



Evaluation of Child Outcomes in Nine Child-Parent Centers: Report for 2018–19

August 2021

Prepared for:

IFF Pay For Success I, LLC

333 S. Wabash Avenue, Suite 2800

Chicago, Illinois 60604

Attention: Dana Lieberman, Sr. Vice President, Capital Solutions

E-mail: dliberman@iff.org

SRI Education™

A DIVISION OF SRI INTERNATIONAL

Copy to:

IFF

333 S. Wabash Ave., Suite 2800.

Chicago, Illinois 60604

Attention: Charles M. Biggam III, Chief Legal Counsel

E-mail: cbiggam@iff.org

Prepared by:

SRI International

Erika Gaylor

Kate Ferguson

Patrick Thornton

Xin Wei

Donna Spiker

Suggested citation:

Gaylor, E., Ferguson, K., Thornton, S.P., Wei, X., & Spiker, D. (2021). *Evaluation of child outcomes in nine Child-Parent Centers: Report for 2018-19*. Prepared for IFF Pay for Success I, LLC. Menlo Park, CA: SRI International.

Contents

- Background..... 1
- CPC Program Model.....2
 - CPC Model Description2
 - Expected Outcomes from the CPC Program Model.....4
 - School Readiness4
 - Third-Grade Reading and Literacy4
 - Reduced Special Education Use.....5
 - Chicago PFS Project (PFS-CPC Project)6
- Evaluation Design7
 - Analysis Approach.....9
 - CPC Intervention Sample Included in Analysis9
 - Analyzing Impact on Kindergarten Readiness.....13
 - Analyzing Impact on Third Grade Reading.....15
 - Analyzing Impact on Special Education Placement.....17
 - Selecting a Comparison Group18
- Results.....23
 - Kindergarten Readiness23
 - Special Education Placement.....24
 - Cohort 4 Kindergarten Special Education Placement.....24
 - Cohort 3 First Grade Education Placement25
 - Cohort 2 Second Grade Special Education Placement25
 - Cohort 1 Third Grade Special Education Placement25
 - Third Grade Reading27
 - Cohort 1 Third Grade Reading27
- Discussion27
 - Special Education Placement Findings.....28
 - Third Grade Reading Findings.....33
 - Limitations34
 - Conclusion35
- References36
- Appendices43

Appendix A: Chicago Child-Parent Center Social Impact Bond Evaluation Plan..... A-1
Appendix B: Timing of Cohorts..... B-1
Appendix C: Propensity Score Approach C-1

List of Exhibits

Exhibit 1. CPC Program Model Components	3
Exhibit 2. Description of Participating CPC Sites, by Project Year	6
Exhibit 3. Enrollment at PFS-funded CPC sites, by year	11
Exhibit 4. Children Attending PFS-funded CPC Sites, by Cohort and Exclusion Criteria	12
Exhibit 5. Intervention Sample Characteristics During their Preschool Year, by Cohort.....	13
Exhibit 6. Annual attrition, by Cohort and comparison group	22
Exhibit 7. Percent of Children Meeting Kindergarten Readiness Criteria, Across Domains, by Cohort	24
Exhibit 8. Percent of Children Meeting Kindergarten Readiness Criteria, by Cohort and Domain.....	24
Exhibit 9. Percent of Children in Special Education, by Cohort and Grade	26

Background

The Child-Parent Center (CPC) model, one of the longest running early childhood intervention models in the United States, has produced some of the most robust long-term academic and social outcomes for children (Reynolds, 2000; Reynolds & Temple, 2008). Beginning in January 2012, as part of a U.S. Department of Education Investing in Innovation (i3) grant to the University of Minnesota, the city of Chicago and Chicago Public Schools (CPS) received funding to (1) increase the number of children who could attend existing CPC sites and (2) increase the availability of CPC programs by adding 16 new sites.

Pay for Success (PFS), previously referred to in this evaluation as a Social Impact Bond (SIB), is a funding mechanism whereby private business and philanthropic partners purchase Social Impact Bonds (SIBs) to support public programs. Government entities, such as a state's Department of Education, pay investors only when a program meets its pre-determined outcomes. The investors bear the full risk of the investment and if a program fails to meet its goals, taxpayers owe nothing. To this end, PFS initiatives typically have an independent evaluator to help determine whether the outcomes have been realized and outcome payments to the private investor need to be made. Beginning in 2014–15, the IFF Pay for Success project funded additional CPC preschool slots at six CPS schools. In 2015–16, three sites (identified by CPS and approved by the city of Chicago) were added to the PFS project. Nine sites remained in 2016-17 and 2017-18. SRI International (SRI) was contracted to conduct the evaluation of the child outcomes for this project, referred to as the "PFS-CPC project." The project served four cohorts of preschool children across the nine sites over four school years—Cohort 1: 2014–15, Cohort 2: 2015–16, Cohort 3: 2016–17, and Cohort 4: 2017–18.

This fifth SRI project evaluation report describes third grade special education placement outcomes of Cohort 1 children, second grade special education placement outcomes of Cohort 2, and first grade special education outcomes for Cohort 3 and kindergarten special education outcomes for Cohort 4. The report also includes third grade literacy outcomes for Cohort 1. The report also includes all outcomes reported to date.

The report begins with a description of the CPC program and its expansion efforts using PFS funding, including evidence about the impacts of the CPC program model on children’s school readiness and school achievement at the beginning of the project. Next, we describe the approach to the PFS-CPC program evaluation. The last section presents the results of the evaluation for Cohorts 1, 2, 3, and 4 through the 2018-19 school year. We then conclude this report with a brief discussion of how the results fit in the larger context of research.

CPC Program Model

CPC Model Description

The CPC program model is designed to promote school readiness, parent involvement, and early learning that, in turn, will translate into long-term academic achievement, higher graduation rates, and career success for CPC students. The CPC model is unique in that it is designed to (1) provide full- or part-time high-quality preschool experiences for 3- and 4-year-old children and (2) combine those educational experiences with family support services and parent engagement activities. CPC programs deliver synergized services for children and their families from preschool through third grade. Indeed, the CPCs emphasize the provision of comprehensive services and parental involvement—program features that are considered to be strongly associated with program quality (Reynolds & Hayakawa, 2011; Reynolds, Magnuson, & Ou, 2010). A typical CPC site has the components listed in Exhibit 1.

The CPC program model components are explained fully at <https://humancapitalrc.org/midwest-cpc/cpc-resources> (Human Capital Research Collaborative, 2015). For this report, the components listed in Exhibit 1 were taken from the draft evaluation plan in the PFS-CPC expansion agreement (see Appendix A, pp. 9–11). Note that the CPC model as conceptualized in the current PFS expansion project focuses primarily on providing high-quality preschool education, engaging parents in their child’s education through a parent resource teacher (PRT) at the child’s preschool, and promoting continuity and stability from pre-K through the primary grades. Because the focus for the PFS-CPC project is on preschool programming, SRI’s evaluation is designed to measure the impact of the preschool components on children’s short- and long-term outcomes.

Exhibit 1. CPC Program Model Components

Effective Learning Experiences

- Offer pre-K classes that are limited to 34 children for half-day classrooms (two sessions of 17 children each) and have a minimum of two teaching staff. Full-day classrooms, if available, will be limited to 20 children per session.
- Provide highly qualified educational staff who will deliver the classroom instruction and parent engagement activities. For example, classroom teachers are certified with a bachelor's degree (or higher). Overall, program staff must adhere to the requirements set forth by the CPS Talent office, in accordance with collective bargaining unit agreements, and state regulations. Any changes in CPS education and certification requirements will be complied with.
- Use data to drive instruction by effectively documenting the organization and implementation of instructional practices to monitor quality and adherence to the program, which is completed by all program staff where appropriate.
- Program staff meet with parents over the course of each school year to review their child's progress and discuss parent program opportunities with the Parent Resource Teacher (PRT).

Aligned Curriculum

- Implement a CPS District curriculum and formative assessment that are aligned to standards, domains of learning, assessments, and learning activities.
- Collaborate with the PRT and classroom teachers to ensure that opportunities to engage families in student learning are available, appropriate, and aligned to the program and parents' needs.
- CPS and, most specifically, the Office of Early Childhood Education provide meaningful professional development and ongoing coaching and feedback for teachers, aides, and other staff members that facilitates high-quality instructional practices.

Parent Involvement and Engagement

- Engage a PRT and School-Community Representative to work closely with the head teacher and liaisons to maintain a consistently supportive parent program.
- Encourage parents to sign a CPC school-home agreement at the start of the school year outlining a plan for fostering learning at home and participating in CPC activities.
- Offer and engage families in monthly activities. PRTs create and distribute a monthly parent involvement calendar and conduct parent/teacher conferences over the year to review progress in the parent program.
- Provide a resource room dedicated to parent and family activities through kindergarten when possible.
- Provide culturally responsive learning opportunities for families that provide flexibility for families' needs and schedules.

Collaborative Leadership Team

- Engage a program leadership team that includes the head teacher, PRT, and school-community representative.
- Meet regularly, under the direction of the principal, to discuss operations and best practices within the CPC.
- Meet regularly, under the direction of the Office of Early Childhood Education (OECE) management team, with staff from across sites to share challenges, experiences, and best practices and make frequent on-site visits to monitor the quality and effectiveness of the program.
- Establish meaningful partnerships with community providers to strengthen service delivery and enlist local universities in training opportunities.

Exhibit 1. CPC Program Model Components (concluded)

Continuity and Stability

- CPC pre-K classrooms are collocated in the same building as kindergarten classrooms, when possible, to promote familiarity and integration for students as they transition to kindergarten.
- Provide a structure of communication, planning, and joint activities under the direction of the principal, leadership team, and OECE management team from pre-K through the primary grades.
- Provide a part-time kindergarten aide when funding is available to support the transition into kindergarten.

Professional Development System

- Offer ongoing professional development opportunities on current trends and needs in early childhood education classrooms, through the OECE and the CPC leadership teams, including topics such as quality curriculum and instruction, data-driven instruction, learning environment, social and emotional needs, and parent engagement.
- Meet regularly and create professional learning communities to review ways to support instruction in the classroom and with other teachers.

Source: Adapted from Chicago Child-Parent Center Social Impact Bond Evaluation Plan, dated December 2, 2014 (in Appendix A).

Expected Outcomes from the CPC Program Model

SCHOOL READINESS

Previous research on the CPC program showed significant positive effects on children's kindergarten readiness (Reynolds, 1995; Reynolds, Temple, Robertson, & Mann, 2002). Examination of a more recent cohort of CPC participants indicated that they had significantly higher scores on a measure of language proficiency at the end of the program than children enrolled in other publicly funded preschool programs (Reynolds, 2002). More recently, Reynolds and colleagues reported that CPC participants are more likely to meet kindergarten readiness standards in four of six educational focus areas, or 'domains' on a teacher-rated measure (70%) compared with preschool children in the school district who did not attend CPC preschool classrooms (52%) (Reynolds, Richardson, Hayakawa, Englund, & Ou, 2016).

THIRD-GRADE READING AND LITERACY

The Chicago Longitudinal Study (CLS) followed children over time using administrative records to examine attendance, achievement, and graduation rates in CPC participants compared with children who did not attend CPC preschool. One study found a significant positive impact on third-grade reading achievement for pre-K to third-grade participants (.53 standard deviation) compared with participants who attended CPC only for pre-K and kindergarten (Reynolds, 1994). Smaller studies of

high-quality preschool interventions have found similar impacts on later school achievement compared with a no-preschool Comparison group (e.g., the Abecedarian study: Campbell, Ramey, Pungello, Sparling, & Miller-Johnson, 2002; Perry preschool project: Belfield, Nores, Barnett, & Schweinhart, 2006).

REDUCED SPECIAL EDUCATION USE

The long-term CLS study showed that extended CPC participation (defined as 4 to 6 years) resulted in reductions in the use of special education. Among children 6 to 18 years old, CPC participants had an average rate of special education placement of 14.4% compared with 24.6% for children in the comparison group (who did not attend CPC preschool), indicating that CPC participants had a 41% lower rate of special education placement (Reynolds, Temple, & Ou, 2003). This finding is consistent with another analysis using the CLS sample that compared the average rates of special education placement over time for children who had attended a CPC preschool with those of children who attended a full-day non-CPC kindergarten classroom: special education placement rates of 12.5% and 18.4%, respectively (Conyers, Reynolds, & Ou, 2003). It is noteworthy that these estimates average special education placement rates over a wide age range extending beyond the early school years. A more recent study of North Carolina's state-funded preschool program used statewide population data from 1995 to 2010 to show that third-grade special education rates across multiple cohorts were reduced by as much as 39% for children who participated in the preschool program, even after taking into account a variety of child and family risk factors, types of special education categories, and funding levels that varied by year (Muschkin, Ladd, & Dodge, 2015). Other reviews of a variety of preschool program models also reported reductions in special education placement as one of the many cost savings results from participation in high-quality preschool programs like the CPC model (Karoly et al., 1998; Lynch, 2007).

In summary, positive impacts on kindergarten readiness, third-grade reading achievement, and special education placements have been cited extensively to demonstrate the short- and long-term benefits for the individual child and savings for society that come from investing in early childhood education. These studies, often including older samples of children followed extensively, were used as the basis for identifying the selected outcomes in the current evaluation and for calculating the repayments that will be made in the Chicago PFS-CPC project.

Chicago PFS Project (PFS-CPC Project)

During 2017–18, the PFS expansion of the CPC model involved funding for part-day or full-day CPC preschool at nine sites. Exhibit 2 indicates the year each site began receiving PFS funding and whether the site expanded an existing CPC program or began implementing the CPC program for the first time (i.e., a “new” CPC site).¹ In 2017–18, the PFS funding provided preschool to an additional 718 3- and 4-year-olds across the nine sites (see Exhibit 3). The funding paid for the expansion of classroom programming at each site as well as enhanced resources and instructional materials to implement the CPC model. The CPC program typically serves both 3- and 4-year-olds, sometimes in mixed-age classrooms. Thus, the investor funding was used to provide CPC preschool and enhanced services to both 3- and 4-year-olds.

Exhibit 2. Description of Participating CPC Sites, by Project Year

Site	Year 1 2014–15	Year 2 2015–16	Year 3 2016–17	Year 4 2017–18
De Diego	Expanded	Continued	Continued	Continued
Peck	Expanded	Continued	Continued	Continued
Melody	Expanded	Continued	Continued	Continued
Fiske	Expanded	Continued	Continued	Continued
Thomas	Expanded	Continued	Continued	Continued
Hanson Park	New ²	Continued	Continued	Continued
Edwards		Expanded	Continued	Continued
Tonti		New	Continued	Continued
Davis		New	Continued	Continued

Note: “Expanded” indicates that a site used PFS funding in that year to expand an existing CPC program, “New” indicates a program received PFS funding to begin implementing a CPC program for the first time, and “Continued” indicates that a site continued to receive PFS funding for an additional year.

The project administrators anticipated that four cohorts of children would be served across the nine sites, identified by the school year in which children begin preschool

¹ Three of the six sites in Year 1 had been providing CPC services since 2012, at the start of the i3 federal grant, and two had been providing CPC services since 2013, when the original sites from the 1970s were merged with the current site. For the three additional sites in year 2, two were new to providing CPC services and one had been providing CPC services since 2012.

² This site only operated for half of the year due to delays in hiring. Thus, the first full year of implementing CPC was 2015–16.

(Cohort 1: 2014–15, Cohort 2: 2015–16, Cohort 3: 2016–17, Cohort 4: 2017-18) (see Appendix B for grade levels of children in the four cohorts across years).

Evaluation Design

Because government pays investors only when outcomes are achieved, PFS initiatives typically have an independent evaluator to help determine whether the outcomes have been realized. SRI is conducting the independent evaluation of the outcomes of the PFS-CPC expansion project. The SRI evaluation team developed the evaluation methodology building on a draft evaluation design written by a team that included the Harvard Social Impact Bonds Technical Assistance Lab. The project also has an oversight committee of early education and research experts. The evaluation team has been charged with independently documenting the outcomes-based performance measures of the initiative. This kind of evaluation is not intended to test the impact of the CPC model against other preschool models; rather it is tracking the outcomes of the participating children against specific outcome metrics. The evaluation is addressing three performance-based outcome questions:

1. What is the rate of kindergarten readiness in children participating in the PFS-CPC sites as defined by performance on the Teaching Strategies (TS) *GOLD*[™] instrument (completed by teachers in the spring of the preschool before a child enters kindergarten)?
2. What is the rate of third-grade literacy as defined by performance in meeting or exceeding grade-level performance on the state- or district-administered third-grade assessment in reading?
3. What is the rate at which students are identified with special education needs and placed in special education services (starting in kindergarten) compared with a matched comparison group of children?

Kindergarten readiness was measured in the spring of preschool for CPC participants (as described below), and third-grade literacy was measured in the spring of third grade after the administration of required state achievement tests. SRI began measuring special education placement in kindergarten and continued through spring

of 2019 (in spring 2019, Cohort 1 reached the third grade; Cohort 2, second grade; Cohort 3, first grade; and Cohort 4, kindergarten).

For the evaluation of the PFS-CPC project, SRI is using two designs to track the primary outcomes: a descriptive study for the *kindergarten readiness* and *third-grade literacy* outcomes³ and a quasi-experimental design for the *special education outcomes from first to fourth grades* (see analysis approach for further information). Specifically, for the kindergarten readiness and third-grade literacy outcomes, there will be no comparison group for evaluating the outcomes and calculating the subsequent repayment. Evaluation of these two primary outcomes will be based on the intervention group only, and payments will be calculated using outcomes relative to national standards. In the planning phase, it was determined that both the kindergarten readiness and literacy outcomes had normative information so that children's performance on these measures could be used to identify whether they were performing at or above normative trends for comparable same-age peers. The decision was to use this kind of standard rather than compare performance with that of a comparison group of children. In addition, the kindergarten readiness data are not available for children with no preschool experience, given that the kindergarten readiness measure is collected during the spring of pre-K in Chicago Public Schools and only for children who attended preschool.

For special education outcomes (kindergarten to fourth grades), children are identified as receiving the intervention (i.e., attending a CPC preschool classroom) in the year they are in preschool and then are matched to children with similar demographic characteristics but who did not attend any preschool (CPC or otherwise) in CPS. This no pre-K comparison group was identified when the children were in kindergarten for each of the four cohorts. That is, SRI created a no pre-K comparison group for each cohort of intervention children using propensity score weighting processes.

³ The approach used here is sometimes referred to as the "rate card" approach where success payments are made on a per-student basis.

Analysis Approach

CPC INTERVENTION SAMPLE INCLUDED IN ANALYSIS

Children were included in the intervention cohorts if they attended one of the PFS-CPC sites that was fully implementing the CPC program during their preschool year,⁴ were enrolled in either a full- or half-day pre-K classroom, were not identified as having a severe disability, were income eligible (i.e., eligible to receive free or reduced-price lunch), and were at least 4 years old in September of their preschool year. Additionally, a child needed to have attended a CPC pre-K classroom for at least 66% of the days (not consecutively) in a given school year—a percentage considered a sufficient amount or dose of the intervention to affect child outcomes.

The project is based on the hypothesis that high-quality early childhood education will prevent or reduce a future need for special education services for children considered at-risk for developing delays or mild disabilities. As such, children diagnosed with severe disabilities were excluded from the project. Early childhood education and intervention also may reduce the need for children with mild delays or speech and language impairments in preschool from needing additional special education services in kindergarten and beyond. The project does not expect to prevent children with severe disabilities or needs from receiving special education services. Children were categorized as having no disability, a mild disability, or severe disability based on *a priori* decisions of the evaluation team in the planning and evaluation design phase. A severe disability could include autism, specific learning disability, deaf-blindness, deafness, hearing impairment, orthopedic impairment, other health impairment, traumatic brain injury, visual impairment, and multiple disabilities. A mild disability could include developmental delay, speech and language impairment, specific learning disability, and educational support accommodations or modifications for children with no other disability (mild or severe). Additionally, children were excluded from the intervention cohort if they were in a separate classroom for students with special education needs. Finally, children were excluded if they were identified with special needs and already had an Individualized Education Plan (IEP) prior to starting their preschool year at age 4 and were specifically assigned to one of the CPC sites because the site had blended classrooms (i.e., based on a CPS district

⁴ There were five sites for Cohort 1 and nine sites for Cohorts 2, 3, and 4.

policy, some school sites had general education classrooms with additional supports to better serve children with IEPs, which are referred to as “blended classrooms”).

Each cohort includes children from all PFS-funded sites that were providing the CPC model to 3- and 4-year-olds during that cohort’s 4-year-old preschool year. Inclusion of all eligible 4-year-olds in this group increases the sample size for the study to provide a more reliable and valid assessment of kindergarten readiness at the CPC sites. All the children across all classrooms received the full CPC model. That is, the experience of all 4-year-olds enrolled in these CPCs is similar, with a common curriculum, professional development, and parent engagement aligned through monthly Collaborative Leadership Training by all CPCs, including high-quality preschool and family support services and parent engagement activities. Thus, the evaluation does not distinguish between preschool slots funded by PFS versus other CPC funding sources.

Each PFS-CPC cohort was defined as meeting the eligibility criteria above and these children became the cohort to be tracked for outcomes in kindergarten and in later grades. Each cohort also was used to identify a matched comparison group of children in kindergarten for comparing special education outcomes at the end of kindergarten and in later grades.

Exhibit 3 shows the enrollment information for CPC preschool slots at PFS-funded sites, by year. At the end of Year 4 (the 2017–18 school year), administrative enrollment data showed that 1,287 3- and 4-year-old children were attending preschool at these nine sites (480 3-year-olds; 807 4-year-olds). PFS expansion funding covered the costs of providing CPC preschool for 718 of these 1,287 children.

Exhibit 3. Enrollment at PFS-funded CPC sites, by year

	Year 1 Enrollment (2014–15)	Year 2 Enrollment (2015–16)	Year 3 Enrollment (2016–17)	Year 4 Enrollment (2017–18)
Number of sites	6	9	9	9
Total enrollment	653	1,378	1,309	1,287
3-year-olds	267	537	502	480
4-year-olds	386	841	807	807
PFS-funded seats	374	782	798	718 ⁵

For Cohort 4 intervention group, SRI requested a data export that included all students ever enrolled in grade PK (the CPS designation for 4-year-olds in preschool) in the PFS-funded CPC sites at any time in the 2017–18 school year. Overall, 733 PK students were ever enrolled at one of the nine CPC sites during 2017–18.⁶ Across the total sample of 733 PK children, 592 or 77% met all the eligibility criteria. Exhibit 4 indicates the exclusions from the original sample of PK children ever enrolled in one of the sites that resulted in the final sample of children included in the analysis for each cohort.

⁵ The method of calculating funded slots in Year 4 may differ from previous years due to staffing changes in the district.

⁶ The number of children ever enrolled is different from enrollment estimates at any given point in the year. As children left a CPC site, new children were enrolled. The 773 includes all children ever enrolled during the 2017–18 year. Based on enrollment in May/June 2018, CPS reported that 807 4-year-old children were enrolled at the nine CPC sites at the end of the year.

Exhibit 4. Children Attending PFS-funded CPC Sites, by Cohort and Exclusion Criteria

	Cohort 1 2014–15 (n or %)	Cohort 2 2015–16 (n or %)	Cohort 3 2016–17 (n or %)	Cohort 4 2017–18 (n or %)
Number of sites included	5	9	9	9
Number of four-year old children attending these sites (grade PK)	449	1,004	818	773
Number of four-year old children meeting eligibility criteria	313⁷	654⁸	637⁹	592¹⁰
Reason for exclusion	%	%	%	%
Did not attend 66% of days	21%	21%	7%	7%
Severe disability and/or enrolled in separate special education classroom	3%	2%	3%	2%
Had an IEP prior to PK year	3%	7%	8%	8%
Not eligible for free or reduced-price lunch or insufficient documentation	2%	1%	4%	5%
Under 4 years old in September of PK year	<1%	4%	1%	<1%

Cohort 1 (2014–15) included 313 children, Cohort 2 (2015–16) included 654 children, Cohort 3 (2016–17) included 637 children, and Cohort 4 (2017–18) included 592 children. The demographic characteristics of all four Cohorts are described in Exhibit 5. Cohorts 1, 2, 3, and 4 are also described in previous reports (Gaylor, Ferguson, McCracken, Wei, & Spiker, 2017; Gaylor et al., 2016). The final samples of children who were included were similar in many ways to the children who did not meet the eligibility criteria with a few exceptions. The children who were included in each intervention cohort were more likely to be Hispanic and more likely to speak Spanish at home compared with the children who were excluded from that cohort.

⁷ Of the 313 children selected for Cohort 1 in their PK year (2014–15), 289 were enrolled in CPS on the 20th day of school during their kindergarten year, for a Cohort 1 kindergarten retention rate of 92.33%.

⁸ Of the 654 children selected for Cohort 2 in their PK year (2015–16), 619 were enrolled in CPS on the 20th day of school during their kindergarten year, for a Cohort 2 kindergarten retention rate of 94.65%.

⁹ Of the 637 children selected for Cohort 3 in their PK year (2016–17), 593 were enrolled in CPS on the 20th day of school during their kindergarten year, for a Cohort 3 kindergarten retention rate of 93.09%.

¹⁰ Of the 592 children selected for Cohort 4 in their PK year (2017–18), 556 were enrolled in CPS on the 20th day of school during their kindergarten year, for a Cohort 4 kindergarten retention rate of 93.92%.

Exhibit 5. Intervention Sample Characteristics During their Preschool Year, by Cohort

Characteristic	Cohort 1 2014–15 (n = 313) (percent)	Cohort 2 2015–16 (n = 654) (percent)	Cohort 3 2016–17 (n = 637) (percent)	Cohort 4 2017–18 (n = 592) (percent)
Male	50	48	47	50
Hispanic	67	79	82	84
African-American	30	17	15	14
Caucasian	1	2	2	1
Other ethnicity	1	<1	1	1
Designated as English Language Learner (ELL)	44	57	60	57
Identified mild developmental delay or disability	4	4	3	5
Enrolled in full-day Pre-K classrooms	37	40	42 ¹¹	36 ¹²

ANALYZING IMPACT ON KINDERGARTEN READINESS

SRI examined kindergarten readiness in school year’s 2014-15 through 2017-18 using Teaching Strategies (TS) *GOLD*TM scores from the spring before the child entered kindergarten.¹³ TS *GOLD* is a teacher-reported measure of young children’s skills across six developmental domains: literacy, language, mathematics, cognitive development, socio-emotional well-being, and physical health. We are using this measure because it was the only available child assessment that CPS routinely collects and was therefore selected as the measure of kindergarten readiness by the PFS planning team.¹⁴ It is used routinely in the CPS preschool programs, and there was no alternative CPS-wide measure of kindergarten readiness that is completed as children are entering kindergarten in the fall of the school year at the time of the PFS implementation design.

¹¹ Half- or full-day status was unknown for 7% of the 637 students.

¹² Half- or full-day status was unknown for 9% of the 592 students.

¹³ The TS *GOLD* assessment was developed to monitor the skills of children attending a child care or preschool program so teachers can adjust their instructional strategies depending on the children’s progress on a variety of skills and behaviors. TS *GOLD*TM was not developed as a measure of kindergarten readiness.

¹⁴ The methodology involved in most SIB projects relies on use of available administrative data rather than additional data collection to evaluate outcomes.

Calculating Impact on Kindergarten Readiness

As described, we calculated the impact of CPC on kindergarten readiness by comparing children's performance on the measure with national norms. We selected this approach for two reasons. First, adequate normative data enables us to identify whether children in the sample were performing at or above a widely accepted standard. Second, creating an appropriate comparison group within CPS was not possible; kindergarten readiness data are not available for children without preschool experience (our comparison group), given that the kindergarten readiness measure is collected during the spring of pre-K in Chicago Public Schools.

The metric for kindergarten readiness is the percentage of children who are performing at or above national trends across at least five of these six domains.¹⁵ A child is determined to be ready for kindergarten if he or she is rated by the teacher as demonstrating levels of skill or knowledge that are expected for a child at a particular age; the reference point for such expectations come from the observed abilities of other children from a representative sample of same-aged peers in the United States. We categorized children as kindergarten ready on each domain by the criterion of meeting or exceeding the 50% percentile on the standard score for that domain using scores from the most recently published technical manual (Lambert, Kim, & Burts, 2014a). Then, we calculated the percentage of children who met this criterion on five of six domains.¹⁶

For Cohorts 1, 2, and 3, TS *GOLD*TM observation data were administered (meaning the same items made up each domain) and scored the same way using a system for converting raw rating scores to standard scores. Between Cohort 3 and Cohort 4, the developers changed the TS *GOLD*TM system from a Birth to Kindergarten assessment to a Birth to Third Grade assessment. This change to expand the age range of this

¹⁵ No data are available on which domains of the TS *GOLD* assessment to use to reliably and validly determine kindergarten readiness. The decision to define kindergarten readiness as performing at or above national trends on five of six domains (and not four of six) aligns with the National Research Council's definition of school readiness, which includes age-level skills across multiple domains (National Research Council, 2008). The threshold of five of six domains also takes into account that a child may not meet a standard for all six domains, especially in the spring of preschool, as these skills are emerging during this time period.

¹⁶ Teacher-reported assessments have some unknown sources of variability, and the *GOLD* assessment is no different. Research on the *GOLD* assessment indicates that between 17% and 25% of the variance in scale scores is accounted for by unmeasured differences between classroom and teachers, including rater effects (Lambert et al., 2014). At this time, there is no consensus on how to calculate kindergarten readiness using *GOLD* assessment scores. Thus, we continue to use the a priori definition and benchmark.

assessment tool involved adding items to some domains to represent the full-range of skills and abilities that would be observed in children from birth to 9 years of age. It also involved re-scaling the standard scores to represent the full range of raw and standard scores across the expanded age range. Unfortunately, these changes in the item pool and the scoring mean that for Cohort 4 children, we have no way to compare their TS *GOLD*[™] scores to scores derived earlier for the other cohorts. Thus, the new version of the assessment measure is still under development and does not have the same reference points to allow for reliable statements about children's kindergarten readiness skills.

ANALYZING IMPACT ON THIRD GRADE READING

As part of the original evaluation plan, SRI was charged with calculating the number (and percentage) of students classified as “reading at or above grade level” which was defined at the time as scoring at or above the 25th percentile on the English Language Arts/Literacy portions of the state administered achievement test, PARCC (Partnership for Assessment of Readiness for College and Careers). This criteria was set prior to SRI's involvement but is supported by previous research (Lestnick, Goerge, Smithgall, & Gwynne, 2010).¹⁷ At that time, it was decided that if the PARCC was no longer being used, SRI would follow the same procedures on the assessment that is administered in the district for each cohort (see evaluation plan, Appendix A, p. 13):

“At the time of drafting this analysis, the PARCC test has yet to be officially implemented in CPS schools. Given the uncertainty of performance on this test and how its outcomes will compare to past tests taken by CPS students, the evaluator may suggest amendments to the definition of reading “on grade level” that could include utilizing a different test or metric. Any modifications must be made prior to the first cohort starting Third Grade, and must be approved by CPS, the City, the Project Coordinator, and Approved by the Lender Committee.”

Thus, during the 2018-19 school year, the evaluation team reviewed the availability of third grade testing options for calculating literacy outcomes for the first cohort of participants to reach third grade. As a reminder, as part of the PFS-CPC project, SRI

¹⁷ More information about the PARCC can be found at <http://www.isbe.net/assessment/parcc.htm>.

is tasked with conducting a descriptive study¹⁸ to measure the number of students attending PFS-funded CPC sites who meet or exceed a particular performance standard on a third grade literacy assessment. Repayments to the lenders will be calculated accordingly using the data. Because the available measures of third grade literacy changed since the original PFS-CPC contract and evaluation design were developed, SRI discussed with IFF and relevant parties from CPS, the city, and the lenders which measure to use in the evaluation. Below we describe the two available options and the rationale for the final selection.

The two available measures of third grade literacy during the 2018-19 school year were (a) the Illinois Assessment of Readiness ([IAR](#))¹⁹ and (b) the NWEA MAP ELA assessment ([NWEA MAP](#)) (first implemented in the 2013-14 school year). SRI selected the MAP as the most appropriate choice for this evaluation because 1) it is a standardized test used throughout the U.S. and provides nationally normed percentile ranks of a child's score in comparison to other children of the same grade level in the U.S., and 2) it aligns best with the approach used for the kindergarten readiness measure in the PFS-CPC evaluation which also used a national norm-referenced assessment and scores [i.e., Teaching Strategies (TS) GOLD™]. For the Teaching Strategies (TS) GOLD™ measure, a threshold was established such that students needed to perform at or above the 50th percentile rank on that measure in 5 of 6 domains.

NWEA MAP is a norm-referenced test and can be used to identify whether a student performed at or above normative trends for comparable grade-level peers nationwide. NWEA MAP assessments provide a RIT score²⁰, percentile rank, and growth percentile score for each student, which enables the evaluation team to categorize whether a student is at or above the 25th percentile, which was established as the criterion in the evaluation plan and is typically used to identify a reading difficulty.

¹⁸ Please note that similar to the kindergarten readiness outcome, no comparison group will be available for evaluating third-grade literacy outcome.

¹⁹ The IAR is a criterion-referenced, Illinois state assessment that provides proficiency level scores ranging from 1 to 5 with level 1 for students who did not yet meet expectations and level 5 for students who exceeded expectations. Students performing at levels 4 or 5 on IAR are considered proficient and ready for the next grade level. We did not select the IAR for the PFS-CPC evaluation because these proficiency levels are not directly comparable to national standards.

²⁰ RIT scores (short for Rasch Unit) is an estimation of a student's instructional level and also measures student progress or growth in school.

ANALYZING IMPACT ON SPECIAL EDUCATION PLACEMENT

Special education placement was determined using data on children’s disability designation at any time during the child’s kindergarten year. Children were classified into three categories: children receiving special education services for mild disability, children receiving special education services for severe disability, and children not receiving special education services (no IEP). Recall that our hypothesis is that high-quality preschool via the CPC program will decrease the chances that children who are at risk will need special education services in the future. Because we are not trying to prevent children with severe disabilities from receiving the special education services they need, we restricted our definition of special education outcomes to children who needed special education for mild delays or disabilities defined as those children who had an IEP for the following: speech and language issue (S/L), developmental delay (DD), and specific learning disabilities (SLD), which is the only information available in the administrative dataset describing the type and severity of disability. This helps avoid the perverse incentive of withholding special education services from children with severe disabilities. Below we report the cumulative findings for special education placement outcomes for cohorts 1, 2, and 3 and their peers (i.e., the comparison group of children who did not attend any CPS preschool).

The effect size of the impact on special education placement for Cohort 1 was calculated using the risk difference approach. The equation is the following:

$$ES_{i,t} = SPED_{C,i,t} - SPED_{T,i,t}$$

where $ES_{i,t}$ is the effect size for cohort i in year t , $SPED_{C,i,t}$ is equal to the average of a binary indicator of special education placement among the no pre-K comparison group for cohort i in year t , and $SPED_{T,i,t}$ is the average of a binary indicator of special education placement among the intervention group for cohort i in year t . The same calculation will be used for each cohort for each year through sixth grade as described below.

Based on conversations with special education experts and reviewing existing CPS data, the consensus by the planning committee is that the vast majority of children who have a disability will be identified by the end of sixth grade (Blackorby et al., 2010). As a result, after the sixth-grade effect size has been calculated, IFF (or the district) will average the effect size over the last 3 years (fourth, fifth, and sixth

grades) and lock in that average effect size for the purposes of calculating payments in grades 7 through 12. This lock-in rate will be calculated separately for each intervention cohort. SRI may propose changes to this lock-in methodology in the event that the team determines that it produces skewed results. Any modifications must be approved by CPS, the city of Chicago, the project coordinator, and approved by the lender committee.

SELECTING A COMPARISON GROUP

For the special education placement outcome, we conducted propensity score analysis to identify an appropriate comparison group that had not received CPS preschool in either school- or community-based settings. Propensity score methods are quasi-experimental approaches that were developed to approximate findings obtained from randomized controlled trials (Becker & Inchino, 2002). They have been used increasingly in analysis of observational data to reduce selection bias in estimating treatment, policy, or intervention effects when randomized controlled trials are not feasible or ethical (Rosenbaum & Rubin, 1983, 1984, 1985). In essence, propensity score methods help to identify a comparison group that mimics what might have been obtained using random assignment.

Initial selection

To create the initial comparison group for each cohort, we first restricted our data set to all kindergarteners in CPS who were 5 years old or older on September 30, were eligible for free and reduced-price lunch, did not attend preschool in CPS (either in a CPC or other CPS preschool classroom),²¹ and who were not attending kindergarten at a school with a CPC program.²²

For detail about the Cohort 1, 2 and 3 comparison group selections, see the previous reports for 2015–16, 2016–17, and 2017–18. For the Cohort 4 comparison group, our potential sample was approximately 8,208 or about one-third of the total number of

²¹ The evaluation team only had data about whether children had received preschool in CPC or other CPS-sponsored classrooms. As such, it is possible that some children in this comparison group may have participated in a preschool program such as Head Start outside of the district in a community-based setting.

²² The initial planning team had suggested also excluding children who were enrolled in charter schools, magnet or selective enrollment schools, and schools that serve exclusively a special education population. However, there are no elementary schools that serve exclusively a special education population. The evaluation team did not think it necessary to exclude children who attended charter or magnet schools because we did not have adequate information showing these were higher-performing schools than the “business as usual” elementary schools children could have attended.

children enrolled in kindergarten in CPS in 2018–19 (23,527). The sample was further reduced to 7,805 who had no missing data on any baseline covariates and outcome. We then applied propensity score analysis to identify a propensity score for each of the 7,805 children eligible to be in the comparison group. The propensity score is the predicted probability of being in an intervention based on a set of potentially confounding covariates (e.g., child and neighborhood background characteristics; see below for more detail) using logistic regression. The key advantage of using a propensity score is the ability to balance intervention and comparison groups on a large number of covariates by using a linear combination of covariates for a single score. Simply, the propensity score is a measure of how similar children from the comparison group are to the children in the intervention group on a large number of covariates.

We applied propensity score weighting (PSW) because this approach has the advantage of maximizing power by including all eligible children in the comparison group sample rather than only matched cases. It weights each comparison child by their propensity score, a measure of similarity between intervention and comparison on a large number of covariates; comparison children were weighted higher if they were more similar to intervention children and were weighted lower if they were to less similar to Intervention children.

The way PSW works is that each child is given a weight derived from logistic regression which represents how closely the child matches the intervention group; in this case, how well-matched they were on the selected child and neighborhood characteristics. The weight is not related to the outcome (special education status). This propensity score weighting approach adjusts for confounds using inverse propensity score estimators, as recommended by Curtis, Hammill, Eisenstein, Kramer, and Anstrom (2007), Hirano, Imbens, and Ridder (2003), and Rosenbaum and Rubin (1983). From the logistic regression model, we calculated a probability that the comparison child would be in the intervention group. The weight for intervention students was 1.0, and the weight for comparison students was equal to their propensity score transformed to an odds scale ($\pi/1-\pi$) (Harder, Stuart, & Anthony, 2010; Hirano et al., 2003).

In the PSW approach, children in the comparison group who are more like the intervention group children are weighted more heavily, and comparison group children less like the intervention group children get smaller weights. Outcome data for each child is given a numerical weight based on the child's baseline demographic characteristics. For example, if the intervention group has a higher proportion of Hispanic children and a lower proportion of White children than the comparison group, the Hispanic children in the comparison group will be weighted more and the white children in the comparison group will be weighted less.²³ The final comparison sample comprises all 7,805 children, weighted to closely match children in the intervention group.

An important aspect of estimating the propensity score is the selection of covariates. Researchers suggest that covariates that affect both intervention participation and outcomes should be included in the estimation of the propensity score (Caliendo & Kopeinig, 2008; Heckman, Ichimura, Smith, & Todd, 1998; Lechner, 2002; Ravallion, 2001). Covariates included in this study were selected based on findings from studies that have examined neighborhood effects associated with child outcomes (Harding, 2003; Root & Humphrey, 2014; Sampson, Sharkey, & Raudenbush, 2008; ten Bensel, Gibbs, & Lytle, 2015). Our covariates came from four data sources: school district data for the 2015-2016 school year, census tract data for 2013, community area public health data for 2009, and police district crime report data for 2010. The PSW did result in well-match intervention and comparison groups. Additional details about the PSW analysis for the Cohort 4 comparison group are contained in Appendix C.

Adjusting for attrition

The initial selection of the comparison groups for Cohorts 1, 2, and 3 were described in the Year 1, Year 2, and Year 3 reports (April 2017, March 2018, May 2019). In Year 5, Cohort 1 and their comparison group moved into third grade, Cohort 2 and their comparison group moved into second grade, Cohort 3 and their comparison group moved into first grade, Cohort 4 and their comparison group were in kindergarten, and our analyses of their kindergarten through third grade special

²³ This is an oversimplification because the approach actually uses all of the demographic data available to create the weight for each comparison child. That is, instead of matching on each of the covariates; children are given a weight that is based on the combination of all of the covariates.

education outcomes were adjusted to account for attrition²⁴ as students left the district after their previous school year. As shown in Exhibit 6, between 2017–18 and 2018–19, the attrition rates were lower in some cohorts than in their comparison groups, but higher in others; rates between the pairs were similar, with Cohort 1 and Cohort 2 both at 5% and their comparison groups at 6% and 7%, respectively, and Cohort 3 at 3% and their comparison group at just 1%. We observed higher attrition from preschool to Kindergarten for Cohort 4 (18%) compared with Cohorts 1 and 2 as they moved to Kindergarten (5% and 9%, respectively), but lower compared with Cohort 3 between preschool and kindergarten (20%). To make sure the groups were still similar after this attrition, we re-weighted the remaining children in each of the comparison groups to match the remaining children in each cohort on the selected child and neighborhood background characteristics. Additional details about the re-weighting analysis for the Cohort 1, 2, and 3 comparison groups are contained in Appendix C in the Year 4 report.

²⁴ Attrition here is defined as students not attending CPS in the expected grade level in a particular school year. Children are included if they are no longer enrolled in the school district or are enrolled in another grade.

Exhibit 6. Annual attrition, by Cohort and comparison group

Group	School Year 2014–15	School Year 2015–16	School Year 2016–17	School Year 2017–18	School Year 2018–19
Cohort 1 and their comparison group					
<i>Cohort 1</i>	Preschool <i>n</i> = 313 original sample	Kindergarten <i>n</i> = 297 5% attrition	First Grade <i>n</i> = 256 14% attrition	Second Grade <i>n</i> = 241 6% attrition	Third Grade <i>n</i> = 229 ²⁵ 5% attrition
<i>Comparison</i>	n/a	Kindergarten <i>n</i> = 9,445 original sample	First Grade <i>n</i> = 7,076 25% attrition	Second Grade <i>n</i> = 6,385 10% attrition	Third Grade <i>n</i> = 6,025 6% attrition
Cohort 2 and their comparison group					
<i>Cohort 2</i>	n/a	Preschool <i>n</i> = 654 original sample	Kindergarten <i>n</i> = 595 9% attrition	First Grade <i>n</i> = 562 6% attrition	Second Grade <i>n</i> = 536 ²⁶ 5% attrition
<i>Comparison</i>	n/a	n/a	Kindergarten <i>n</i> = 7,126 original sample	First Grade <i>n</i> = 6,331 11% attrition	Second Grade <i>n</i> = 5,911 7% attrition
Cohort 3 and their comparison group					
<i>Cohort 3</i>	n/a	n/a	Preschool <i>n</i> = 637 original sample	Kindergarten <i>n</i> = 506 20% attrition	First Grade <i>n</i> = 490 ²⁷ 3% attrition
<i>Comparison</i>	n/a	n/a	n/a	Kindergarten <i>n</i> = 7,890 original sample	First Grade <i>n</i> = 7,822 1% attrition
Cohort 4 and their comparison group					
<i>Cohort 4</i>	n/a	n/a	n/a	Preschool <i>n</i> = 592 original sample	Kindergarten <i>n</i> = 493 ²⁸ 17% attrition
<i>Comparison</i>	n/a	n/a	n/a	n/a	Kindergarten <i>n</i> = 7,805 original sample

Note: Attrition rates are calculated using the number of students enrolled in that grade at the end of the designated school year as the numerator and the number of students enrolled the previous year as the denominator.

²⁵ Of the 313 children selected for Cohort 1 in their PK year (2014–15), 229 were enrolled in CPS at the end of their third grade year, for a Cohort 1 second grade retention rate of 73.16%.

²⁶ Of the 654 children selected for Cohort 2 in their PK year (2015–16), 536 were enrolled in CPS at the end of their second grade year, for a Cohort 2 second grade retention rate of 81.96%.

²⁷ Of the 637 children selected for Cohort 3 in their PK year (2016–17), 490 were enrolled in CPS at the end of their first grade year, for a Cohort 3 first grade retention rate of 76.92%.

²⁸ Of the 592 children selected for Cohort 4 in their PK year (2017–18), 493 were enrolled in CPS at the end of their kindergarten year, for a Cohort 4 kindergarten retention rate of 83.28%.

Results

We first present the kindergarten readiness results for all cohorts and then the results for kindergarten special education placement outcomes for Cohort 4 children, first grade special education placement for Cohort 3 children, second grade special education placement for Cohort 2 children, and third grade special education placement for Cohort 1 children. Lastly, we present third grade reading outcome results for Cohort 1 children.

Kindergarten Readiness

Kindergarten readiness results for Cohorts 1, 2, and 3 are presented in Exhibits 7 and 8. The evaluation team was unable to calculate kindergarten readiness for Cohort 4. When the evaluation team analyzed the TS *GOLD*[™] Spring 2018 data for Cohort 4, we found that the percentage of children meeting the kindergarten readiness criterion was dramatically lower than the percentages from previous years. When investigating this difference, the team learned that the developers had changed the measure by adding items and modifying the algorithm that converts the raw scores to standard (scale) scores. These item and scoring differences mean that the evaluation team could not use an identical procedure for calculating the percentage of children meeting the kindergarten readiness criterion. That is, as described above, we categorized children as kindergarten ready on each domain by the criterion of meeting or exceeding the 50% percentile on the standard score for that domain using scores from the most recently published technical manual (Lambert et al., 2014a). Then, we calculated the percentage of children who met this criterion on five of six domains. The standard scores we received for Cohort 4 were not derived in the same way as the standard scores from the earlier cohorts; when we categorized children using the earlier 50% percentiles, we found much lower rates of kindergarten readiness. We also know that there were no large differences in other factors that could affect kindergarten readiness rates (e.g., no changes in this Cohort 4 on implementation of the CPC program model, other school or district characteristics, sample characteristics, teacher training on the use of the measure, etc.). Given these results, we have no data on kindergarten readiness to report for Cohort 4.

Exhibit 7. Percent of Children Meeting Kindergarten Readiness Criteria, Across Domains, by Cohort

Number of domains meeting or exceeding the 50th percentile	Cohort 1 2014–15 (percent)	Cohort 2 2015–16 (percent)	Cohort 3 2016–17 (percent)	Cohort 4 2017–18 (percent)
0	9%	14%	19%	not available
1	3%	11%	7%	not available
2	7%	11%	9%	not available
3	12%	11%	7%	not available
4	9%	11%	13%	not available
5	10%	19%	15%	not available
6	51%	23%	29%	not available
5 or 6	61%	42%	44%	not available

Exhibit 8. Percent of Children Meeting Kindergarten Readiness Criteria, by Cohort and Domain

Domain	Cohort 1 2014–15 (percent)	Cohort 2 2015–16 (percent)	Cohort 3 2016–17 (percent)	Cohort 4 2017–18 (percent)
Cognitive	82%	65%	65%	not available
Language	66%	50%	53%	not available
Literacy	75%	62%	60%	not available
Math	81%	72%	68%	not available
Physical	60%	30%	42%	not available
Social-emotional	79%	63%	61%	not available

Special Education Placement

COHORT 4 KINDERGARTEN SPECIAL EDUCATION PLACEMENT

After ensuring the two groups were equivalent on child and neighborhood characteristics and weighting appropriately, we examined the kindergarten special education placement rates for mild and moderate developmental delay or disability for Cohort 4 and its comparison group. The special education rate was lower in the intervention group than the comparison group (5.48% for Cohort 4 in kindergarten

and 7.81% for children in the comparison group in kindergarten). This is a difference of 2.33 percent.

COHORT 3 FIRST GRADE EDUCATION PLACEMENT

After ensuring the two groups were equivalent on child and neighborhood characteristics and weighting appropriately, we examined the first grade special education placement rates for mild and moderate developmental delay or disability for Cohort 3 and its comparison group. The special education rate was lower in the intervention group than the comparison group (4.08% for Cohort 3 in first grade and 8.60% for children in the comparison group in first grade). This is a difference of 4.52 percent.

COHORT 2 SECOND GRADE SPECIAL EDUCATION PLACEMENT

After ensuring the two groups were equivalent on child and neighborhood characteristics and weighting appropriately, we examined the second grade special education placement rates for mild and moderate developmental delay or disability for Cohort 2 and its comparison group. The special education rate was lower in the intervention group than the comparison group (5.78% for Cohort 2 in second grade and 8.71% for children in the comparison group in second grade). This is a difference of 2.93 percent.

COHORT 1 THIRD GRADE SPECIAL EDUCATION PLACEMENT

After ensuring the two groups were equivalent on child and neighborhood characteristics and weighting appropriately, we examined the third grade special education placement rates for mild and moderate developmental delay or disability for Cohort 1 and its comparison group. The special education rate was lower in the intervention group than the comparison group (8.73% for Cohort 1 in third grade and 10.34% for children in the comparison group in second grade). This is a difference of 1.61 percent.

Exhibit 9. Percent of Children in Special Education, by Cohort and Grade

Group	Kindergarten (GK)	First Grade (G1)	Second Grade (G2)	Third Grade (G3)
Cohort 1 (PK n = 313)	2015-16, n = 297 4.38%	2016-17, n = 256 3.13%	2017-18, n = 241 4.15%	2018-19, n = 229 8.73%
Comparison	2015-16, n = 9,445 4.94%	2016-17, n = 7,076 6.17%	2017-18, n = 6,385 7.93%	2018-19, n = 6,025 10.34%
<i>Difference</i>	-0.56%	-3.04%	-3.78%	-1.61%
Cohort 2 (PK n = 654)	2016-17, n = 595 3.36%	2017-18, n = 562 4.27%	2018-19, n = 536 5.78%	TBD
Comparison	2016-17, n = 7,126 5.09%	2017-18, n = 6,331 7.20%	2018-19, n = 5,911 8.71%	TBD
<i>Difference</i>	-1.73%	-2.93%	-2.93%	TBD
Cohort 3 (PK n = 637)	2017-18, n = 506 3.95%	2018-19, n = 490 4.08%	TBD	TBD
Comparison	2017-18, n = 7,890 6.26%	2018-19, n = 7,822 8.60%	TBD	TBD
<i>Difference</i>	-2.31%	-4.52%	TBD	TBD
Cohort 4 (PK n = 592)	2018-19, n = 493 5.48%	TBD	TBD	TBD
Comparison	2018-19, n = 7,805 7.81%	TBD	TBD	TBD
<i>Difference</i>	-2.33%	TBD	TBD	TBD

TBD = to be determined

Third Grade Reading

COHORT 1 THIRD GRADE READING

Of the original 313 students in Cohort 1, 249 had third grade reading scores on the NWEA MAP assessment. Most (197 of 249 or 79.12%) met the 25th percentile or higher criterion that was established prior to calculating the percentage.

Discussion

Sociodemographic risk factors—the most extensively studied of which is poverty—are highly predictive of developmental trajectories. Children from low-socioeconomic-status (SES) households are less likely to enter kindergarten with the pre-academic and social skills needed to succeed and are more likely to require later special education services later (Brooks-Gunn, Rouse, & McLanahan, 2007; Hogan, Msall, Rogers, & Avery, 1997; Isaacs, 2012; Lee & Burkam, 2002). Early childhood programs potentially mitigate the risks endemic to children from disadvantaged backgrounds, with studies showing that the strongest positive short- and long-term outcomes result from intensive and comprehensive programs targeting low-income children (Burger, 2010; Institute for Research on Poverty, 1997; Reynolds et al., 2010). Indeed, prior studies have highlighted early childhood as a critical and sensitive period for the development of brain architecture and neurochemistry (e.g., Knudsen, Heckman, Cameron, & Shonkoff, 2006) and subsequent academic and socio-emotional well-being (Shonkoff & Phillips, 2000).

First implemented in Chicago in 1967, the CPC model has a long history of offering innovative, targeted approaches to school reform including a comprehensive system of educational and family support services during the preschool through third grade years for young children in low-income neighborhoods. The intervention promoted young children’s school success through language enrichment and intensive, mandatory parent involvement within a system of comprehensive support services for children and their families. The CPC model, integrated into the CPS system since its inception in 1967, has been systematically evaluated for its impact on child and family outcomes. A notable by-product of the CPC program’s efforts is the Chicago Longitudinal Study (CLS), which has supported researchers’ efforts to develop a deeper understanding of the “active” ingredients of early dual-generation

interventions and early childhood interventions more generally. The following are key relevant findings from analyses conducted on the CLS samples:

- Nearly half of children (44%) attending a CPC for one year were considered ready for kindergarten compared with 28% of children who had no preschool (unpublished data, A. Reynolds, personal communication, February 25, 2015).
- Children having one or two years of CPC preschool experience were less likely than those having no CPC preschool experience to have received special education throughout the elementary school years (Reynolds, 1995).

The expectation for the PFS-funded expansion of CPC to new sites in CPS and increasing the number of available CPC preschool slots at existing sites in CPS was based on previous research showing positive impacts on kindergarten readiness and school achievement, and reductions in special education placements over time. Other new CPC data also show positive impacts on kindergarten readiness (Reynolds et al., 2017). In a meta-analysis of 9 high-quality experimental and quasi-experimental studies conducted in the last 50 years researchers found that participation in ECE leads to reductions in special education placement, that is, about 8% fewer children need special education services when rates are averaged across studies (McCoy et al., 2017). Below we discuss the findings from the year 5 evaluation outcomes, including some of the limitations in interpreting the data.

Special Education Placement Findings

Based on the extensive CLS analyses as well as reviewing existing data for CPS in the years leading up to the PFS-CPC project, the PFS project was built on the hypothesis that high-quality preschool through the CPC program would help prevent or reduce the need for special education in the intervention group. In general, across cohorts and grades, we find that students who attended the PFS-CPC sites are less likely to need special education services (Exhibit 9). The difference ranges from 0.56% in kindergarten for Cohort 1 and its comparison group to 4.52% in first grade for Cohort 3 and its comparison group. For the latter finding, this is an 111% decrease for the intervention group relative to the comparison group. Other notable findings include an increase in the percentage of children receiving special education services across the grades (i.e., higher rates in older grades) and a narrowing of the gaps over by third grade. We suspect these trends reflect typical referral,

assessment, and placement policies and practices as students are confronted with more demanding educational material and have higher needs as the demands become greater. For example, in anticipation of and after third grade achievement testing (which is the beginning of required testing in most districts), students may be more likely to be referred for a learning disability or delay and receive an IEP. Many teachers and parents may adopt a “wait and see” approach in the early elementary grades before referring students for an assessment.

We also see larger differences favoring the intervention groups for Cohorts 2, 3, and 4 compared to Cohort 1 in kindergarten. This may suggest the quality or other positive characteristics of preschool programming offered in the later cohorts was somewhat better and/or had a greater impact on student’s learning and development.

We want to acknowledge a few data limitations. As described earlier, SRI was unable to identify with certainty children from the comparison group who may have received pre-K programming in non-CPS funded settings in order to exclude them from the comparison group sample. We did however exclude from the comparison group all children who had a CPS identification number when they entered kindergarten, which serves as an indication that they attended a CPS preschool program and/or were receiving services from the school district. In addition, with the propensity score weighting procedure, we used a large comparison sample to reduce the impact of possible preschool attendance on our estimation of comparison group outcomes. That is, there is likely some proportion of children in the comparison group samples who did attend some type of public preschool program (e.g., Head Start) or private child care program using child care subsidies and/or tuition. It is not clear how inclusion of some children with preschool experience in the comparison group is affecting the special education rate in that group. We also want to note that in the original design of this PFS project using the CPC model, the planning team decided that the targeted population was high risk children but not necessarily those with identified disabilities prior to participation in preschool. Therefore, children who were identified prior to enrollment in PreK were excluded from the sample that is tracked over time.

Finally, the evaluation team used disability categories to define who had a mild versus a severe disability, based on the assumption that the intervention was only

targeting prevention of the need for later special education services for children at-risk of a mild disability. Because the disability categories do not provide information on the severity of a child's disability or delay, the team used additional information about placement in special day classes and blended classrooms to identify children as having more "severe" delays or disabilities. While the information sources used are proxies for identifying those with "severe" delays or disabilities who were excluded from the analyses, rather than direct measurements of the severity of a child's disability, they were agreed upon with input from CPS as appropriate proxies to identify severity. We recognize that some children with developmental delay may eventually develop a more severe disorder and need services as could some of the children with mild speech and language delays and specific learning disabilities.

To put the findings about rates of special education placement for CPC cohorts and comparison group samples into context, we searched for comparable data and organized our search around three questions. First, to what extent are the findings similar to those reported in the Chicago Longitudinal Study of the earlier CPC cohort? Second, how do kindergarten special education placement rates for the intervention and comparison groups compare with kindergarten special education placement rates in the Chicago Public Schools? Third, to what extent are the findings similar to the national data about special education placement rates for children who attend preschool versus those who do not?

TO WHAT EXTENT ARE THE FINDINGS SIMILAR TO THOSE REPORTED IN THE CHICAGO LONGITUDINAL STUDY OF THE EARLIER CPC COHORT?

Studies of the Chicago Longitudinal Study (CLS) cohort of children who attended CPC preschool in 1984 showed that special education placement rates were lower for CPC recipients compared with children who did not participate in CPC preschool. In one CLS analysis that looked at children 6 to 18 years old, CPC participants had an *average* special education placement rate of 14.4% compared with 24.6% for children in the comparison group (Reynolds et al., 2003). This rate was calculated by averaging rates across school years, and therefore is not directly comparable to the kindergarten special education rates described in this report. In another CLS analysis that compared average special education placement rates from school entry through eighth grade for children who had attended a CPC preschool with a comparison group, special education placement rates were 12.5% versus 18.4%, respectively

(Conyers et al., 2003). Significant differences in special education placement rates between children in the CLS who did and did not attend CPC preschool programs emerged as early as first grade (0.5% versus 3.2%) (Conyers et al., 2003). We see difference in special education rates also emerging by the end of first grade, a 111% decrease for the intervention group relative to the comparison group for Cohort 3. These earlier CLS findings suggest that special education rates may rise over the early elementary grades and we may continue to detect a positive intervention effect for reduced special education placement from attendance in CPC preschool across all the early grades.²⁹ Other changes in how children are identified and placed in special education have occurred since the 1990s when these latter analyses were completed making these comparisons less appropriate than more contemporary data. These changes include but are not limited to improvements in early screening and programming prior to entering kindergarten as well as during the early elementary school years. For instance, with increasing attention to signs of early reading difficulties once children enter elementary school, and the growing research about the importance of skilled reading by 3rd grade for continuing school success (Snow & Matthews, 2016), many school districts may be identifying at risk children and providing extra instructional supports that prevent milder disabilities from emerging. Such improvements may mean that contemporary children who would have needed special education in the past are being identified during preschool or the early elementary grades and given more intensive or focused supports that can prevent those early difficulties from becoming more severe delays and learning problems requiring special education services.

The percentage of children needing special education for developmental delay, specific learning disability, or speech/language issues across all groups ranged from 3% in first grade for Cohort 1 to close to 10% for the Cohort 1 comparison group in third grade. Interestingly, the majority of children in the intervention cohort were identified during the preschool year while enrolled in CPC preschool classrooms. Because of changes in how children are identified for services and how important assessment and early intervention are in early care and education settings, we

²⁹ Support for this suggestion comes from earlier CPC studies that found that students with extended CPC program participation (through second or third grade) had lower rates of special education placements than CPC students having fewer years of intervention through middle (Reynolds & Temple, 1998) and high school (Reynolds, Temple, Robertson, & Mann, 2001).

believe as part of best practices in a high-quality preschool classroom, more children may be identified for services (and appropriately so) early in their school careers than may have occurred previously when the CLS was conducted. The hypothesized role of CPC in preventing or reducing the need for special education is still valid; however, in contemporary settings, there may be an initial increase in identification and placement before we see long-term reductions and/or average reductions in this outcome.

HOW DO KINDERGARTEN SPECIAL EDUCATION PLACEMENT RATES FOR THE INTERVENTION AND COMPARISON GROUPS COMPARE WITH KINDERGARTEN SPECIAL EDUCATION PLACEMENT RATES IN CHICAGO PUBLIC SCHOOLS?

In the previous 2015–16 school year, the kindergarten special education placement rate for CPS overall for what we are identifying as *mild delays and disorders* was 7.4%. Our findings show that both Cohort 1 and their comparison group have rates that are lower than the rate in CPS overall during the same year. We do not have a good explanation for this difference. Historical data from CPS demonstrate that special education placement rates during preschool have been relatively high for children attending CPS preschool (10–14% for the last 5 school years)³⁰ and that the overall CPS rates decrease in kindergarten (7–8%) and steadily climb through elementary and middle school grades; by third grade, for example, overall special education rates in CPS have been 10–11% for the last 5 school years (Chicago Public Schools, 2016). These special education rates in CPS suggest that we will see increases in special education rates as children move through the early elementary years, probably for both groups, which indeed is what is being found across cohorts over time (Exhibit 9). We hypothesized that the rate for the intervention group would increase at a slower rate and be lower than the rate in the comparison group in the later elementary years if the preschool CPC experiences have helped intervention group children have better developmental trajectories. Continued follow-up of the cohorts will address this question more fully, but the data up through third grade reported here suggest that it is being supported.³¹

³⁰ For the previous 2015-16 school year, we found that for low-income children who had attended any CPS preschool, the overall IEP special education rate in kindergarten was 13.7% (10.7% for mild and 3% for not mild delays and disorders).

³¹ The special education rates in CPS preschool may be higher than rates for kindergarten because children identified with a delay or disability prior to kindergarten would be likely to be referred to CPS and when determined to be eligible for an IEP, they would be served in a CPS preschool program.

TO WHAT EXTENT ARE THE FINDINGS SIMILAR TO THE NATIONAL DATA ABOUT SPECIAL EDUCATION PLACEMENT RATES FOR CHILDREN WHO ATTEND PRESCHOOL VERSUS THOSE WHO DO NOT?

Data collected annually from every state as part of federal reporting required under the Individual with Disabilities Education Act by the Office of Special Education Programs give us another perspective on the special education rates. The most recent national data from these sources indicate that 8% of children were receiving special education services at age 5, the kindergarten year for most students, 10% of 6 year olds (or equivalent to first grade students), and 11% of 7 year olds (equivalent to second grade students)³². Other data from the nationally representative Head Start Family and Child Experiences Survey (FACES) for children participating in Head Start preschool programs (Barton, Spiker, & Williamson, 2012) and the nationally representative ECLS-B study (Parsons, 2016)³³ found rates higher than those for our comparison sample (8% and 7%, respectively). However, findings from the FACES study also suggest that more children in Head Start without an Individual Education Plan (IEP) meet other criteria for disability or delay than are served (about 33% meet criteria indicative of delays such as very low assessment scores on standard measures, but only 8% had an IEP and were receiving special education services) (Barton et al., 2012). These national data and other research suggest that children from low-SES families are at greater risk of developmental delay and low levels of kindergarten readiness but may be less likely than higher SES children to be receiving special education services in early childhood (Morgan et al., 2019). Taken together, these other data suggest that the rates we are seeing in both the intervention and comparison groups are lower than would be expected for our high risk samples.

Third Grade Reading Findings

Of the original 313 students in Cohort 1, 294 had third grade reading scores on the NWEA MAP assessment. Most (197 of 294; 79.12%) met the 25th percentile or higher criterion. In examining CPS district-wide data, we find that 74.77% of third grade students in the district scored at or above the 25th percentile on the NWEA MAP

³² Data sources: U.S. Department of Education, *EDFacts* Data Warehouse (EDW): "IDEA Part B Child Count and Educational Environments Collection," 2017-18. Data extracted as of July 11, 2018 from file specifications 002 and 089. U.S. Bureau of the Census. "2017 State Population Estimates by Age, Sex, Race, and Hispanic Origin". Data accessed July 2018 from <http://www.census.gov/popest>

³³ ECLS-B children entered kindergarten in 2006 or 2007.

reading assessment across the 2017-2019 school years (P. Leonard, personal communication, June 21, 2021). This suggests that SIB-CPC participants have slightly higher scores on this reading assessment compared to the district overall.

Limitations

The evaluation is limited to data already collected in the district data and as such, must use for a kindergarten readiness measure, an assessment tool that was not developed for these purposes. The TS *GOLD*TM assessment is vulnerable to large (and maybe small) changes in training protocols and teacher turnover and both reliability and validity may suffer. In addition, the use of the TS *GOLD*TM measure with ELL populations is problematic because some ELL children can be assessed on all six domains in English while others cannot. Furthermore, there are not scale scores for children who have the TS *GOLD*TM language and literacy assessment domains in Spanish and therefore we did not have complete data needed for making the kindergarten readiness determination for those children. Because of the changes in items and scoring procedures with the TS *GOLD*TM data for Cohort 4, the evaluation team could not calculate percentages for kindergarten readiness in this new cohort of entering kindergarteners.

Use of a teacher-report measure as the indicator of kindergarten readiness can be seen as a limitation of this study. As described above, teacher effects or bias is also a limitation in the evaluation. The accuracy of teacher ratings of young children's behavior has been questioned by some early childhood researchers (Baker, Tichovolsky, Kupersmidt, Voegler-Lee, & Arnold, 2015; Mashburn & Henry, 2004). Additionally, there is a need in the field for more research about the amount and types of teacher training and knowledge needed to assess young children reliably and accurately, even while it is widely acknowledged that teacher training and knowledge are critical (Institute of Medicine & National Research Council, 2015; National Association Early Childhood Specialists in State Departments of Education & National Association for the Education of Young Children, 2003; National Association for the Education of Young Children, 2009).

For the special education outcome at kindergarten, to the best of our ability we created a comparison group that was weighted to match closely the characteristics of children in the intervention cohort. However, both propensity score methods—PSW

and PSM—have some disadvantages. One disadvantage is that these methods only account for observed (and observable) covariates. They cannot balance intervention and comparison groups on unobservable characteristics (for example, parent education levels or parent involvement in the children’s early learning). Second, we can only identify those children who attended a CPS-funded preschool program prior to kindergarten. Many entering kindergarten children who meet the income and age criteria may have attended other preschool programs either public (e.g., state-funded preschool outside of the city of Chicago) or private (e.g., using child care subsidies for example). The evaluation team had no way of identifying which children attended these other settings because this information is not routinely collected at kindergarten entry in CPS.

Conclusion

Together, these findings show that early childhood education in the form of CPC is associated with rates of kindergarten readiness of 61% in Cohort 1, 42% in Cohort 2, and 44% in Cohort 3. The findings also show that there is a significant decrease in special education starting in kindergarten for the intervention group for Cohorts 2, 3, and 4. While the overall special education rates in both the intervention and comparison group children are very low at the end of kindergarten for all four cohorts, the rates are lower in the intervention group than the comparison group in all four cohorts and across all grades. These data suggest that perhaps over time children served in the CPC preschool classrooms in the year prior to entering kindergarten received several supports and services that led to a decreased need for special education services in the first four years of elementary school. In particular, we may see a group of children who were referred and assessed for special education supports during preschool and kindergarten then remediating by first and second grade such that they no longer need those additional supports. Additional follow-up of cohorts through third and fourth grade and beyond would help better understand these trajectories in both the CPC students as well as their peers who did not receive district preschool.

References

- Baker, C. N., Tichovolsky, M. H., Kupersmidt, J. B., Voegler-Lee, M. E., & Arnold, D. H. (2015). Teacher (mis)perceptions of preschoolers' academic skills: Predictors and associations with longitudinal outcomes. *Journal of Educational Psychology, 107*(3), 805-820. doi:10.1037/edu0000008
- Barton, L., Spiker, D., & Williamson, C. (2012). Characterizing disability in Head Start programs: Not so clearcut. *Early Childhood Research Quarterly, 27*(4), 569-612.
- Becker, S. O., & Inchino, A. (2002). Estimation of average treatment effects based on propensity scores. *Stata Journal, 2*, 358-377.
- Belfield, C., Nores, M., Barnett, W. S., & Schweinhart, L. J. (2006). The High/Scope Perry Preschool Program: Cost-benefit analysis using data from the age-40 followup. *Journal of Human Resources, 41*(1), 162-190.
- Blackorby, J., Schiller, E., Mallik, S., Hebbeler, K., Huang, T., Javitz, H., . . . Williamson, C. (2010). *Patterns in the identification of and outcomes for children and youth with disabilities*. Menlo Park, CA: SRI International.
- Brooks-Gunn, J., Rouse, C. E., & McLanahan, S. (2007). Racial and ethnic gaps in school readiness. In R. C. Pianta, M. J. Cox, & K. L. Snow (Eds.), *School readiness and the transition to kindergarten in the era of accountability* (pp. 283-305). Baltimore, MD: Paul H. Brookes Publishing.
- Burger, K. (2010). How does early childhood care and education affect cognitive development? An international review of the effects of early interventions for children from different social backgrounds. *Early Childhood Research Quarterly, 25*(2), 140-165.
- Burts, D. C., & Kim, D.-H. (2014). The Teaching Strategies GOLD Assessment System: Measurement Properties and Use. *NHSA Dialog, 17*(3).
- Caliendo, M., & Kopeinig, S. (2008). Some practical guidance for the implementation of propensity score matching. *Journal of Economic Surveys, 22*(1), 31-72.
- Campbell, F. A., Ramey, C. T., Pungello, E., Sparling, J., & Miller-Johnson, S. (2002). Early childhood education: Young adult outcomes from the Abecedarian Project. *Applied Developmental Science, 6*(1), 42-57.
- Chicago Public Schools. (2016). School data. In: Chicago Public Schools.
- Conyers, L. M., Reynolds, A. J., & Ou, S. (2003). The effect of early childhood intervention and subsequent special education services: Findings from the Chicago Child-Parent Centers. *Educational Evaluation and Policy Analysis, 25*, 75-95.
- Curtis, L. H., Hammill, B. G., Eisenstein, E. L., Kramer, J. M., & Anstrom, K. J. (2007). Using inverse probability-weighted estimators in comparative effectiveness. Analyses with observational databases. *Medical Care, 45*(2), 158-165.
- Decker, C. G. (2013). *Teaching Strategies GOLD: Testing reliability and validity using the Bracken School Readiness Assessment*. Unpublished report of CAP Tulsa.

- Gaylor, E., Ferguson, K., McCracken, M., Wei, X., & Spiker, D. (2017). *Evaluation of kindergarten readiness in nine Child-Parent Centers: Report for 2015-16*. Menlo Park, CA: SRI International.
- Gaylor, E., Kutaka, T., Ferguson, K., Williamson, C., Wei, X., & Spiker, D. (2016). *Evaluation of kindergarten readiness in five Child-Parent Centers: Report for 2014-15*. Menlo Park, CA: SRI International.
- Hair, E. C., Halle, T., Terry-Humen, E., Lavelle, B., & Calkins, J. (2006). Children's school readiness in the ECLS-K: Predictions to academic, health, and social outcomes in first grade. *Early Childhood Research Quarterly, 21*, 431-454.
- Harder, V. S., Stuart, E. A., & Anthony, J. C. (2010). Propensity score techniques and the assessment of measured covariate balance to test causal associations in psychological research. *Psychological Methods, 15*(3), 234-249.
- Harding, D. (2003). Counterfactual Models of Neighborhood Effects: The Effect of Neighborhood Poverty on Dropping Out and Teenage Pregnancy. *American Journal of Sociology, 109*(3), 676-719. doi:10.1086/379217
- Heckman, J., Ichimura, H., Smith, J., & Todd, P. (1998). Characterizing selection bias using experimental data. *Econometrica, 66*(5), 1017-1098. doi:10.2307/2999630
- Hirano, K., Imbens, G. W., & Ridder, G. (2003). Efficient estimation of average treatment effects using the estimated propensity score. *Econometrica, 71*(4), 1161-1189. doi:10.1023/a:1020371312283
- Hogan, D. P., Msall, M. E., Rogers, M. L., & Avery, R. C. (1997). Improved disability population estimates of functional limitation among American children aged 5–17. *Maternal and Child Health Journal, 1*(4), 203-216. doi:10.1023/a:1022354126797
- Human Capital Research Collaborative. (2015). *2015-16 program guidelines and requirements. CPC-PK3 program*. Minneapolis, MN: University of Minnesota. Retrieved from <https://humancapitalrc.org/midwest-cpc/cpc-resources>
- Institute for Research on Poverty. (1997). Investing in young children. *Focus, 19*(1), 1-4.
- Institute of Medicine & National Research Council. (2015). *Transforming the workforce for children birth through age 8: A unifying foundation*. Washington, DC: National Academic Press.
- Isaacs, J. B. (2012). *Starting school at a disadvantage: The school readiness of poor children*. Washington DC: Center on Children and Families at Brookings.
- Karoly, L. A., Greenwood, P. W., Everingham, S. S., Hoube, J., Kilburn, M. R., Rydell, C. P., . . . Chiesa, J. (1998). *Investing in our children: What we know and don't know about the costs and benefits of early childhood interventions* (MR-898-TCWF). Santa Monica, CA: RAND Corporation.
- Kim, D. H., Lambert, R. G., & Burts, D. C. (2013). Evidence of the validity of Teaching Strategies GOLD® assessment tool for English language learners and children with disabilities. *Early Education & Development, 24*(4), 574-595.

- Knudsen, E. I., Heckman, J. J., Cameron, J. L., & Shonkoff, J. P. (2006). Economic, neurobiological, and behavioral perspectives on building America's future workforce. *Proceedings of the National Academy of Sciences*, *103*(27), 10155–10162.
- Lambert, R. G., Kim, D. H., & Burts, D. C. (2014a). *Technical manual for the Teaching Strategies GOLD® assessment system (3rd edition)*. Center for Educational Measurement and Evaluation, UNCC, Charlotte.
- Lambert, R. G., Kim, D. H., & Burts, D. C. (2014b). Using teacher ratings to track the growth and development of young children using the Teaching Strategies GOLD® assessment system. *Journal of Psychoeducational Assessment*, *32*(1), 27-39.
- Lechner, M. (2002). Some practical issues in the evaluation of heterogeneous labour market programmes by matching methods. *Journal of the Royal Statistical Society Series a-Statistics in Society*, *165*, 59-82. doi:10.1111/1467-985X.0asp2
- Lee, R., Zhai, F., Brooks-Gunn, J., Han, W.-J., & Waldfogel, J. (2014). Head Start participation and school readiness: Evidence from the Early Childhood Longitudinal Study-Birth cohort December 10, 2012. *Developmental Psychology*, *50*(1), 202-215.
- Lee, V. E., & Burkam, D. T. (2002). *Inequality at the starting gate: Social background differences in achievement as children begin school*. Washington, DC: Economic Policy Institute.
- Lynch, R. G. (2007). *Enriching children, enriching the nation: Public investment in high-quality prekindergarten*. Washington, DC: Economic Policy Institute.
- Mashburn, A. J., & Henry, G. T. (2004). Assessing school readiness: Validity and bias in preschool and kindergarten teachers' ratings. *Educational Measurement: Issues and Practice*, *23*(4), 16-30.
- McCoy, D. C., Yoshikawa, H., Ziol-Guest, K. M., Duncan, G. J., Schindler, H. S., Magnuson, K., . . . Shonkoff, J. P. (2017). Impacts of early childhood education on medium- and long-term educational outcomes. *Educational Researcher*, *46*(8), 474–487. doi:10.3102/0013189X17737739
- Morgan, P. L., Farkas, G., Hillemeier, M. M., Wang, Y., Mandel, Z., DeJarnett, C., & Maczuga, S. (2019). Are students with disabilities suspended more frequently than otherwise similar students without disabilities? *Journal of School Psychology*, *72*, 1-13. doi:<https://doi.org/10.1016/j.jsp.2018.11.001>
- Muschkin, C. G., Ladd, H. F., & Dodge, K. A. (2015). Impact of North Carolina's early childhood initiatives on special education placements in third grade. *Educational Evaluation and Policy Analysis*. doi:10.3102/0162373714559096
- National Association Early Childhood Specialists in State Departments of Education, & National Association for the Education of Young Children. (2003). *Early childhood curriculum, assessment, and program evaluation: Building an effective, accountable system in programs for children birth through age 8* (Position Statement with expanded resources). Washington, DC: National Association Early Childhood Specialists in State Departments of Education (NAECS/SDE), National Association for the Education of Young Children. Retrieved from www.naeyc.org/resources/position_statements/pscape.pdf

- National Association for the Education of Young Children. (2009). *Developmentally appropriate practice in early childhood programs serving children from birth through age 8* (Position Statement). Washington, DC: Author.
- National Research Council. (2008). *Early childhood assessment: Why, what, and how?* Committee on Developmental Outcomes and Assessments for Young Children, C. E. Snow & S. B. Van Hemel, (Eds.). Washington, DC: The National Academies Press.
- Parsons, S. (2016). *Panel paper: Factors and processes predicting placement into special education in early childhood* Paper presented at the 38th Annual Fall Research Conference: The Role of Research in Making Government More Effective, Washington, DC. <https://appam.confex.com/appam/2016/webprogram/Paper18665.html>
- Ravallion, M. (2001). The mystery of the vanishing benefits: An introduction to impact evaluation. *The World Bank Economic Review*, 15(1), 115-140.
doi:10.1093/wber/15.1.115
- Reynolds, A. J. (1994). Effects of a preschool plus follow-on intervention for children at risk. *Developmental Psychology*, 30, 787-804.
- Reynolds, A. J. (1995). One year of preschool intervention or two: Does it matter? *Early Childhood Research Quarterly*, 10, 1–31.
- Reynolds, A. J. (2000). *Success in early intervention: The Chicago Child-Parent Centers*. Lincoln, NE: University of Nebraska Press.
- Reynolds, A. J. (2002). *Chicago's preschool programs: Do they promote children's school success? Executive Summary II*. Chicago, IL: Chicago Board of Education.
- Reynolds, A. J., & Hayakawa, C. M. (2011). Why the Child-Parent Center education program promotes life-course development. In E. Zigler, W. S. Gilliam, & W. S. Barnett (Eds.), *The pre-k debates. Current controversies and issues*. New York, NY: Brookes.
- Reynolds, A. J., Hayakawa, M., Ou, S., Mondri, C. F., Englund, M. M., Candee, A. J., & Smerillo, N. E. (2017). Scaling and sustaining effective early childhood programs through school–family–university collaboration. *Child Development*, 88(5), 1453-1465.
- Reynolds, A. J., Magnuson, K. A., & Ou, S. R. (2010). Preschool-to-third grade programs and practices: A review of research. *Children and Youth Services Review*, 32(8), 1121-1131.
- Reynolds, A. J., Richardson, B. A., Hayakawa, M., Englund, M. M., & Ou, S.-R. (2016). Multi-site expansion of an early childhood intervention and school readiness. *Pediatrics*, 138(1), 1-11.
- Reynolds, A. J., Rolnick, A. J., & Temple, J. A. (Eds.). (2014). *Health and education in early childhood: predictors, interventions, and policies*. Cambridge, United Kingdom: Cambridge University Press.
- Reynolds, A. J., & Temple, J. A. (1998). Extended early childhood intervention and school achievement: Age thirteen findings from the Chicago longitudinal study. *Child Development*, 69, 231-246.
- Reynolds, A. J., & Temple, J. A. (2008). Cost-effective early childhood development programs from preschool to third grade. *Annual Review of Clinical Psychology*, 4, 109-139.

- Reynolds, A. J., Temple, J. A., & Ou, S.-R. (2003). School-based early intervention and child well-being in the Chicago Longitudinal Study. *Child Welfare, 82*(5), 633-656.
- Reynolds, A. J., Temple, J. A., Robertson, D. L., & Mann, E. A. (2001). Long-term effects of an early childhood intervention on educational achievement and juvenile arrest: A 15-year follow-up of low-income children in public schools. *Journal of the American Medical Association, 285*(18), 2339-2346.
- Reynolds, A. J., Temple, J. A., Robertson, D. L., & Mann, E. A. (2002). Age 21 cost-benefit analysis of the Title I Chicago Child-Parent Centers. *Educational Evaluation and Policy Analysis, 24*(4), 267-303.
- Root, E. D., & Humphrey, J. L. (2014). Neighborhood racial composition and trajectories of child self-rated health: An application of longitudinal propensity scores. *Social Science & Medicine, 120*, 31-39. doi:<http://dx.doi.org/10.1016/j.socscimed.2014.09.006>
- Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika, 70*(1), 41-55.
- Rosenbaum, P. R., & Rubin, D. B. (1984). Reducing bias in observational studies using subclassification on the propensity score. *Journal of the American Statistical Association, 79*, 516-524.
- Rosenbaum, P. R., & Rubin, D. B. (1985). Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *American Statistician, 39*, 33-38.
- Russo, J. M., Williford, A. P., Markowitz, A. J., Vitiello, V. E., & Bassok, D. (2019). Examining the validity of a widely-used school readiness assessment: Implications for teachers and early childhood programs. *Early Childhood Research Quarterly, 48*, 14-25.
- Sampson, R. J., Sharkey, P., & Raudenbush, S. W. (2008). Durable Effects of Concentrated Disadvantage on Verbal Ability Among African-American Children. *Proceedings of the National Academy of Sciences, 105*, 845-852.
- Shonkoff, J., & Phillips, D. (Eds.). (2000). *From neurons to neighborhoods: The science of early childhood development*. Washington, DC: National Academy Press.
- Snow, C. E., & Matthews, T. J. (2016). Reading and language in the early grades. *The Future of Children, 26*(2), 57-74.
- Soderberg, J. S., Stull, S., Cummings, K., Nolen, E., McCutchen, D., & Joseph, G. (2013). *Inter-rater reliability and concurrent validity study of the Washington Kindergarten Inventory of Developing Skills (WaKIDS)*. Unpublished report prepared for the State of Washington Office of Superintendent of Public Instruction.
- State of Washington, Office of Superintendent of Public Instruction (n.d.). *Washington State report card. Washington kindergarten inventory of developing skills (WaKIDS)*. Olympia, WA: Author. Retrieved from <http://reportcard.ospi.k12.wa.us/WaKidsDetailPage.aspx?domain=WaKIDS&year=2014-15&wakidsyr=201314&schoolId=1&waslCategory=1&numberOrChart=1&yrs=2014-15%29>

Teaching Strategies. (2013). *Teaching Strategies GOLD assessment system. Concurrent validity*. Bethesda, MD: Author. Retrieved from <https://teachingstrategies.com/wp-content/uploads/2017/03/GOLD-Concurrent-Validity-2013.pdf>

ten Bensel, T., Gibbs, B., & Lytle, R. (2015). A Propensity Score Approach towards Assessing Neighborhood Risk of Parole Revocation. *American Journal of Criminal Justice*, 40(2), 377-398. doi:10.1007/s12103-014-9269-z

Appendices

Appendix A: Chicago Child-Parent Center Social Impact Bond Evaluation Plan

Appendix B: Timing of Cohorts

Appendix C: Propensity Score Weighting Process