-2.033 (3.899)	8.163***	.477	
24. Planning and Organization	-8.572 (4.706)	5.100 (2.385)*	-4.711 (2.974)	16.978***	.670	
25. Task Monitoring	-17.996 (6.648)**	1.100 (3.369	-.489 (4.202)	13.158***	.607	Gender: B = 4.648 (2.166)*
26. Organizing Materials	-4.706 (5.232)	5.100 (2.651)	-6.378 (3.307)	11.760***	.578	

Note. † = Overall Score (with sub-scores underneath). ‡ = Overall Composite Score (with sub-scores underneath, including relevant Overall Scores). * $p \leq .05$. ** $p \leq .01$. *** $p < .001$. B = unstandardized regression coefficient. SE = standard error. F-statistic = ratio of variance. Adjusted R2 = coefficient of determination adjusted for the number of predictors in the model and the sample size

Additionally, condition approached significance in predicting Planning and Organization abilities (B = -8.572, p = .075), suggesting that students in different conditions may also have differed in how well they could plan and organize their work, though this effect did not reach the threshold for statistical significance. Once again, scores in both conditions increased on average from baseline (MPB = 41.722, MC = 55.100) to the end of the year (MPB = 42.111, MC = 60.200), but followed similar patterns as Task Monitoring with the children in the contemporary classroom just crossing the threshold into mildly elevated concern. While the condition variable did not have a significant main effect on Organizing Materials (B = -4.706, p = .373) or Comprehension of Informational Text (B = -31.310, p = .416), negative trends were observed in both cases, implying that students in different conditions may have exhibited lower performance in these areas, though the differences were not statistically conclusive. Children in the play-based classroom improved (i.e., decreased) their mean Organizing Materials scores from pre- (MPB = 43.500) to posttest (MPB = 42.222), but children in the contemporary classroom had mean scores with increased levels of concern during the same timeframe (MC = 54.600 and 59.700 respectively). Again, while not a statistically significant difference, children in the play-based classroom had higher Comprehension of Informational Text mean scores at pretest (MPB = 389.389, MC = 373.700) and posttest (MPB = 479.278, MC = 417.300) with a greater mean change over time (MPB = 89.889, MC = 43.600). Overall, the findings suggest that the educational or environmental condition significantly impacted certain cognitive and academic skills, particularly task monitoring and vocabulary, with trends also observed for planning and organizational skills.

Main effect for time. The analysis revealed that time had a significant positive effect on multiple outcomes across academic and executive function domains. Time had a significant positive impact on Receptive Vocabulary, all reading scores except the Vocabulary sub score (B = 9.300, p = .573), and all i-Ready math scores. The only BRIEF2 executive function score significantly impacted by time was Planning and Organization (B = 5.100, p = .038). These results suggest that students developed stronger academic skills over time and demonstrated substantial improvements in these areas, but there was a negative trend in the Planning and Organization outcomes (i.e., higher scores are not desirable).

While the effect of time on Emotional Control (B = 8.800, p = .162) and Inhibitory Control (B = 6.300, p = .239) was not statistically significant, there were notable, albeit concerning, trends in these areas as well. Children in the play-based class showed improvement on average from pretest (MPB = 53.444) to posttest (MPB = 49.944) with a mean difference score of -3.500 (remember lower scores indicate greater health) where the mean difference score for children in the contemporary class increased by 8.800, putting the average posttest score in the potentially clinically elevated range (MC = 57.100 to 65.900 from baseline to the end of the year). Overall, the results suggest that time had a robust effect on students' academic performance, as might be expected, as well as on key aspects of executive functioning, although not always in a positive manner.

Interaction effect for condition*time. Several interaction effects between condition and time were identified across reading domains. The effect of time on Overall Reading scores differed between the play-based and contemporary classrooms (B = 50.800, p < .001). Specifically, students in the play-based classroom improved more over time compared to students in the contemporary classroom with mean difference scores of MPB = 100.889 and MC = 50.800 respectively. There were similar outcomes for High Frequency Words (B = 51.389, p = .015, MPB = 118.889 and MC = 67.500) and Vocabulary (B = 80.589, p < .001, MPB = 89.889 and MC = 9.300) despite the previously mentioned higher BOY Vocabulary scores for children in the contemporary classroom. The interaction between condition and time approaches significance for Comprehension of Informational Text (B = 46.289, p = .061, MPB = 89.889 and MC = 43.600), as does Algebraic Thinking B = 26.111, p = .059, MPB = 102.800 and MC = 31.000). This suggests that these scores may differ in meaningful ways over time between the play-based and contemporary classrooms, though this finding is not significant at the conventional α = .05 level. These are key findings for the DiD analysis, indicating that the play-based condition had a stronger impact on growth over time across multiple variables, while the significant positive interaction coefficient suggests that students in the play-based classroom improved more from pretest to posttest compared to students in the contemporary classroom.

Covariates. The covariate analyses revealed several significant relationships between the covariates (age, race, gender, and FRPL status) and various outcomes across the domains of reading, math, and executive function. Age was a significant covariate for multiple outcomes where each statistically significant unstandardized coefficient (B) value suggests a predictive improvement of that many points for every monthly increase in age. For example, Overall Reading scores (B = 3.818, p = .002) increased almost four points for every month older a child was. Phonics (B = 5.692, p = .008), Vocabulary (B = 4.961, p = .011), Overall Math (B = 2.329, p = .005), Measurement and Data (B = 2.962, p = .012), and Geometry (B = 3.491, p = .001) scores all indicate that older students tended to perform better in these areas. Age was also approaching significance suggesting a potential positive effect of age on Comprehension of Informational Text (B = 3.491, p = .094) and Number and Operation (B = 1.716, p = .074), scores, though these results are not significant at the traditional α = .05 level. It is important to note that children in the play-based classroom were slightly older at BOY (MPB = 66.67 months) than children in the contemporary classroom (Mc = 66.67 months), but there were no statistically significant differences in age between the two classrooms, t(52.855) = -.519, p = .606. See Table 3 for a complete reporting of mean scores at both time points across variables disaggregated by covariate.

Gender was significant for Task Monitoring (B = 4.648, p = .037), with male students demonstrating better task monitoring skills (i.e., lower scores; BOY MBoy = 47.36, MGirl = 52.12, EOY MBoy = 47.64, MGirl = 53.24). Gender approaches significance for both Inhibit (B = 5.978, p = .085, BOY MBoy = 47.00, MGirl = 52.18, EOY MBoy = 47.91, MGirl = 54.47) and Behavior Regulation Index (B = 5.305, p = .100, BOY MBoy = 48.09, MGirl = 52.47, EOY MBoy = 48.73, MGirl = 54.47) scores, and while gender is not a significant predictor, the positive Self-Monitor coefficient (B = 3.450, p = .173) suggests a potential trend where gender may be associated with higher scores (BOY MBoy = 49.82, MGirl = 52.12, EOY MBoy = 49.73, MGirl = 53.47), but this requires further exploration. Of these four executive function variables, three can be grouped together as Inhibit and Self-Monitoring are the two sub-skills measured by the categorical Behavior Regulation Index value whereas Task Monitoring falls under the Cognitive Regulation Index category. Again, there are no statistically different proportions of male and female students by classroom condition as determined by a chi-square test for homogeneity, p = .935, nor were there any statistically different proportions of racial or ethnic student compositions by classroom condition, p = .129.

Race/Ethnicity was significant for Receptive Vocabulary (B = -5.684, p = .045), indicating that race had a meaningful effect on vocabulary performance suggesting children in this sample from certain racial groups scored significantly lower on PPVT-4 raw scores compared to others. Specifically, the mean scores at pre- (MAsian = 356.00, MHispanic = 399.33, MWhite = 367.39, and MBlack = 331.00) and posttest (MAsian = 458.00, MHispanic = 435.33, MWhite = 427.00, and MBlack = 437.67) suggest on average the three Black students had lower mean Vocabulary scores at pretest, but they also had the largest mean change score at 106.67 points. A difference, even one that is statistically significant, based on three students should be interpreted cautiously, and no other variables were significantly predicted by children's race or ethnicity

While there was no statistically significant association between FRPL status and classroom condition as assessed by Fisher's exact test, p = .519, FRPL status, serving as a proxy for students' socio-economic status, also emerged as a significant covariate, particularly for Emotional Regulation (B = -3.891, p = .022, BOY MFRPL = 55.26, EOY MFRPL = 55.32, BOY MNon = 46.89, EOY MNon = 48.56) and Emotional Control (B = -5.042, p = .018, BOY MFRPL = 57.68, EOY MFRPL = 58.26, BOY MNon = 48.56, EOY MNon = 50.11), two related scores, suggesting that students eligible for free or reduced-price lunch tended to have higher scores in these areas (i.e., greater concern). FRPL status, as a covariate, approaches significance (B = -2.102, p = .064, BOY MFRPL = 51.42, EOY MFRPL = 50.37, BOY MNon = 45.33, EOY MNon = 46.67), indicating that students from lower socioeconomic backgrounds (as indicated by FRPL status) may higher lower Shift scores (i.e., greater concern), but this is not statistically confirmed. While the negative coefficient suggests that higher SES (as indicated by FRPL status) might be associated with slightly lower Phonics scores, this effect is not statistically significant (B = -9.553, p = .114, BOY MFRPL = 354.00, EOY MFRPL = 459.05, BOY MNon = 349.89, EOY MNon = 430.00). Overall, these covariate analyses underscore the importance of demographic factors in students' academic and cognitive outcomes, although the interpretation of these results requires thoughtful consideration.

Correlations between outcomes

Due to outliers and non-normal bivariate distributions, a Spearman's Rank Correlation Coefficient test was used to evaluate relationships between posttest scores. Significant correlations were found for 403 of the possible relationships at p ≤ .05, 57 at p ≤ .01, and 145 at p < .001. Moderate effect sizes (.40 < rs < .59) were found in 109 cases, strong effect sizes (.60 < rs < .79) in 96 cases, and very strong effect sizes (.80 < rs < 1.0) in 44 cases. All measures of reading (i.e., i-Ready scores) were at least moderately correlated with each other with statistical significance at p ≤ .05 as were all measures of math (i.e., i-Ready scores). All measures of executive function were similarly correlated with each other. As one might expect like scores to be correlated with like scores, this is not surprising. What is interesting is the extent to which student outcome measures were correlated across academic and executive function scores (see Table 4). For example, receptive vocabulary did not consistently correlate strongly with other academic scores (Comprehension of Literary Text rs = .423, p = .025, Shift rs = -.402, p = .034), but Vocabulary was significantly and at least moderately correlated with both Comprehension variables, Overall Math and Number and Operation, and ten of the Executive Function variables including all of the sub-elements of the Cognitive Regulation Index (-.400 < rs < -.641, < .001 < p < .014). While no single posttest variable was significantly correlated with all other posttest variables with moderate to very strong effect sizes, all were correlated with at least some others, and the extent to which each of these variables is related suggests important connections. For example, lower executive function concerns were associated with higher reading and math outcomes and students who were academically stronger in one area, tended to also be stronger in others.

Table 3. Mean Scores at Pretest and Posttest by Condition and Covariate

	Receptive Vocabulary		i-Ready Reading Overall		Phonological Awareness		Phonic		High-Frequency Words		Vocabulary		Comprehension: Literary Text		Comprehension: Informational Text		i-Ready Math Overall		Number & Operations		Algebraic Thinking		Measurement & Data		Geometry	
	BOY	EOY	BOY	EOY	BOY	EOY	BOY	EOY	BOY	EOY	BOY	EOY	BOY	EOY	BOY	EOY	BOY	EOY	BOY	EOY	BOY	EOY	BOY	EOY	BOY	EOY
Classroom Condition																										
Contemporary (n = 10)	94.20	111.50	350.60	401.40	354.70	429.30	326.10	380.60	314.30	381.80	383.70	393.00	366.50	415.70	373.70	417.30	350.20	384.80	340.50	385.90	345.40	376.40	360.00	387.60	362.30	395.50
Play-Based (n = 18)	109.83	121.78	373.33	474.22	375.89	473.00	367.44	488.11	363.06	481.94	362.28	452.17	389.17	474.50	389.39	479.28	365.44	412.83	362.50	403.00	362.39	419.50	373.33	418.78	366.83	412.06
Age in Months at BOY																										
62 (n = 1)	89.00	102.00	294.00	466.00	292.00	451.00	314.00	498.00	314.00	494.00	276.00	447.00	240.00	481.00	322.00	441.00	316.00	408.00	329.00	410.00	309.00	413.00	314.00	399.00	305.00	406.00
63 (n = 2)	110.50	128.50	372.50	465.50	393.50	502.00	337.50	451.50	350.50	454.00	362.50	439.50	409.00	448.00	398.50	504.00	366.50	400.50	370.50	391.50	365.50	411.00	364.00	403.50	364.00	396.00
64 (n = 4)	101.00	113.25	338.25	453.50	348.50	468.50	307.25	449.50	323.50	445.25	322.75	438.50	379.25	465.00	359.50	455.75	358.75	394.50	358.75	385.00	353.75	397.75	371.50	398.25	356.25	400.75
65 (n = 8)	99.00	120.63	369.75	429.13	375.63	448.13	366.88	424.25	357.25	423.13	381.50	402.13	370.50	449.38	372.88	442.00	354.38	401.75	345.88	397.88	353.25	408.25	359.63	402.38	363.63	399.38
66 (n = 2)	96.00	105.00	356.50	396.50	343.50	406.50	336.00	368.50	289.50	391.00	372.00	402.50	412.00	408.00	399.50	404.00	355.00	381.50	338.50	372.50	354.50	366.50	368.50	414.00	367.00	392.50
67 (n = 1)	125.00	135.00	370.00	409.00	362.00	415.00	287.00	397.00	299.00	383.00	456.00	429.00	440.00	416.00	452.00	412.00	363.00	385.00	356.00	401.00	356.00	358.00	388.00	379.00	357.00	407.00
68 (n = 2)	100.00	117.50	356.00	458.50	337.00	467.00	371.50	443.50	315.00	508.00	354.50	418.00	371.00	470.50	393.50	474.50	363.00	410.50	350.50	388.50	358.00	422.50	375.50	418.00	377.00	417.00
70 (n = 4)	108.50	113.50	373.00	449.25	365.25	452.25	367.00	471.50	371.25	455.00	374.75	438.25	381.25	438.50	394.25	441.50	366.50	404.75	355.75	408.75	371.00	397.25	370.75	401.25	376.25	412.75
71 (n = 4)	115.50	124.00	397.25	497.50	412.75	480.50	388.00	523.00	379.50	481.00	401.50	487.25	400.00	484.75	397.75	514.25	374.00	422.50	375.25	409.75	357.75	424.50	389.00	437.00	377.25	424.75
Gender																										
Male (n = 11)	104.55	120.73	364.00	447.73	374.64	450.55	345.55	456.27	343.55	454.73	376.45	437.91	377.09	450.45	374.91	446.55	360.27	401.00	357.45	394.27	349.73	404.18	370.55	402.45	368.73	406.27
Female (n = 17)	104.06	116.41	366.00	448.53	364.24	461.82	357.29	445.47	347.00	440.65	365.71	426.59	383.65	455.47	389.53	464.00	359.82	404.00	352.82	398.59	360.59	404.06	367.29	411.00	362.94	406.06
Race/Ethnicity																										
Asian (n = 1)	115.00	119.00	390.00	496.00	375.00	437.00	350.00	581.00	379.00	494.00	356.00	458.00	463.00	479.00	447.00	532.00	366.00	411.00	347.00	416.00	365.00	407.00	381.00	408.00	383.00	411.00
Hispanic (n = 6)	111.33	123.83	373.67	449.67	358.83	447.67	352.83	461.67	367.50	459.50	399.33	435.33	398.33	445.50	382.83	445.33	368.33	406.50	363.67	401.50	366.17	407.67	381.83	409.33	367.17	409.67
White (n = 18)	101.94	118.11	364.94	442.39	373.22	459.94	354.94	433.89	340.83	426.44	367.39	427.00	377.94	453.83	384.61	458.72	356.17	401.28	352.00	394.78	350.50	402.33	363.67	407.72	361.94	404.72
Black (n = 3)	100.33	106.33	341.67	464.33	355.67	468.33	339.67	477.00	319.67	522.00	331.00	437.67	338.00	459.00	359.67	446.33	364.33	402.00	355.00	394.00	368.67	406.67	367.33	403.67	375.00	406.00
FRPL Status																										
FRPL Eligible (n = 18)	104.53	115.68	366.21	453.47	366.32	463.21	354.00	459.05	345.58	449.89	367.21	438.89	383.79	455.37	391.16	457.11	359.79	402.79	353.89	399.37	354.26	401.84	371.05	408.74	365.26	404.16
FRPL Non-Eligible (n = 9)	103.67	123.22	363.11	437.11	372.56	445.11	349.89	430.00	345.78	438.33	375.67	414.44	375.33	449.56	368.22	457.22	360.44	402.89	356.22	391.67	360.67	408.89	363.33	405.33	365.11	410.33
Total Mean (n = 28)	104.25	118.11	365.21	448.21	368.32	457.39	352.68	449.71	345.64	446.18	369.93	431.04	381.07	453.50	383.79	457.14	360.00	402.82	354.64	396.89	356.32	404.11	368.57	407.64	365.21	406.14

	BRIEF2 Global executive composite		Behavior regulation index		Inhibit		Self-monitor		Emotion regulation index		shift		Emotional control		Cognitive regulation Index		initiate		Working memory		Planning & organization		Task monitoring		Organizing materials	
	BOY	EOY	BOY	EOY	BOY	EOY	BOY	EOY	BOY	EOY	BOY	EOY	BOY	EOY	BOY	EOY	BOY	EOY	BOY	EOY	BOY	EOY	BOY	EOY	BOY	EOY
Classroom condition																										
Contemporary (n = 10)	60.10	65.60	59.90	65.20	56.50	62.80	62.70	65.30	60.60	67.60	60.60	63.80	57.10	65.90	58.00	61.30	53.90	56.60	57.50	60.20	55.10	60.20	62.30	63.40	54.60	59.70
Play-Based (n = 18)	44.39	43.50	45.67	45.00	46.61	45.83	44.83	44.61	48.11	45.11	43.28	41.06	53.44	49.94	42.78	42.94	45.67	44.61	44.33	45.00	41.72	42.11	43.56	44.17	43.50	42.22
Age in Months at BOY																										
62 (n = 1)	44.00	42.00	44.00	42.00	46.00	43.00	42.00	42.00	54.00	48.00	45.00	41.00	64.00	56.00	41.00	41.00	43.00	43.00	43.00	43.00	41.00	41.00	41.00	41.00	43.00	43.00
63 (n = 2)	46.00	44.00	52.00	46.00	53.50	49.00	49.00	42.00	43.00	44.00	41.00	41.00	46.00	48.00	45.00	43.50	50.50	45.50	48.00	45.50	41.00	41.00	45.50	47.50	45.50	43.00
64 (n = 4)	49.25	50.50	49.00	54.25	48.50	52.25	49.50	56.25	53.75	52.50	46.50	45.25	58.50	59.75	47.50	47.75	47.25	46.50	47.00	49.75	47.50	46.75	50.50	48.25	45.75	48.25
65 (n = 8)	49.25	52.50	47.75	49.13	46.63	48.38	49.25	49.88	48.00	50.13	47.25	50.13	49.12	49.25	50.25	54.00	50.00	52.62	50.62	55.63	47.75	52.00	53.13	56.00	47.25	50.13
66 (n = 2)	61.50	73.50	67.50	81.00	65.50	80.50	67.00	75.50	65.00	83.50	64.50	70.00	62.00	85.00	55.00	61.00	50.50	57.00	54.00	55.50	48.00	62.00	66.00	69.50	58.00	58.00
67 (n = 1)	68.00	77.00	69.00	78.00	71.00	78.00	65.00	72.00	74.00	82.00	67.00	67.00	75.00	88.00	62.00	69.00	54.00	60.00	59.00	67.00	58.00	67.00	67.00	67.00	60.00	68.00
68 (n = 2)	44.00	45.00	45.00	49.50	48.50	51.00	41.50	48.00	41.50	44.00	40.50	41.50	44.00	47.00	45.00	44.00	48.50	49.50	47.50	46.00	41.50	41.50	47.50	43.00	43.50	42.00
70 (n = 4)	58.50	53.00	58.75	53.00	54.25	52.50	62.75	53.25	66.75	57.00	63.50	54.75	66.25	57.00	52.25	49.50	52.00	48.00	54.75	50.25	51.75	49.75	51.25	50.50	51.25	50.00
71 (n = 4)	40.00	40.25	41.50	41.75	42.00	41.25	42.25	43.50	44.75	44.00	40.75	40.75	49.50	48.25	39.00	39.50	42.00	40.75	40.00	40.75	40.75	40.25	37.25	39.50	41.50	41.50
Gender																										
Male (n = 11)	48.36	48.82	48.09	48.73	47.00	47.91	49.82	49.73	52.27	51.73	49.91	47.55	54.82	54.73	46.82	47.73	46.64	47.18	47.55	48.55	46.64	46.55	47.36	47.64	45.27	47.36
Female (n = 17)	51.06	53.06	52.47	54.47	52.18	54.47	52.12	53.47	52.76	54.06	49.18	50.24	54.71	56.24	49.12	50.65	49.88	50.00	50.00	51.65	46.41	49.88	52.12	53.24	48.88	49.18
Race/Ethnicity																										
Asian (n = 1)	44.00	41.00	48.00	42.00	46.00	43.00	51.00	42.00	48.00	42.00	45.00	41.00	52.00	44.00	41.00	41.00	43.00	43.00	43.00	43.00	41.00	41.00	41.00	41.00	43.00	43.00
Hispanic (n = 6)	50.83	48.50	49.67	48.83	48.17	47.67	51.83	50.33	58.83	52.17	53.83	47.33	63.17	55.67	47.33	46.50	48.33	44.33	48.33	46.33	47.33	47.17	47.00	48.17	47.33	47.00
White (n = 18)	51.44	54.44	52.72	55.11	52.06	54.83	52.72	54.22	52.22	55.39	49.83	51.56	53.06	57.22	50.00	52.28	49.67	51.33	50.56	53.39	47.50	50.78	53.22	54.39	48.67	50.50
Black (n = 3)	41.67	42.33	42.00	45.00	44.00	45.67	41.00	45.33	43.67	45.33	40.00	41.33	49.00	50.00	41.67	41.67	44.67	45.33	43.33	43.33	40.67	40.67	42.00	40.00	42.00	41.00
FRPL Status																										
FRPL Eligible (n = 18)	51.16	53.11	51.74	54.68	51.63	54.58	51.37	53.79	55.26	55.32	51.42	50.37	57.68	58.26	48.47	50.37	48.21	49.79	49.11	51.00	46.68	49.63	51.32	52.11	48.05	48.95
FRPL Non-Eligible (n = 9)	47.56	47.78	48.67	47.00	47.00	46.22	50.89	48.22	46.89	48.56	45.33	46.67	48.56	50.11	47.67	47.67	49.44	47.00	48.89	49.22	46.11	46.33	48.00	48.78	46.22	47.44
Total Mean (n = 28)	50.00	51.39	50.75	52.21	50.14	51.89	51.21	52.00	52.57	53.14	49.46	49.18	54.75	55.64	48.21	49.50	48.61	48.89	49.04	50.43	46.50	48.57	50.25	51.04	47.46	48.46

Note. The BRIEF2 evaluates executive function health, and the higher the score, the more increased the degree of executive dysfunction. This is opposite to the academic scores, and scores between 60-64, 65-69, and ≥70 indicate mildly, potentially clinically, and clinically elevated concern (Gioia et al., 2015).

Discussion

The current study seeks to extend prior work (Allee et al., 2023; Allee-Herndon et al., 2022) by exploring the impact of pedagogical differences—specifically, play-based learning versus contemporary, teacher-directed instruction—on the receptive vocabulary, executive functioning, and academic performance of students in a kindergarten setting on a larger number of specific pre- and posttest variables using a DiD approach. The DiD approach was selected as a thoughtful and effective way to address the sample size and to simulate the advantages of an experimental design within this naturalistic context. Given the small sample and non-randomized nature of the group allocations, a traditional experimental design was not feasible. However, the DiD method allows for a robust comparison of changes over time between the two groups—those in the play-based classroom and those in the contemporary classroom—by controlling for pre-existing differences and time-related effects. By focusing on the differences in trends between the groups, the DiD analysis effectively isolates the impact of the pedagogical intervention, minimizing potential biases from confounding variables. This approach helps mitigate concerns about the small sample size and non-random group assignment by leveraging both within-group and between-group comparisons to simulate the control provided by a randomized controlled trial. As a result, the DiD method provides a rigorous means of evaluating the intervention's effects in a real-world educational setting. The findings highlight several important insights regarding the role of pedagogy in early childhood development, especially in the domains of vocabulary, reading, math, and executive function.

A summary of the results

Main effects of time and condition. The analysis showed that both groups made significant progress over time in key academic domains, particularly in reading and math. For example, both groups demonstrated significant improvements in Phonological Awareness and High-Frequency Word recognition, consistent with the expected developmental trajectory for early readers (Ehri, 2005). However, the time effect was more pronounced in the play-based classroom, where students exhibited larger gains, especially in reading comprehension and math operations. These findings align with previous research suggesting that active, playful learning environments foster deeper engagement and skill acquisition, particularly in early literacy and numeracy (Hirsh-Pasek et al., 2009). The play-based classroom also demonstrated significant advantages over the contemporary classroom in vocabulary acquisition. While both groups showed growth, students in the play-based classroom exhibited substantially higher Vocabulary scores by the end of the year. This may be due to the open-ended, exploratory nature of play, which allows for more opportunities to engage with language in meaningful contexts (Weisberg et al., 2013).

Interaction effects between condition and time. Significant interaction effects were found in areas such as Vocabulary and Phonics, with the play-based group showing greater improvements over time compared to the contemporary group. The interaction effect for Vocabulary ($B = 80.589$, $p < .001$) was particularly striking. It indicates that both groups began with relatively similar scores, the play-based group outperformed their peers over the course of the year. This suggests that the pedagogical approach not only influences immediate outcomes, but also shapes the students' long-term learning trajectories. The interaction effects in math also support the idea that playful learning contributes to better conceptual understanding. For example, in Algebraic Thinking and Measurement and Data, the play-based group made greater gains compared to the contemporary classroom, further highlighting the cognitive benefits of incorporating playful, hands-on learning experiences (Clements & Sarama, 2014).

Covariate Effects. Covariate analyses provided additional insight into the role of demographic variables on student outcomes. Age emerged as a significant predictor of performance in several domains, including Receptive Vocabulary and Math, confirming the importance of developmental stage for skill acquisition ($B = 5.692$, $p = .002$ for phonics; $B = 4.961$, $p = .01$ for vocabulary). However, no significant age differences were found between the two classrooms, indicating that the observed pedagogical effects were not confounded by age-related developmental differences. Gender and socioeconomic status (indicated by FRPL status) also played a role in the expression of the results, particularly in executive functions. For example, students eligible for FRPL demonstrated higher scores in Task Monitoring and Cognitive Regulation, which is consistent with research suggesting that children from lower-income backgrounds face additional challenges in executive function development ($B = -5.042$, $p = .018$ for task monitoring; Noble et al., 2007). These findings highlight the need for targeted interventions that address both academic and social-emotional needs, particularly for students from disadvantaged backgrounds.

Table 4. Aggregate Spearman's Rho Correlations of Posttest Variables

Variables (n = 31)	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
Receptive Vocabulary																										
1. Receptive Vocabulary	–	.325	.165	.196	.163	.350	.423*	.284	.232	.062	.221	.209	.092	-.182	-.183	-.088	-.242	-.372	-.402*	-.231	-.127	-.016	-.028	-.193	-.112	-.117
Reading																										
2. i-Ready Reading Overall†		–	.738***	.921***	.752***	.863***	.711***	.765***	.717***	.572**	.626***	.432*	.459*	-.644***	-.521**	-.387*	-.570**	-.495**	-.659***	-.310	-.689***	-.520**	-.614***	-.730***	-.669***	-.584**
3. Phonological Awareness			–	.628***	.466*	.556**	.435*	.491**	.615***	.618***	.562**	.296	.361	-.360	-.259	-.122	-.374	-.274	-.388*	-.184	-.407*	-.223	-.345	-.428*	-.388*	-.342
4. Phonics				–	.764***	.796***	.598***	.653***	.764***	.683***	.666***	.463*	.479**	-.759***	-.675***	-.566**	-.697***	-.546**	-.671***	-.350	-.790***	-.657***	-.738***	-.800***	-.743***	-.683***
5. High Frequency Words					–	.575**	.593***	.480**	.639***	.433*	.567**	.449*	.395*	-.739***	-.639***	-.516**	-.674***	-.543**	-.627***	-.377*	-.748***	-.575**	-.680***	-.786***	-.770***	-.696***
6. Vocabulary						–	.511**	.555**	.493**	.433*	.342	.320	.359	-.492**	-.400*	-.288	-.439*	-.331	-.521**	-.090	-.597***	-.459*	-.532**	-.641***	-.564**	-.531**
7. Comprehension: Literary Text							–	.583***	.445*	.240	.403*	.184	.324	-.452*	-.403*	-.311	-.430*	-.363	-.516**	-.249	-.388*	-.242	-.322	-.476*	-.385*	-.304
8. Comprehension: Info. Text								–	.560**	.228	.597***	.385*	.347	-.559**	-.424*	-.343	-.409*	-.497**	-.628***	-.374*	-.557**	-.498**	-.484**	-.594***	-.565**	-.414*
Math																										
9. i-Ready Math Overall†									–	.786***	.897***	.737***	.680***	-.749***	-.679***	-.546**	-.721***	-.658***	-.663***	-.516**	-.722***	-.537**	-.665***	-.677***	-.718***	-.485**
10. Number & Operations										–	.580**	.419*	.505**	-.472*	-.449*	-.361	-.486**	-.368	-.336	-.313	-.444*	-.318	-.422*	-.368	-.436*	-.252
11. Algebraic Thinking											–	.662***	.488**	-.719***	-.662***	-.546**	-.706***	-.640***	-.674***	-.497**	-.688***	-.555**	-.622***	-.658***	-.659***	-.558**
12. Measurement & Data												–	.423*	-.552**	-.507**	-.382*	-.529**	-.593***	-.571**	-.462*	-.498**	-.320	-.452*	-.455*	-.513**	-.330
13. Geometry													–	-.414*	-.350	-.308	-.322	-.216	-.167	-.167	-.452*	-.387*	-.477*	-.431*	-.440*	-.181
Executive Function																										
14. Global Executive Composite‡														–	.953***	.895***	.937***	.818***	.807***	.665***	.946***	.855***	.878***	.919***	.938***	.791***
15. Behavior Regulation Index†															–	.960***	.965***	.820***	.744***	.706***	.853***	.802***	.777***	.815***	.832***	.762***
16. Inhibit																–	.878***	.756***	.640***	.706***	.792***	.843***	.743***	.742***	.744***	.728***
17. Self-Monitor																	–	.833***	.784***	.686***	.851***	.729***	.761***	.831***	.825***	.749***
18. Emotional Regulation Index†																		–	.888***	.922***	.687***	.573**	.572**	.658***	.667***	.642***
19. Shift																			–	.700***	.735***	.586**	.602***	.737***	.727***	.653***
20. Emotional Control																				–	.491**	.454*	.401*	.439*	.464*	.471*
21. Cognitive Regulation Index†																					–	.913***	.968***	.983***	.962***	.846***
22. Initiate																						–	.917***	.879***	.837***	.822***
23. Working Memory																							–	.942***	.903***	.802***
24. Planning & Organization																								–	.941***	.851***
25. Task Monitoring																									–	.766***
26. Organizing Materials																										–

Note.† = Overall Score (with sub-scores underneath). ‡ = Overall Composite Score (with sub-scores underneath, including relevant Overall Scores). * = Correlation is significant at $p \leq .05$ (2-tailed). ** = Correlation is significant at $p \leq .01$ (2-tailed). *** = Correlation is significant at $p < .001$ (2-tailed). **Bolded** = .40-.59 "moderate," .60-.79 "strong," and .80-1.0 "very strong" correlations. Info = Informational.a

Implications

Researchers widely accept that experiments with a randomized controlled trial design are the "gold standard" of research (Hariton & Locascio, 2018), and quasi-experimental, naturalistic studies such as the one presented here are not sufficient because they do not use randomized selection or allocation and do not even necessarily use a carefully designed, evidence-based intervention. Even when they do, however, it is notoriously difficult to affect teacher practices and harder still to see effects on student outcomes (Korest & Carlson, 2022; Warmbold-Brann et al., 2017). The reality in most of the US is that state, district, and school-based educational decision-makers or leaders are reluctant to allow children who are identified as academically at-risk to play because they believe that playful learning and academic rigor aligned to instructional standards are mutually exclusive. The rarity of play-based instruction in Title I schools in the US, coupled with the challenges to access classrooms for research that has only increased since the COVID-19 pandemic (Greenberg, 2004, 2010; Waechter et al., 2023), meant that certain challenges had to be accepted in order to proceed with the study. Despite these challenges (i.e., the principal's decision to place the most advanced entering kindergarten students into one class, having only two classrooms for comparison with a small sample size, and district restrictions on the curriculum teachers were allowed to use), it is not hyperbolic to say the results are incredibly exciting and reinforce prior empirical results.

Alignment with prior studies. This study employed the use of specific assessment instruments for assessing executive function, language and literacy (i.e., reading), and math to generate explicit and realistic results aligned with a schoolified culture and illustrates the challenges in finding public Title I kindergartens using playful learning approaches. The findings from this study align with extant literature describing:

- the increased academization of kindergarten and the decrease in play (Bailey et al., 2019; Bassok et al., s 2016; Pyle et al., 2018; Repko-Erwin, 2017),
- the connections between school readiness and academic achievement (Blair & Raver, 2016: Madrick, 2020; Roos et al., 2019),
- and the potential for active, playful learning pedagogical approaches—in conjunction with standards-aligned instruction and assessment—to yield improved outcomes for children (Allee et al., 2023; King & Newstead, 2021; Pyle et al., 2018), particularly in a world still recovering from the COVID-19 pandemic (Hirsh-Pasek et al., 2024).

Statistically significant correlations between the dependent variables and classroom conditions explained the greater growth rates in the playful learning classroom. These findings add to the scholarship on instructional practices that build executive function skills and academic achievement, especially for vulnerable children (Allee-Herndon & Roberts, 2019). We understand increasingly that adversity often delays the development of skills critical to school readiness and academic achievement, and we recognize that pre-existing economically driven, systemic, and other adversity-related disparities at kindergarten entry often widen throughout a child's K-12 experience (Allee-Herndon et al., 2022; Bailey et al., 2019; Gilkerson et al., 2018; Mazzocco & Claessens, 2020). It is critical to address delays in the development of skills crucial for school readiness and academic achievement to mitigate these disparities and close these opportunity gaps (Allee-Herndon et al., 2022; Bailey et al., 2019; Gilkerson et al., 2018; Mazzocco & Claessens, 2020).

The curriculum and the pedagogy. Some curricula or interventions have been shown to have statistically significant effects on young children's outcomes, particularly those students experiencing economic disadvantage and other threats to development. For example, a 2019 systematic literature review (Allee-Herndon & Roberts) identified eight experimental or quasi-experimental studies that explored interventions designed to improve the executive function of children experiencing poverty or economic instability. Three of the studies analyzed curricular interventions: Head Start REDI and Preschool PATHS (Bierman et al., 2008), Incredible Years (Webster-Stratton et al., 2008), and Tools of the Mind (Blair & Raver, 2014). Since the systematic literature review, more evidence has been published supporting particular curricula (Incredible Years; Korest & Carlson, 2022) and pedagogical approaches like active, playful learning (Golinkoff & Hirsh-Pasek, 2016; Nesbitt et al., 2023) to push back against the schoolification of early childhood education. Each of these curricula, generally, employs the active, playful learning philosophies of blending what scientists have learned about how and what children need to learn.

The executive team of the Active, Playful Learning! Project (n.d.) explained that children learn the 6Cs (Golinkoff & Hirsh-Pasek, 2016; Hirsh-Pasek et al., 2020) through meaningful, joyful, socially interactive, active, engaging, and iterative learning experiences aligned with learning goals (Nesbitt et al., 2023). This evidence for these innovative, playful approaches is compelling, yet scripted curricula, such as those used in these two study classrooms, are often required of US kindergarten teachers, and the option to select a different curriculum such as Tools of the Mind, for example, is not available to most kindergarten teachers. This is partly why the results from this study and from similar, albeit much more experimental and larger-scaled work (the Active, Playful Learning! Project) is so exciting. This via media approach of using guided or purposeful play, initiated and designed by skillful teachers to meet specific learning goals but enacted upon by children, allows learning to be both joyful and rigorous without requiring the purchase of any new, packaged materials.

While this study was US-based, the challenges US kindergarten teachers have been facing in recent years seem to have only become more difficult and are certainly not exclusive to the US. The perception that play and rigor are mutually exclusive (Bassok et al., 2016; Dealey & Stone, 2018; Nitecki & Chung, 2013; Pyle et al., 2018) complicates cultural shifts in US education, though many countries prioritize both play and assessment (Allee et al., 2023; King & Newstead, 2021; Pyle et al., 2018; Synodi, 2010). The significant differences found in this study encourage further exploration of playful learning pedagogy alongside academic standards (Allee-Herndon et al., 2022) to impact a broad spectrum of child outcomes. Addressing the unmet goals of past decades, play, which has been dismissed in the name of standardizing assessment outcomes, could help close adversity-related gaps (Prioletta & Pyle, 2017; Sharkins et al., 2017; Walker et al., 2020; White et al., 2021). Understanding how to support this balanced approach is crucial, especially post-COVID-19, where didactic instruction has increased for many marginalized students (Donnelly & Patrinos, 2021; Dorn et al., 2020; Engzell et al., 2021).

Presenting a model for future studies. Decades of research in the science of learning and development have built upon and validated Vygotsky's (1978) social constructivist learning theory. "Humans learn best when they can be active and engaged in learning that is meaningful, socially interactive, iterative, and joyful" (Nesbitt et al., 2023, How We Learn section), particularly when they are learning the skills necessary for success across the lifespan (Golinkoff & Hirsh-Pasek, 2016; Hirsh-Pasek et al., 2020). Playful learning supports healthy development in all domains because it is more engaging, mentally active, collaborative, culturally connective, content-rich, and creative compared to contemporary practices, which are often passive, didactic, isolated, and boring. Given these benefits, one might assume that it would be an easy decision to ensure that all children learn this way, especially those who are considered to be at risk academically or otherwise. However, as discussed in the background section, this is not the case. In the US, it is often the children who are believed to be sufficiently high-performing who are allowed to engage in playful learning, much like in Teacher A's classroom. In contrast, "at-risk" children are more likely to lose play or playful learning opportunities due to behavioral infractions, academic deficiencies, or other perceived gaps. While the evidence is clear that play supports learning (Bailey et al., 2019; Blinkoff et al., 2023; Colliver et al., 2022; DeLuca et al., 2020; Hirsh-Pasek et al., 2020; Hirsh-Pasek et al., 2024), educational stakeholders require even more evidence to trust that play is not mutually exclusive to rigorous learning; in fact, play is rigorous learning.

Fortunately, the empirical literature on the benefits of active, playful learning for children's outcomes (e.g., cognitive, affective, behavioral) is growing. A search of the ERIC database using the terms "purposeful play or playful learning or active playful learning or guided play or play pedagogy" and "kindergarten or primary school or elementary school learning," yielded 75 empirical, full-text, peer-reviewed articles on play since 2002, when No Child Left Behind was passed. Our work to develop an evidence base with which to convince skeptical educational policy makers or leaders, while nascent, is emerging. This study, particularly with the more nuanced analysis of subscales on each valid and reliable assessment and across academic and cognitive domains, adds to the existing and growing evidence base that play effectively supports children's learning, particularly those who may need extra support to build their 6Cs and academic capacity (Golinkoff & Hirsh-Pasek, 2016; Hirsh-Pasek et al., 2020; Nesbitt et al., 2023). Much more work remains to be done, however, particularly to determine more precisely the specific playful learning curricular, environmental, or pedagogical components or factors that are most likely to yield positive outcomes for children.

Limitations and future directions

Although the results of this study provide compelling evidence in favor of play-based learning, there are several limitations to note. First, the sample size was relatively small (n = 28), which may limit the generalizability of the findings. Additionally, although age was controlled for in the analysis, other unmeasured factors such as classroom environment or teacher experience could have influenced the results. While the DiD analysis mitigated these limitations, it is unlikely that these approaches completely removed selection bias. Teacher A was likely able to play despite the district requirements for explicit, didactic, direct instruction because of her longer tenure at the school and because those students began the school year with an academic advantage, lowering their perceived "risk". Teacher effects outside of the pedagogical approach (e.g., personality, connections to students, professional training) may also have influenced the difference in outcomes by classroom condition. The successful use of a playful learning approach in kindergartens is predicated on leveraging high-quality learning environments and interactions (Pyle et al., 2018), but the observational and environmental data collected were limited in measuring teacher and classroom quality and were not formally included in this analysis. Analyzing this data more deeply, and using similar observational scales (the Classroom Assessment Scoring System; Pianta et al., 2008 and the Play Observation Scale; Rubin, 2001) would add nuance to our understanding of how environmental factors and teacher and student behaviors may also contribute to improved student outcomes. Future studies should aim for larger sample sizes and experimental designs with random selection and assignment while maintaining an analytical focus on multi-domain outcomes as in this study.

Future research should also aim to replicate these findings with larger, more diverse samples and explore the potential moderating effects of teacher-student interactions. In addition, Spearman's Rank Correlation analysis revealed strong associations between posttest measures of reading, math, and executive functions, suggesting that these skills may develop together. A closer examination of these relationships in future studies could provide valuable insights into how the different domains of learning support one another in early childhood. In addition to what has already been considered, it is important that future research better operationalize the elements of playful learning and contemporary classrooms in elementary schools as most of the research has been conducted in preschools (i.e., typically for non-compulsory programs in the US for children ages 3–5, with varying attendance patterns). Using other assessments of academic achievement in reading, math, and executive function may also influence future study outcomes and should be explored as different dependent variables may also shed light on the appropriateness of specific measures to assess academic achievement aligned with academic standards. Further exploration in hyper-academic, schoolified US kindergartens is warranted. Defining elements of playful learning and contemporary classrooms in elementary schools, using varied assessments, and exploring different dependent variables could provide deeper insights.

Conclusion

Despite limitations, the results support expanding studies on play for positive child outcomes (Fitzpatrick et al., 2014; Moreno et al., 2017; Raver et al., 2011; Skibbe et al., 2019). These studies connect playful learning pedagogy to improved academic and executive function outcomes (Gimbert et al., 2019; Meixner et al., 2019; Morgan et al., 2019), but it is probable using a playful learning, via media approach may also be helpful to reduce externalizing behaviors (e.g., discipline referrals, suspensions, absenteeism) and improve children's social-emotional skills (e.g., pro-social behaviors, cooperation, approaches to learning; Allee-Herndon et al., 2019). These would be desirable outcomes that may be connected to a more developmentally appropriate, child-centered, constructivist, and playful approach using the via media maxim (Allee et al., 2023). We know there are multiple threats to vulnerable, marginalized children, and disproportionate inequities in life and education have long-reaching impacts (Chetty et al., 2011; Dodge et al., 2015; Gimbert et al., 2019; Meixner et al., 2019; Morgan et al., 2019; Nesbitt et al., 2019; Skibbe et al., 2019). It is insufficient to understand these connections without also knowing more about effective strategies to mitigate the risks and threats to reducing inequality in our schools.

Overall, this study provides compelling evidence of the benefits of play-based learning in early childhood education. The significant main and interaction effects for vocabulary, reading, and math underscore the importance of creating a learning environment that encourages active, meaningful engagement with content. Additionally, the findings highlight the need to address the social-emotional needs of students from disadvantaged backgrounds, as evidenced by the covariate effects related to FRPL status. By integrating playful learning strategies into the classroom, educators can better support the holistic development of young learners. A playful learning, via media approach to learning may reduce externalizing behaviors and improve social-emotional skills while simultaneously supporting increased academic outcomes compared to contemporary, schoolified approaches. Addressing threats to marginalized children and reducing educational inequities requires effective strategies, and further research is needed to mitigate these risks to achieve equity, and joyful learning, in schools.

Conflicts of interest

The authors declare no conflicts of interest.

References

Active playful learning. (n.d.). Active Playful Learning. https://activeplayfullearning.com.

Agorastos, A., Pervanidou, P., Chrousos, G. P., & Baker, D. G. (2019). Developmental trajectories of early life stress and trauma: a narrative review on neurobiological aspects beyond stress system dysregulation. *Frontiers in psychiatry*, 118. https://doi.org/10.3389/fpsyt.2019.0018

American Institutes for Research (2020). *i-Ready® Diagnostic Reading.* https://charts.intensiveintervention.org/screening/tool/?id=7534542c4f422f85.

Ansari, A., & Gottfried, M. A. (2020). Early childhood educational experiences and preschool absenteeism. *The Elementary School Journal, 121(1),* 34–51. https://doi.org/10.1086/709832

Au, W. (2007). High-stakes testing and curricular control: A qualitative metasynthesis. *Educational Researcher, 36(5),* 258–267. https://doi.org/10.3102/0013189X07306523.

Allee, K. A., Clark, M. H., Roberts, S. K., & Hu, B. Y. (2023). Exploring play, executive function, and academic achievement in two Title I kindergartens. *Journal of Research in Childhood Education*, *37(2),* 197–215. https://doi.org/10.1080/02568543.2022.2139784.

Allee-Herndon, K. A., & Roberts, S. K. (2019). Poverty, self-regulation and executive function, and learning in K-2 classrooms: A systematic literature review of current empirical research. *Journal of Research in Childhood Education, 33(3),* 345–362. https://doi.org/10.1080/02568543.2019.1613273.

Allee-Herndon, K. A., & Roberts, S. K. (2021). The power of purposeful play in primary grades: Adjusting pedagogy for children's needs and academic gains. *Journal of Education*, *201(1),* 54–63. https://doi.org/10.1177/0022057420903272.

Allee-Herndon, K. A., Roberts, S. K., Hu, B. Y., Clark, M. H., & Stewart, M. L. (2022). Let's talk play! Exploring the possible benefits of play-based pedagogy on language and literacy learning in Title I kindergarten classrooms. *Early Childhood Education Journal, 50(1),* 119–132. https://doi.org/10.1007/s10643-020-01141-6.

Allee-Herndon, K. A., Taylor, D. D., & Roberts, S. K. (2019). Putting play in its place: Presenting a continuum to decrease mental health referrals and increase purposeful play in classrooms. *International Journal of Play, 8(2),* 186–203. https://doi.org/10.1080/21594937.2019.1643993.

Bailey, R., & Jones, S. M. (2019). An integrated model of regulation for applied settings. *Clinical Child and Family Psychology Review, 22(1),* 2–23. https://doi.org/10.1007/s10567-019-00288-y.

Bailey, R., Meland, E. A., Brion-Meisels, G., & Jones, S. M. (2019). Getting developmental science back into schools: Can what we know about self-regulation help change how we think about "No Excuses"? *Frontiers in Psychology*, 10, Article 1885. https://doi.org/10.3389/fpsyg.2019.01885.

Bassok, D., Latham, S., & Rorem, A., (2016). Is kindergarten the new first grade? *AERA Open*, 1(4), 1–31. https://doi.org/10.1177/2332858415616358.

Bishop, P. A., Downes, J. M., Netcoh, S., Farber, K., DeMink-Carthew, J., Brown, T., & Mark, R. (2020). Teacher roles in personalized learning environments. *The Elementary School Journal, 121(2),* 311–336. https://doi.org/10.1086/711079.

Blair, C. (2016). Developmental science and executive function. *Current Directions in Psychological Science, 25(1),* 3–7. https://doi.org/10.1177/0963721415622634.

Blair, C., & Raver, C. C. (2014). Closing the achievement gap through modification of neurocognitive and neuroendocrine function: Results from a cluster randomized controlled trial of an innovative approach to the education of children in kindergarten. *PLOS One, 9(11),* 1–13, e112393. https://doi.org/10.1371/journal.pone.0112393.

Blair, C., & Raver, C. C. (2015). School readiness and self-regulation: a developmental psychobiological approach. *Annual Review of Psychology, 66(2015),* 711–31. https://doi.org/10.1146/annurev-psych-010814-015221.

Blair, C., & Raver, C. C. (2016). Poverty, stress, and brain development: New directions for prevention and intervention. *Academic Pediatrics, 16(3):* S30–S36. https://doi.org/10.1016/j.acap.2016.01.010.

Blinkoff, E., Fletcher, K., Wright, C., Espinoza, S., & Hirsh-Pasek, K. (2023, October 9). *Commentary: Tracking the winds of change on the American education policy landscape: the emergence of play-based learning legislation and its implications for the classroom.* The Brookings Institution. https://www.brookings.edu/articles/tracking-the-winds-of-change-on-the-american-education-policy-landscape-the-emergence-of-play-based-learning-legislation-and-its-implications-for-the-classroom.

Brown, C. P., & Lan, Y.-C. (2015). A qualitative metasynthesis comparing U.S. teachers' conceptions of school readiness prior to and after the implementation of NCLB. *Teaching and Teacher Education, 45,* 1–13. https://doi.org/10.1016/j.tate.2014.08.012.

Brown, C. P., Englehardt, J., Ku, D. H., & Barry, D. P. (2019). "Where's the joy in the classroom?" Families' sensemaking of the changed kindergarten. *The Elementary School Journal, 120(2),* 319–346. https://doi.org/10.1086/705964.

Brownell, M. D., Nickel, N. C., Chateau, D., Martens, P. J., Taylor, C., Crockett, L., Katz, A., Sarkar, J., Burland, E., Goh, C. Y., & the PATHS Equity Team. (2015). Long-term benefits of full-day kindergarten: A longitudinal population-based study. *Early Child Development and Care, 185(2),* 291–316. https://doi.org/10.1080/03004430.2014.913586.

Burchinal, M., McCartney, K., Steinberg, L., Crosnoe, R., Friedman, S. L., McLoyd, V., Pianta, R., & NICHD Early Child Care Research Network. (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(5),* 1404–1420. https://doi.org/10.1111/j.1467-8624.2011.01620.x.

Center on Educational Policy. (2007). *Answering the question that matters most: Has student achievement increased since No Child Left Behind?* Center on Educational Policy.

Chetty, R., Friedman, J. N., Hilger, N., Saez, E., Schanzenbach, D. W., & Yagan, D. (2011). How does your kindergarten classroom affect your earnings? Evidence from Project Star. *Quarterly Journal of Economics, Oxford University Press, 126(4),* 1593–1660. https://doi.org/10.1093.qje/qjr041.

Chiatovich, T., & Stipek, D. (2016). Instructional approaches in kindergarten: What works for whom? *The Elementary School Journal, 117(1),* 1–29. https://doi.org/10.1086/687751.

Clements, D. H., & Sarama, J. (2014). Learning trajectories: Foundations for effective, research-based education. In A. P. Maloney, J. Confrey, & K. H. Nguyen (Eds.), *Learning over time: Learning trajectories in mathematics education* (pp. 1–30). IAP Information Age Publishing.

Coldren, J. T. (2013). Cognitive control predicts academic achievement in kindergarten children. *Mind, Brain, and Education, 7(1),* 40–48. https://doi.org/10.1111/

mbe.12006.
Colliver, Y., Harrison, L. J., Brown, J. E., & Humburg, P. (2022). Free play predicts self-regulation years later: Longitudinal evidence from a large Australian sample of toddlers and preschoolers. *Early Childhood Research Quarterly, 59,* 148–161. https://doi.org/10.1016/j.ecresq.2021.11.011.

Curran, F. C. (2017). Income-based disparities in early elementary school science achievement. *The Elementary School Journal, 118(2),* 207–231. https://eric.ed.gov/?id=EJ1162543.

Curriculum Associates (n.d.). *i-Ready Diagnostic and Florida Standards Assessment* (FSA) *Linking Study Overview*. https://www.curriculumassociates.com/-/media/mainsite/files/i-ready/linking-studies/iready-diagnostic-assessments-linking-study-overview-florida-2018.pdf?la=en&hash=6F65D84B0A896CBECE03C8B93E05CE61.

Curriculum Associates. (2018). *i-Ready Scale Score Placement Tables (2018-2019 School Year).* https://www.esboces.org/cms/lib/NY01914091/Centricity/Domain/533/iready-placement-tables-2018-2019.pdf.

de Bilde, J., Boonen, T., Speybroeck, S., De Fraine, B., Van Damme, J., Goos, M. (2015). Experiential education in kindergarten: Associations with school adjustment and the moderating role of initial achievement. *The Elementary School Journal, 116(2),* 173–197. https://doi.org/10.1086/683802.

Dealey, R. P., & Stone, M. H. (2017). Exploring out-of-school play and educational readiness. *Early Childhood Education Journal, 46(2),* 201–208. https://doi.org/10.1007/s10643-017-0849-7.

Dean, D., & Kuhn, D. (2007). Direct instruction vs. discovery: The long view. *Science Education, 91(3),* 384–397. https://doi.org/10.1002/sce.20194.

Dee, T. S., Jacob, B., & Schwartz, N. L. (2013). The effects of NCLB on school resources and practices. *Educational Evaluation and Policy Analysis, 35(2),* 252–279. https://doi.org/10.3102/0162373712467080.

DeLuca, C., Pyle, A., & Valiquette, A., & LaPointe-McEwan, D. (2020). New directions for kindergarten education: Embedding assessment in play-based learning. *The Elementary School Journal, 120(3),* 455–479. https://doi.org/10.1086/707008.

Dodge, K. A., Bierman, K. L., Cole, J. D., Greenberg, M. T., Lochman, J. E., McMahon, R. J., & Pinderhughes, E. E. (2015). Impact of early intervention on psychopathology, crime, and well-being at age 25. *The American Journal of Psychiatry, 172(1),* 59–70. https://doi.org/10.1176/appi.ajp.2014.13060786.

Donnelly, R., & Patrinos, H. A. (2021). Learning loss during COVID-19: An early systematic review. *Prospects,* 1–9. https://doi.org/10.1007/s11125-021-09582-6.

Dorn, E., Hancock, B., Sarakatsannis, J., & Viruleg, E. (2020). COVID-19 and learning loss—disparities grow and students need help. *McKinsey & Company, December, 8,* 6–7. https://wasaoly.org/WASA/images/WASA/5.0%20Professional%20Development/4.2%20Conference%20Resources/Winter/2021/covid-19-and-learning-loss-disparities-grow-and-students-need-help-v3.pdf.

Dunn, L. M., & Dunn, D. M., (2007). *Peabody Picture Vocabulary Test, Fourth Edition.* Pearson Assessments.

Education Commission of the States. (2024). *50-state comparison: State K-3 policies.* https://www.ecs.org/50-state-comparison-state-k-3-policies-2023/.

Egger, F., Benzing, V., Conzelmann, A., & Schmidt, M. (2019). Boost your brain, while having a break! The effects of long-term cognitively engaging physical activity breaks on children's executive functions and academic achievement. *PLoS One, 14(3),* 1–20, e02124882. https://doi.org/10.1371/journal.pone.0212482.

Ellwood-Lowe, M. E., Sacchet, M. D., & Gotlib, I. H., (2016). The application of neuroimaging to social inequity and language disparity: A cautionary examination. *Developmental Cognitive Neuroscience, December (22),* 1–8. https://doi.org./10.1016/j.dcn.2016.10.001.

Engzell, P., Frey, A., & Verhagen, M. D. (2021). Learning loss due to school closures during the COVID-19 pandemic. *Proceedings of the National Academy of Sciences, 118(17),* e2022376118. https://doi.org/10.1073/pnas.202237611.

Ehri, L. C. (2005). Development of sight word reading: Phases and findings. In M. J. Snowling & C. Hulme (Eds.), *The science of reading: A handbook* (pp. 135–154). Blackwell Publishing. https://doi.org/10.1002/9780470757642.ch8.

Every Student Succeeds Act, 20 U.S.C. § 6301 (2015).

Fitzpatrick, C., McKinnon, R. D., Blair, C. B., & Willoughby, M. T. (2014). Do preschool executive function skills explain the school readiness gap between advantaged and disadvantaged children? *Learning and Instruction, 30*(April), 25–31. https://doi.org/j.learninginstruc.2013.11.003.

Gersten, R., & Carnine, D. (1984). Direct instruction mathematics: A longitudinal evaluation of low-income elementary school students. *The Elementary School Journal, 84(4),* 395–407. https://www.jstor.org/stable/1001366 .

Gilkerson, J., Richards, J. A., Warren, S. F., Oller, D. K., Russo, R., & Vohr, B. (2018). Language experience in the second year of life predicts language outcomes in late childhood. *Pediatrics, 142(4),* e20174276. https://doi.org/10.1542/peds.2017-4276.

Gimbert, F., Camos, V., Gentaz, E., & Mazens, K. (2019). What predicts mathematics achievement? Developmental change in 5- and 7-year-old children. *Journal of Experimental Child Psychology, 178*(February), 104–120. https://doi.org/10.1016/j.jecp.2018.09.013.

Gioia, G. A., Isquith, P. K., Guy, S. C., & Kenworthy, L. (2015). *Behavior Rating Inventory of Executive*

Function, Second Edition (BRIEF2) Professional Manual. PAR.

Golinkoff, R. M., & Hirsh-Pasek, K. (2016). *Becoming brilliant: What science tells us about raising successful children.* American Psychological Association.

Greenberg, M. T. (2004). Current and future challenges in school-based prevention: The researcher perspective. *Prevention Science, 5,* 5–13. https://doi.org/10.1023/B:PREV.0000013976.84939.55.

Greenberg, M. T. (2010). School-based prevention: Current status and future challenges. *Effective Education, 2(1),* 27–52. https://doi.org/10.1080/19415531003616862.

Halle, T., Forry, N., Hair, E. C., Perper, K., Wandner, L. D., & Whittaker, J. V. (2009, June). *Disparities in early learning and development: Lessons from the Early Childhood Longitudinal Study – Birth Cohort (ECLS-B).* Child Trends. https://www.childtrends.org/publications/disparities-in-early-learning-and-development-lessons-from-the-early-childhood-longitudinal-study-birth-cohort-ecls-b.

Hariton, E., & Locascio, J. J. (2018). Randomised controlled trials–The gold standard for effectiveness research. *BJOG: An International Journal of Obstetrics and Gynaecology, 125(13),* 1716–1716. https://doi.org/10.1111/1471-0528.15199.

Harms, T., Jacobs, E. V., & White, D. R. (2013). *School-Age Care Environment Rating Scale Updated (SACERS)* (Updated ed.). Teachers College Press.

Hatch, J. A. (2002). Accountability shovedown: Resisting the standards movement in early childhood education. *The Phi Delta Kappan, 83(6),* 457–462. http://www.jstor.org/stable/20440167.

Hirsh-Pasek, K., Golinkoff, R. M., Berk, L. E., & Singer, D. G. (2009). *A mandate for playful learning in preschool: Presenting the evidence.* Oxford University Press.

Hirsh-Pasek, K., Hadani, H., Blinkoff, E., & Golinkoff, R. M. (2020). *A new path to education reform: Playful learning promotes 21st century skills in school and beyond.* The Brookings Institution. https://www.brookings.edu/policy2020/bigideas/a-new-path-to-education-reform-playful-learningpromotes-21st-century-skills-in-schools-and-beyond/.

Hirsh-Pasek, K., Masters, A. S., Christie, S., Gibbs, H. M., Evans, N. S., Fletcher, K. K., Pesch, A., Yang, H., Fan, W., Todaro, R. D., Golinkoff, R. M., & Xu, F. (2022). *Where global science meets playful learning: Implications for home, school, and city planning* [White paper]. The Lego Foundation and the Yidan Foundation. https://learningthroughplay.com/explore-the-research/where-global-science-meets-playful-learning-implications-for-home-school-cities-and-digital-spaces.

Hirsh-Pasek, K., Winthrop, R., & Shah, S. (2024, May 2). *Commentary: Why play is serious work.* The Brookings Institution. https://www.brookings.edu/articles/why-play-is-serious-work/ .

Hu, B., Fan. X., Leong, S. S. L., and Li, K. (2015). Why is group teaching so important to Chinese children's development? *Australasian Journal of Early Childhood, 40(1),* 4–12. https://doi.org/10./1177/183693911504000102.

Improving America's Schools Act, 20 U.S.C. § 6301 et seq. (1994).

Iruka, I. U., Sheridan, S., Koziol, N., Schumacher, R., Kerby, H., Prokasky, A., & Choi, D. (2022). Examining malleable factors that explain the end-of-kindergarten racial/ethnic gaps. *The Elementary School Journal, 122(3),* 378–410. https://doi.org/10.1086/718072.

King, P. & Newstead, S. (Eds.) (2021). *Further perspectives on researching play from a playwork perspective: Process, playfulness, rights-based and critical reflection.* Routledge.

Korest, R., & Carlson, J. S. (2021). A meta-analysis of the current state of evidence of the Incredible Years Teacher Classroom Management program. *Children (Basel), 9(1),* 24. https://doi.org/10.3390/children9010024.

Koslouski, J. B., & Stark, K. (2021). Promoting learning for students experiencing adversity and trauma: The everyday, yet profound, actions of teachers. *The Elementary School Journal, 121(3),* 430–453. https://doi.org/10.1086/712606.

Lee, P., & Bierman, K. L. (2016). Profiles of kindergarten classroom and elementary school contexts: Associations with first-grade outcomes of children transitioning from Head Start. *The Elementary School Journal, 117(1),* 119–142. https://doi.org/10.1086/687813.

Lund, E. (2020). The relation between vocabulary knowledge and phonological awareness in children with cochlear implants. *Journal of Speech, Language, and Hearing Research, 63(7),* 2386–2402. htttps://doi.org/10.1044/2020_JSLHR-19-00259.

Madrick, J. (2020). *Invisible Americans: The tragic cost of child poverty.* Alfred A. Knopf.

Mazzocco, M., & Claessens, A. (2020). Introduction to the special issue: Parents supporting early mathematical thinking. *Early Childhood Research Quarterly, 50,* 1–3. https://doi.org/10.1016/j.ecresq.2019.07.007.

McDonald, P. (2018). Observing, planning, guiding: How an intentional teacher meets standards through play. *Young Children, 73(1).* https://www.naeyc.org/resources/pubs/yc/mar2018/observing-planning-guiding.

McDowell, K., Carroll, J., & Ziolkowski, R. (2013). Relations between phonological awareness and vocabulary in preschoolers. *International Journal of Liberal Arts and Social Science, 1(2),* 75–84. https://ijlass.org/data/frontImages/October2013/8.pdf.

Meixner, J. M., Warner, G. J., Lensing, N., Schiefele, U., & Elsner, B. (2019). The relationship between executive functions and reading comprehension in primary-school students: A cross-lagged-panel analysis. *Early Childhood Research Quarterly, 46(1),* 62–74. https://doi.org/10.1016/j.ecresq.2018.04.010.

Milner, A. R., Sondergeld, T. A., Demir, A., Johnson, C. C., & Czerniak, C. M. (2017). Elementary teachers' beliefs about teaching science and classroom practice: An examination of pre/post NCLB testing in science. *Journal of Science Teacher Education, 23(2),* 111–132. https://doi.org/10.1007/s10972-011-9230-7.

Moreno, A. J., Shwayder, I., & Friedman, I. D. (2016). The function of executive function: Everyday manifestations of regulated thinking in preschool settings. *Early Childhood Education Journal, 45(2),* 143–153. https://doi.org/10.1007/s10643-016-0777-y.

Morgan, P. L., Farkas, G., Wang, Y., Hillemeier, M. M., Oh, Y., & Maczuga, S. (2019). *Executive function deficits in kindergarten predict repeated academic difficulties across elementary school. Early Childhood Research Quarterly,46(1),* 20–32. https://doi.org/10.1016/j.ecresq.2018.06.009.

Morgan, P. L., Li, H., Farkas, G., Cook, M., Pun, W. H., & Hillemeier, M. M. (2017). Executive functioning deficits increase kindergarten children's risk for reading and mathematics difficulties in first grade. *Contemporary Educational Psychology, 50,* 23–32. https://doi.org/10.1016/j.cedpsych.2016.01.004.

Mraz, K., Porcelli, A., & Tyler, C. (2016). *Purposeful play: A teacher's guide to igniting deep and joyful learning across the day.* Heinemann.

Myers, J. M., & Ankrum, J. W. (2018). Explicit vocabulary instruction in kindergarten: Case studies of students with and without language disorders. *Early Childhood Education Journal, 46,* 683–691. https://doi.org/10.1007/s10643-018-0896-8.

National Association for the Education of Young Children. (2021). *Developmentally appropriate practice in early childhood programs serving children from birth through age 8* (S. Friedman, B. L. Wright, M. L. Masterson, B. Willer, & S. Bredekamp, Eds.; 4th ed.).

National Council of Teachers of Mathematics. (2014). *Principles to actions: Ensuring mathematical success for all.*

National Research Council, Division of Behavioral and Social Sciences Education, Center for Education, Board on Science Education, & Kindergarten Through Eighth Grade Committee on Science Learning. (2007). *Taking science to school: Learning and teaching science in grades K-8* (A. W. Shouse, H. A. Schweingruber, & R. A. Duschl, Eds.; 1st ed.). National Academies Press.

Nesbitt, K. T., Blinkoff, E., Golinkoff, R. M., & Hirsh-Pasek, K. (2023). Making schools work: An equation for active playful learning. *Theory Into Practice, 62*(2: Creative Arts, Physical Activity, and Children's Play in Early Childhood), 141–154. https://doi.org/10.1080/00405841.2023.2202136.

Nesbitt, K. T., Fuhs, M. W., & Farran, D. C. (2019). Stability and instability in the co-development of mathematics, executive function skills, and visual-motor integration from prekindergarten to first grade. *Early Childhood Research Quarterly, 46(1),* 262–274. https://doi.org/10.1016/j.ecresq.2018.02.003.

Nitecki, E., & Chung, M.H. (2013). What is not covered by the standards: How to support emergent literacy in preschool classrooms. *Language and Literacy Spectrum, 23,* 46–56. https://files.eric.ed.gov/fulltext/EJ1005285.pdf.

No Child Left Behind Act of 2001, P.L. 107–110, 20 U.S.C. § 6319 (2002).

Noble, K. G., McCandliss, B. D., & Farah, M. J. (2007). Socioeconomic gradients predict individual differences in neurocognitive abilities. *Developmental Science, 10(4),* 464-480. https://doi.org/10.1111/j.1467-7687.2007.00600.x.

Parker, R., Thomsen, B. S., & Berry, A. (2022). Learning through play at school – A framework for policy and practice. *Frontiers in Education,* 7:751801. https://doi.org/10.3389/feduc.2022.751801.

Patton M. Q. (2002). *Qualitative research and evaluation methods* (3rd ed.). Sage Publications.

Piaget, J. (1977). *Biology and knowledge: An essay on the relations between organic regulations and cognitive processes.* University of Chicago Press.

Pianta, R. C., La Paro, K. M., & Hamre, B. K. (2008). *Classroom Assessment Scoring System™: Manual K-3.* Paul H Brookes Publishing.

Prioletta, J., & Pyle, A. (2017). Play and gender in Ontario kindergarten classrooms: implications for literacy learning. *International Journal of Early Years Education, 25(4),* 393-408. https://doi.org/10.1080/09669760.2017.1390446.

Pyle, A., & Danniels, E. (2017). A continuum of play-based learning: The role of the teacher in play-based pedagogy and the fear of hijacking play. *Early Education and Development, 28(3),* 274–289. https://doi.org/10.1080/10409289.2016.1220771.

Pyle, A., Poliszczuk, D., Danniels, E. (2018). The challenges of promoting literacy integration within a play-based learning kindergarten program: Teacher perspectives and implementation. *Journal of Research in Childhood Education, 32(2),* 219–233. https://doi.org/10.1080/02568543.2017.1416006.

Ranz-Smith, D. J. (2007). Teacher perception of play: In leaving no child behind are teachers leaving childhood behind? *Early Education and Development, 18(2),* 271–303. https://doi.org/10.1080/10409280701280425.

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(1),* 362–378. https://doi.org/10.1111/j.1467-8624.2010.01561.x.

Razza, R. A., Martin, A., & Brooks-Gunn, J. (2015). Are approaches to learning in kindergarten associated with academic and social competence similarly? *Child and Youth Care Forum, 44(6),* 757–776. https://doi.org/10.1007/s10566-015-9307-0.

Repko-Erwin, M. E. (2017). Was kindergarten left behind? Examining US kindergarten as the new first grade in the wake of "No Child Left Behind." *Global Education Review, 4(2),* 58–74. https://files.eric.ed.gov/fulltext/EJ1155001.pdf.

Riek, J. (2014). Reinstating P.L.A.Y. pedagogy in early elementary classrooms. *The International Journal of Early Childhood Learning, 22(1),* 1–9.

Ring, E., & O'Sullivan, L. (2018). Dewey: A panacea for the 'schoolification' epidemic. *International Journal of Primary, Elementary and Early Years Education, 46*(4: John Dewey's Democracy and Education: The Relevance for Primary Education 100 Years On), 402–410. https://doi.org/10.1080/03004279.2018.1445474.

Roos, L. L., Wall-Wieler, E., & Lee, J.B. (2019). Poverty and early childhood outcomes. *Pediatrics, 143(6),* e20183426. https://doi.org/10.1542/peds.2018-3426.

Rothwell, J. (2016, December 23). *The declining productivity of education.* The Brookings Institution. https://www.brookings.edu/blog/social-mobility-memos/2016/12/23/the-declining-productivity-of-education/.

Rowe, M. L., Raudenbush, S. W., & Goldin-Meadow, S. (2012). The pace of vocabulary growth helps predict later vocabulary skill. *Child Development, 82(2),* 508–525. https://doi.org/10.111/j.1467-8624.2011.01710.x.

Rubin, K. H. (2001). *The play observation scale (POS) – Revised.* University of Maryland.

Schweinhart, L. J. (2018). Lessons on sustaining early gains from the Life-Course Study of Perry Preschool. In A. J. Reynolds & J. A. Temple (Eds.), *Sustaining early childhood learning gains: Program, school, and family influences: Vol. Part III-School and Family Processes of Impacts over Time* (pp. 254–267). Cambridge University Press. https://doi.org/10.1017/9781108349352.012.

Sharkins, K., Newton, A. B., Causey, C., & Ernest, J. M. (2017). Flipping theory: Ways in which children's experiences in the 21st century can provide insight into the theories of Piaget and Vygotsky. *International Journal of Early Childhood Education Care, 6,* 11–18. http://dx.doi.org/10.37134/saecj.vol6.2.2017.

Shonkoff, J. P. (2011). Protecting brains, not simply stimulating minds. *Science, 333(6045),* 982–983. https://doi.org/10.1126/science.1206014.

Skiba, R. J., Horner, R. H., Chung, C. G., Rausch, M. K., May, S. L., & Tobin, T. (2011). Race is not neutral: A national investigation of African American and Latino disproportionality in school discipline. *School Psychology Review, 40(1),* 85–107. https://doi.org/10.1080/02796015.2011.12087730.

Skibbe, L. E., Montroy, J. J., Bowles, R. P., & Morrison, F. J. (2019). Self-regulation and the development of literacy and language achievement from preschool through second grade. *Early Childhood Research Quarterly, 46(1),* 240–251. https://doi.org/10.1016/j.ecresq.2018.02.005.

Stockard, J., Wood, T. W., Coughlin, C., and Khoury, C. R. (2018). The effectiveness of direct instruction curricula: A meta-analysis of a half century of research. *Review of Educational Research, 88(4),* 479–507. https://doi.org/10.3102/0034654317751919.

Synodi, E. (2010). Play in the kindergarten: The case of Norway, Sweden, New Zealand, and Japan. *International Journal of Early Years Education, 18(3),* 185–200. https://doi.org/10.1080/09669760.2010.521299.

Taylor, J., & Bilbrey, J. (2012). Effectiveness of inquiry based and teacher directed instruction in an Alabama elementary school. Journal of Instructional Pedagogies, 8. https://files.eric.ed.gov/fulltext/EJ1097117.pdf.

Trawick-Smith, J., Swaminathan, S., & Liu, X. (2016). The relationship of teacher-child play interactions to mathematics learning in preschool. *Early Child Development and Care, 186(5),* 1–18. https://doi.org/10.1080/03004430.2015.1054818.

US Department of Education. (2018, November 7). *Improving Basic Programs Operated by Local Educational Agencies (Title I, Part A).* https://www2.ed.gov/programs/titleiparta/index.html.

Vitiello, V. E., & Greenfield, D. B. (2017). Executive functions and approaches to learning in predicting school readiness. *Journal of Applied Developmental Psychology, 53*(November), 1–9. https://doi.org/10.1016/j.appdev.2017.08.004.

Vygotsky, L. (1978). Interaction between learning and development. In *Mind and society* (pp. 79–91). Harvard University Press.

Waechter, N., Kalmus, V., Mascheroni, G., & Opermann, S. (2023). Large-scale comparative school-based survey research: Challenges and solutions for sampling, fieldwork and informed consent. *Methods, Data, Analyses, 17(2),* 303–324. https://doi.org/10.12758/mda.2023.03.

Walker, C. L., & Shore, B. M. (2015). Understanding classroom roles in inquiry education: Linking role theory and social constructivism to the concept of role diversification. *SAGE Open.* https://doi.org/10.1177/2158244015607584.

Wallace, C. E., & Russ, S. W. (2015). Pretend play, divergent thinking, and math achievement in girls: A longitudinal study. *Psychology of Aesthetics, Creativity, and the Arts, 9(3),* 296–305. https://doi.org/10.1037/a0039006.

Warmbold-Brann, K., Burns, M. K., Preast, J. L., Taylor, C. N., & Aguilar, L. N. (2017). Meta-analysis of the effects of academic interventions and modifications on student behavior outcomes. *School Psychology Quarterly, 32(3),* 291–305. https://psycnet.apa.org/doi/10.1037/spq0000207.

Weisberg, D. S., Hirsh & Pasek, K., & Golinkoff, R. M. (2013). Guided play: Where curricular goals meet a playful pedagogy. *Mind, Brain, and Education, 7(2),* 104-112. https://doi.org/10.1111/mbe.12015.

Weisberg, D. S., Hirsh-Pasek, K., Golinkoff, R. M., Kittredge, A. K., & Klahr, D. (2016). Guided play: Principles and practices. *Current Directions in Psychological Science, 25(3),* 177–182. https://doi.org/ 10.1177/0963721416645512

West, J., Hausken, E. G., & Collins, M. (1993, September). *Readiness for kindergarten: Parent and teacher beliefs* (NCES 93–257). National Center for Education Statistics. https://nces.ed.gov/pubs93/93257.pdf.

White, R. E., Thibodeau-Nielsen, R. B., Palermo, F., & Mikulski, A. M. (2021). Engagement in social pretend play predicts preschoolers' executive function gains across the school year. *Early Childhood Research Quarterly, 56,* 103–113. https://doi.org/https://doi.org/10.1016/j.ecresq.2021.03.005.

Wood, C., Wanzek, J., & Schatschneider, C. (2022). The relation between teachers' communicative behaviors and class-level SES. *The Elementary School Journal, 122(4),* 534–556. https://doi.org/10.1086/719506.

Zosh, J. M., Hirsh-Pasek, K., Hopkins, E. J., Jensen, H., Liu, C., Neale, D., Solis, S. L., & Whitebread, D. (2018). Accessing the inaccessible: Redefining play as a spectrum. *Frontiers in Psychology, 9,* 1124. https://doi.org/10.3389/fpsyg.2018.01124.

Zosh, J. M., Hopkins, E. J., Jensen, H., Liu, C., Neale, D., Hirsh-Pasek, K., Solis, S. L., & Whitebread, D. (2017). *Learning through play: A review of the evidence* (white paper). The LEGO Foundation, DK. https://cms.learningthroughplay.com/media/wmtlmbe0/learning-through-play_web.pdf.