

Final Evaluation Results of The Milwaukee Community Literacy Project/ SPARK Early Literacy

Curtis Jones, SREed, University of Wisconsin, Milwaukee

Michael Christian, Education Analytics

Andrew Rice, Education Analytics

March, 3rd, 2016



Presented to the Annual Meeting of the Society for Research on Educational Effectiveness, Washington DC
Project funded by a Department of Education Investing in Innovations (i3) grant. Contract number U396C100694.

Literacy Context in Milwaukee

The need for increased literacy development opportunities for Milwaukee students is urgent. According to the 2013 National Assessment of Educational Progress (NAEP) 15% of 4th grade MPS students were At or Above Proficient in reading.

This need is even more pronounced for low-income and minority students.

- 38% of 4th grade, White MPS students were proficient in reading, compared to 9% of Black and 14% of Hispanic students.
- 11% of 4th grade low-income (free/reduced lunch participants) MPS students were proficient in reading, compared to 39% of non-low-income students.

History of SPARK Early Literacy

2005 - SPARK was created to address this need by Boys & Girls Clubs of Greater Milwaukee and piloted at one site in Milwaukee Public Schools (MPS).

2006 - SPARK was expanded to three MPS schools with funding from the United Way and AmeriCorps.

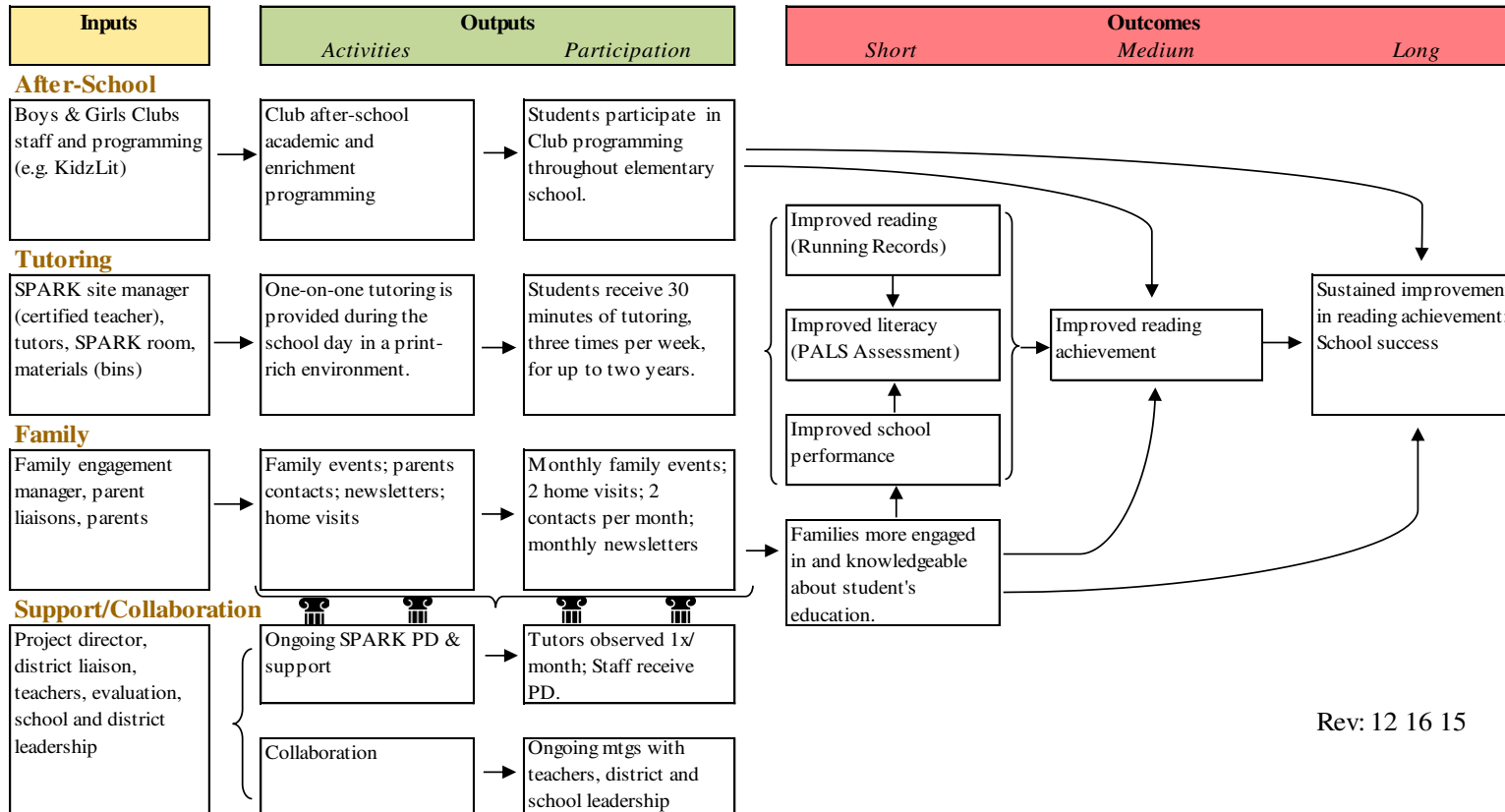
2010 - SPARK received a Department of Education investing in innovation (i3) grant award to expand to seven more schools (10 total).



Program: SPARK Early Literacy Logic Model

Situation: The need for increased literacy opportunities in the Milwaukee area is urgent. According to the 2013 National Assessment of Educational Progress (NAEP) 15% of 4th grade MPS students were At or Above Proficient in reading.

This need is even more pronounced for low-income and minority students. 38% of 4th grade, White MPS students were proficient in reading, compared to 9% of Black and 14% of Hispanic students. 11% of 4th grade low-income (free/reduced lunch participants) MPS students were proficient in reading, compared to 39% of non-low-income students.



Rev: 12 16 15

Assumptions

SPARK adds capacity to a school by supplying tutors and a parents liaison to work with students.

Typical literacy interventions only focus on skill acquisition and ignore the reasons why students fall behind.

The benefits observed in other literacy interventions have typically faded over time.

Boys & Girls Clubs is in a unique position, however, to continue to work with students and families beyond SPARK.

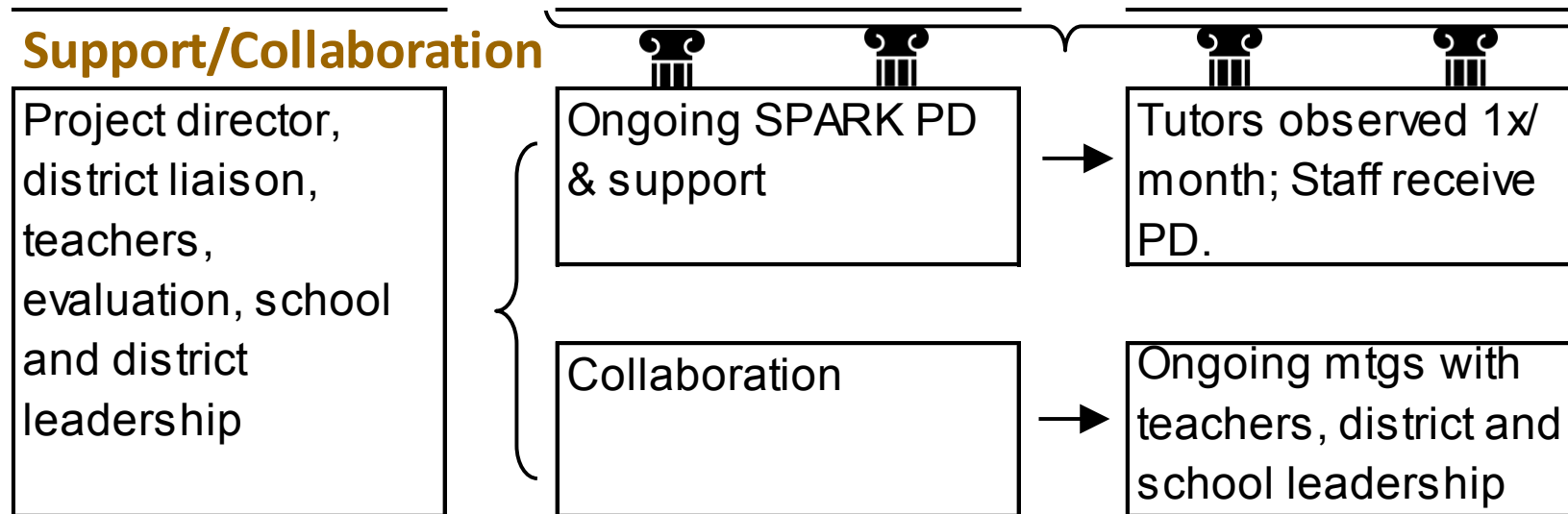
Context

Many districts face staffing and budget shortages that prevent them from using other reading interventions.

The RTI model does not work when 90% of students are below proficiency in reading.

Chronic absenteeism (missing more than 10% of school days) is a serious program in MPS and other urban districts. Many parents do not understand the importance of early grade participation.

Collaboration and Support



Inputs

After-School

Boys & Girls Clubs staff and programming (e.g. KidzLit)

Tutoring

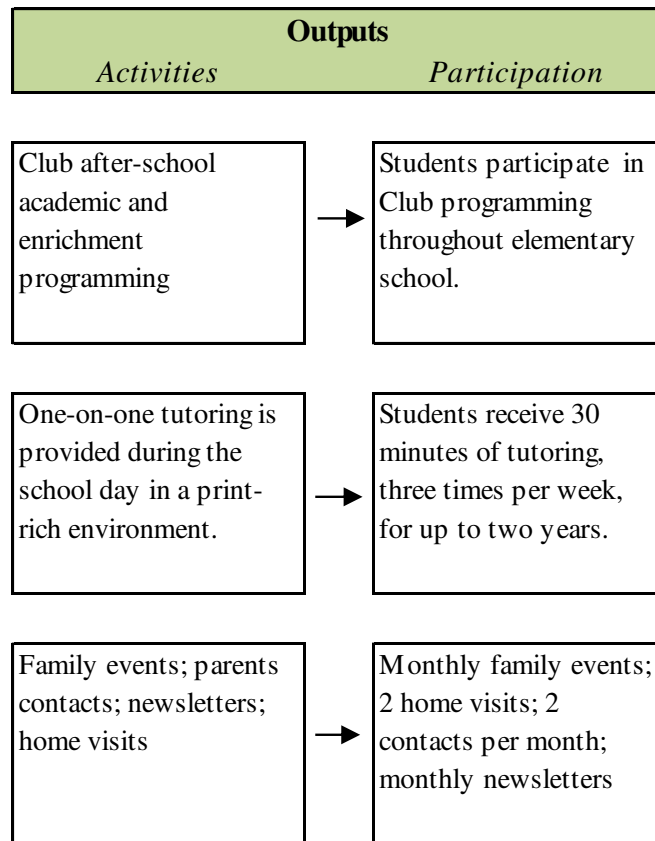
SPARK site manager (certified teacher), tutors, SPARK room, materials (bins)

Family

Family engagement manager, parent liaisons, parents

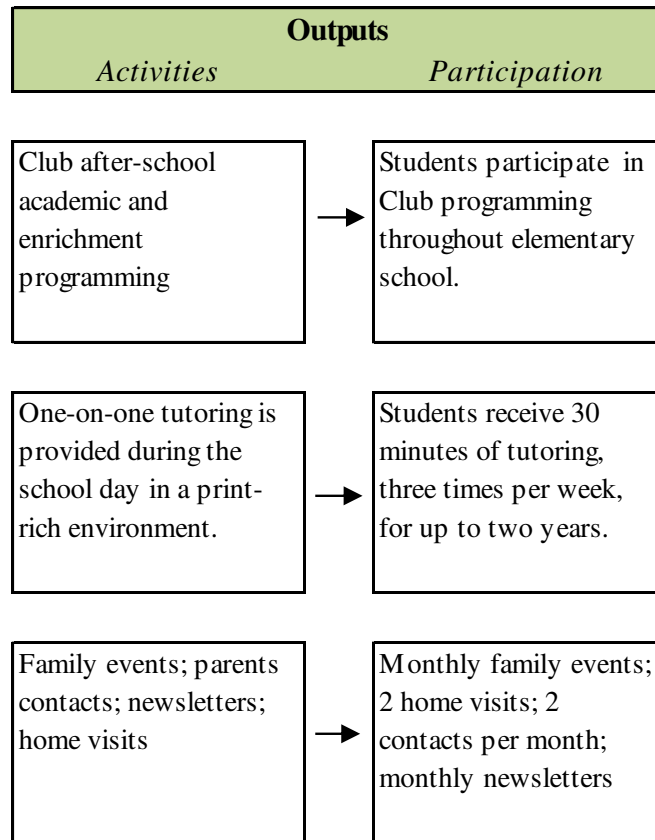
- Certified teacher in each site supports and supervises a cadre of about six tutors.
- Tutors include AmeriCorps members, UWM preservice teachers, and a small number of community members.
- Tutoring is done in a print-rich “SPARK room”
- Each site also has a parent liaison who would work with SPARK families.

Activities - Tutoring



- Students are pulled out of non-core classes and taken to SPARK room three times per week for up to two years.
- Starts with familiar activity.
- Running Records assessment every third session.
- Word play - Students receive instruction using Word Sorts and Making Words. Word Sorts involve students sorting words into categories to increase their understanding of sounds and letters. Making Words involves students using letters to make words so students learn how the sounds of language are put together.
- Reading a book at instructional level.
- Writing sentences - Elkonin boxes are a central piece of SPARK writing and used to help students encode words.
- End with tutor read aloud.

Activities – Family Engagement



The goals of family engagement is to keep families aware of student progress in SPARK, help families promote literacy at home, and address any school attendance issues that arise during the program.

- Monthly family events,
- 2 home contacts per month (email, in person, or by phone),
- 1 home visit per year, usually over the summer,
- Monthly newsletters sent home.

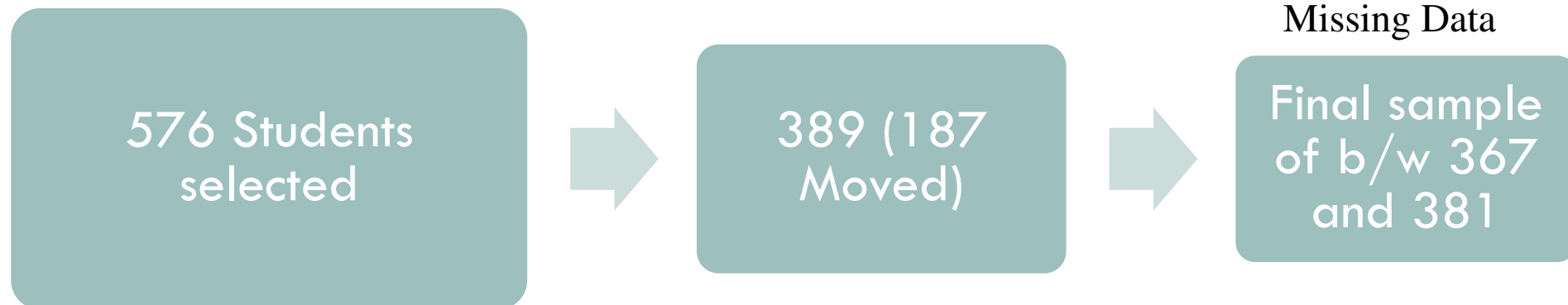
Evaluation Design

The primary goal of the evaluation was to design a study that would meet What Works Clearinghouse (WWC) design standards without reservations.

- Studies have to use a randomized-control framework to qualify.
 - ✓ 576 students were randomly assigned to either the control (290) or SPARK (286) group by the evaluation team. Assignments were stratified by school and grade level within schools. ELL and IEP students were not eligible for study.
- Studies have to use valid and reliable outcome measures.
 - ✓ NWEA Measures of Academic Progress (MAP) Reading Benchmark Assessment,
 - ✓ Phonological Literacy Screening (PALS),
 - ✓ Regular-school-day attendance.
- Studies have to meet WWC attrition and differential attrition standards.

Overall Selection & Attrition Results

33.9% - 36.3% Overall Attrition Rate



Differential Attrition Results

1.9% to 2.3% Differential Attrition Rate

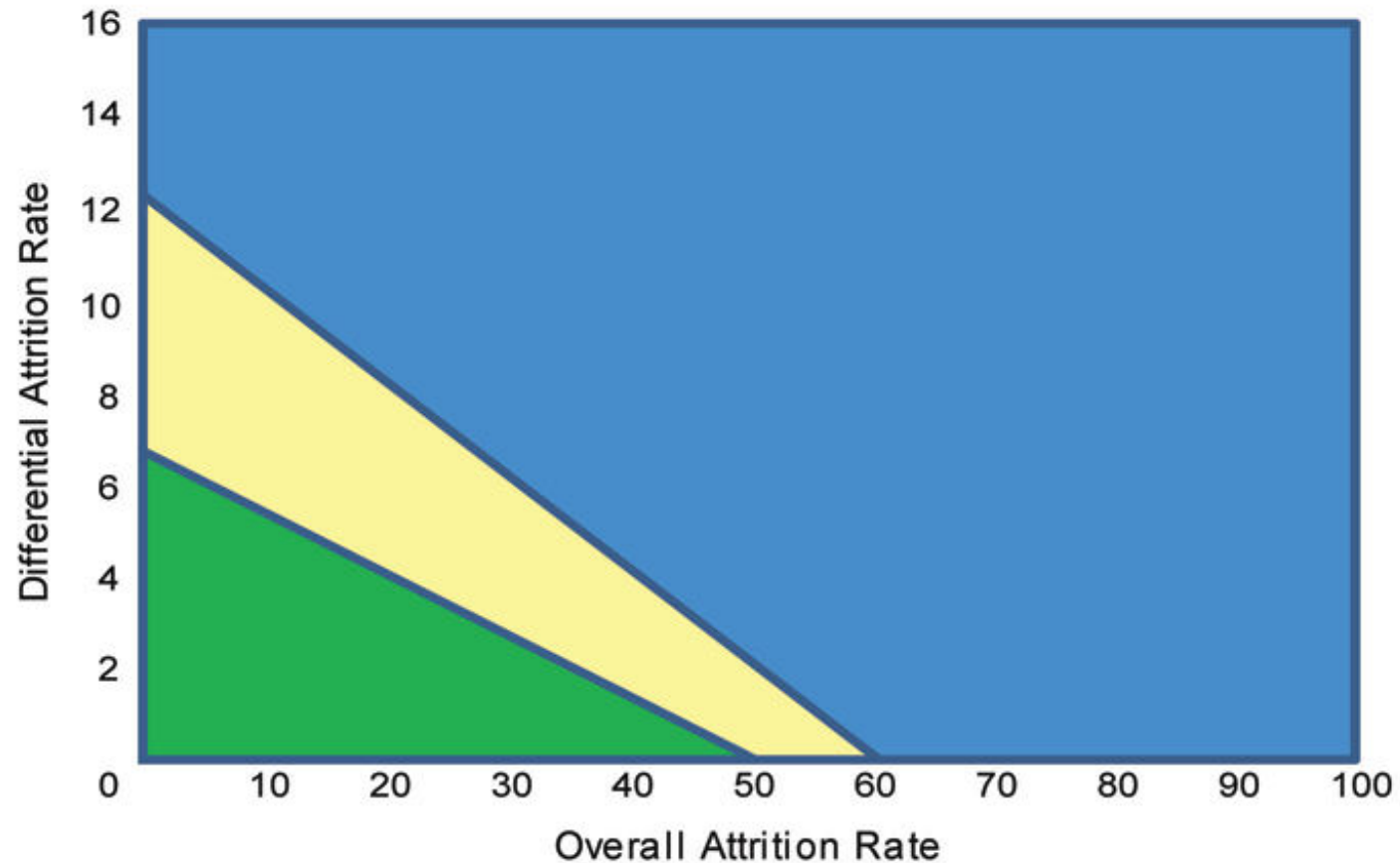
SPARK Participants



Control Students



What Works Clearinghouse (WWC) Attrition Standards



Final Sample

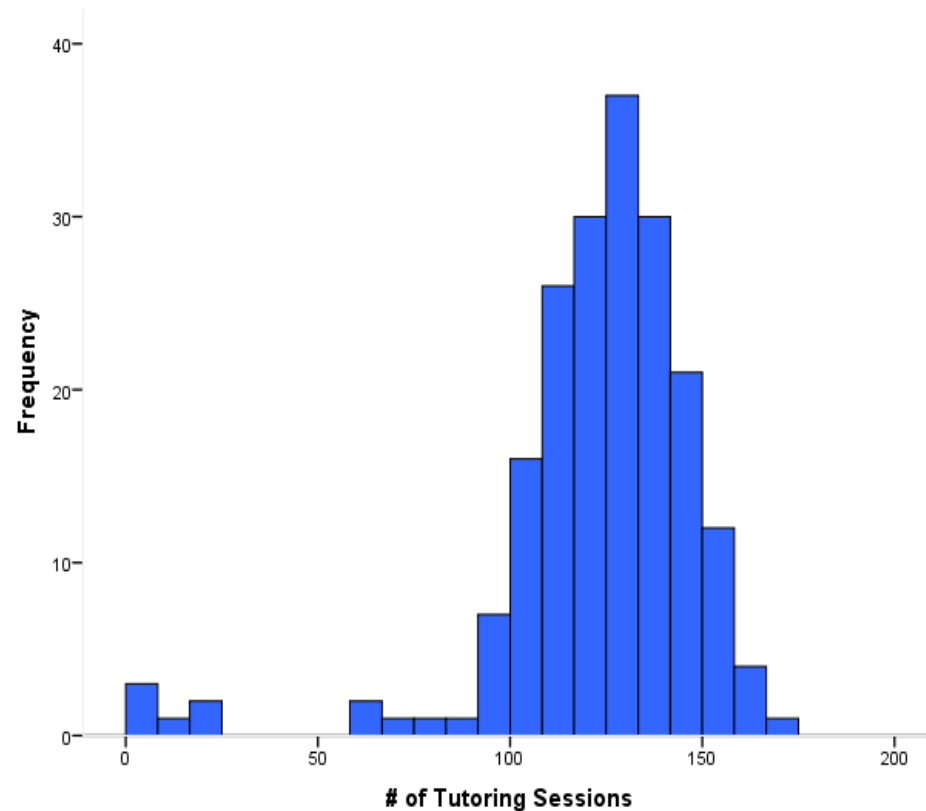
		Control	SPARK	Total
Grade Level	K-1 st	72	66	138
	1 st -2 nd	63	74	137
	2 nd -3 rd	59	55	114
Race/Ethnicity	Black	149	146	295
	Hispanic	29	34	63
	Other	16	15	31
Gender	Female	96	100	196
	Male	98	95	193
F/R Lunch Eligible	No	9	8	17
	Yes	185	187	372
IEP	No	182	184	366
	Yes	12	11	23
Total		194	195	389

Evaluation Design

The primary goal of the evaluation was to design a study that would meet What Works Clearinghouse (WWC) design standards without reservations.

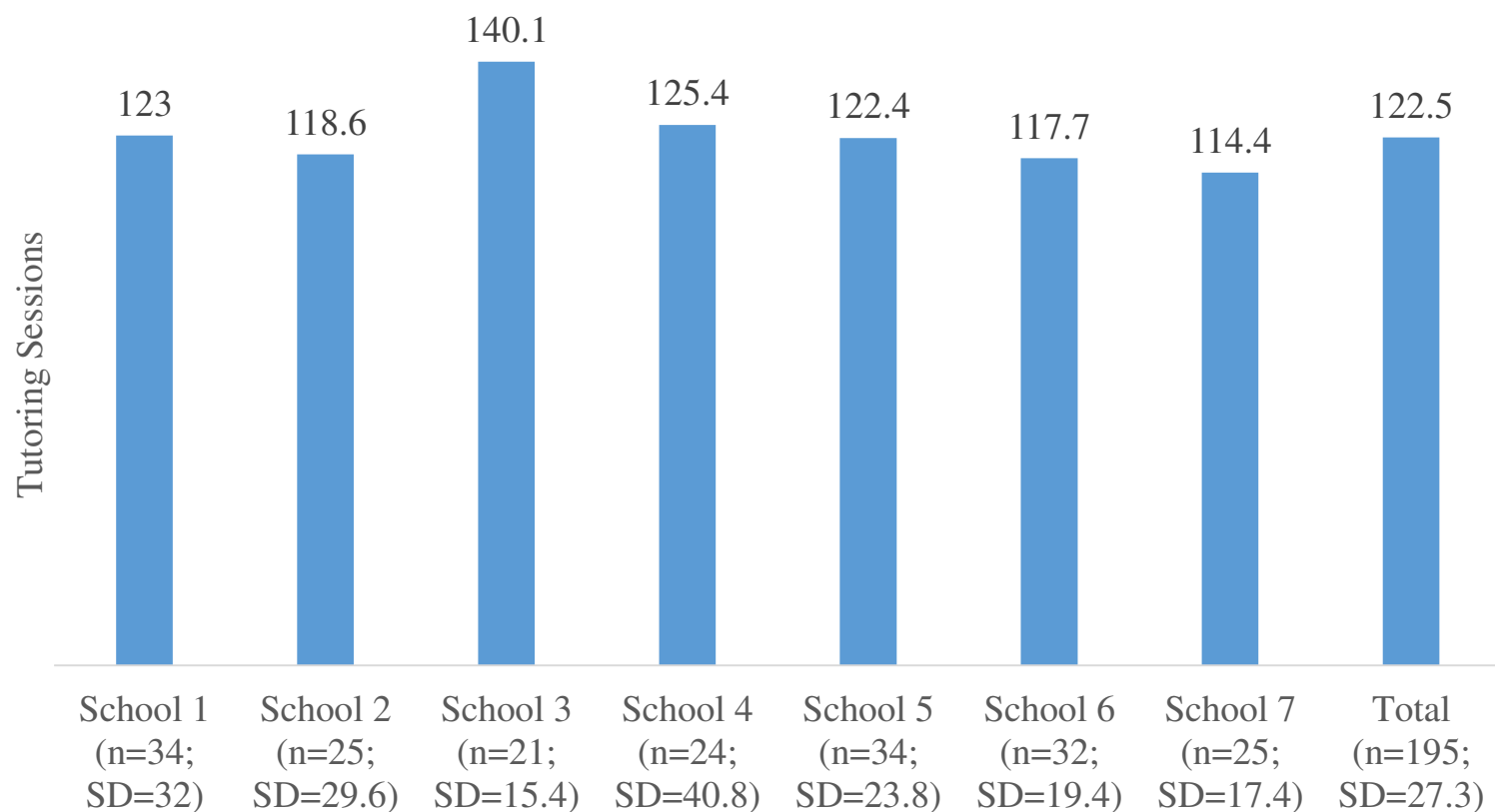
- Studies have to use a randomized-control framework to qualify.
 - ✓ 576 students were randomly assigned to either the control (290) or SPARK (286) group by the evaluation team. Assignments were stratified by school and grade level within schools. ELL and IEP students were not eligible for study.
- Studies have to use valid and reliable outcome measures.
 - ✓ NWEA Measures of Academic Progress (MAP) Reading Benchmark Assessment,
 - ✓ Phonological Literacy Screening (PALS),
 - ✓ Regular-school-day attendance.
- Studies have to meet WWC attrition and differential attrition standards.
 - ✓ Study attrition results were within acceptable ranges. In addition, attrition was exogenous to the program. Thus, the evaluation should meet WWC attrition standards.
- The program has to be implemented as intended.

MCLP/SPARK Implementation - Tutoring



- SPARK students received an intensive amount of services across the two program years while control group students received the “business as usual” reading instruction provided by MPS.
- The average SPARK student received 122.5 tutoring sessions (SD = 27.3).
- Nearly all students received more than the minimum dosage of 90 sessions.

MCLP/SPARK Implementation – Average Tutoring Sessions by School

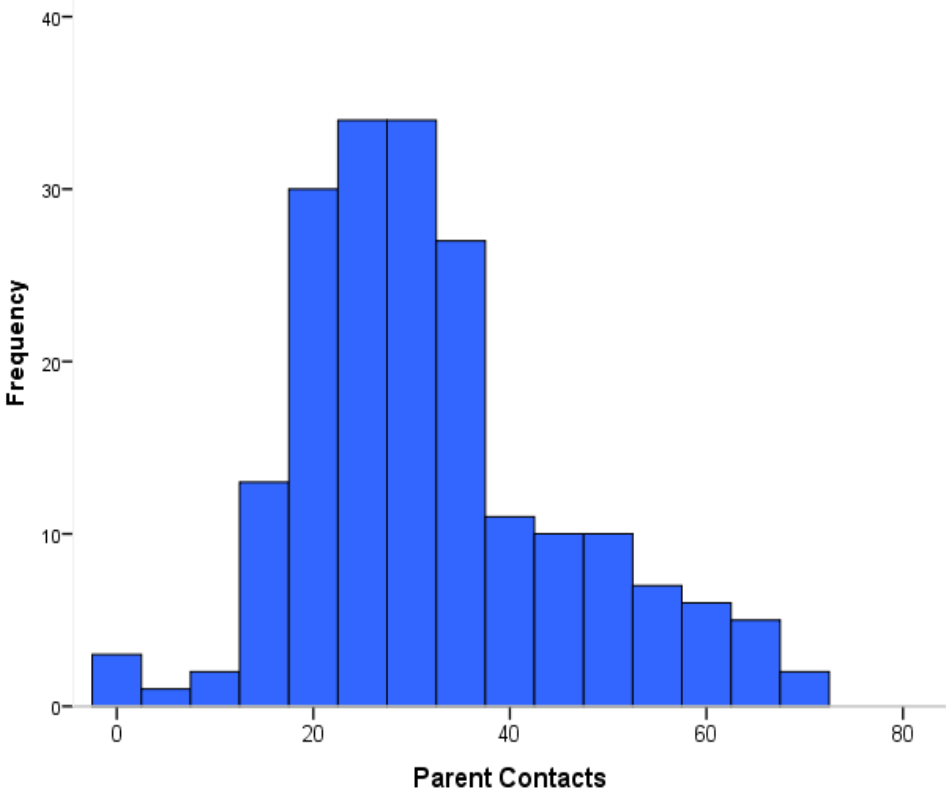
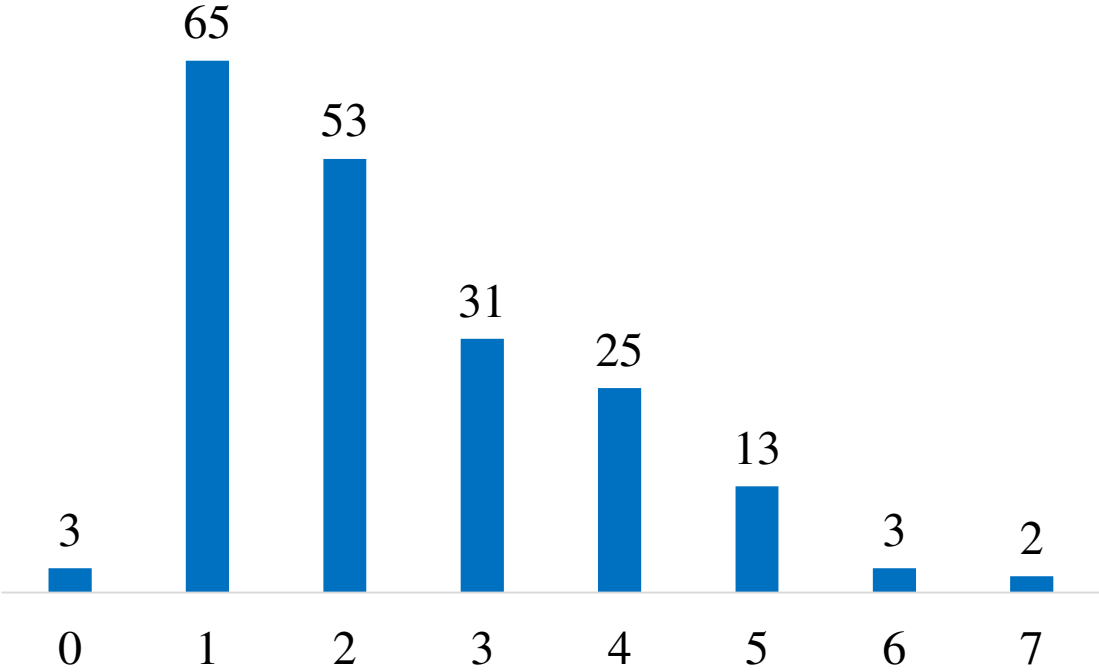


MCLP/SPARK Implementation - Family

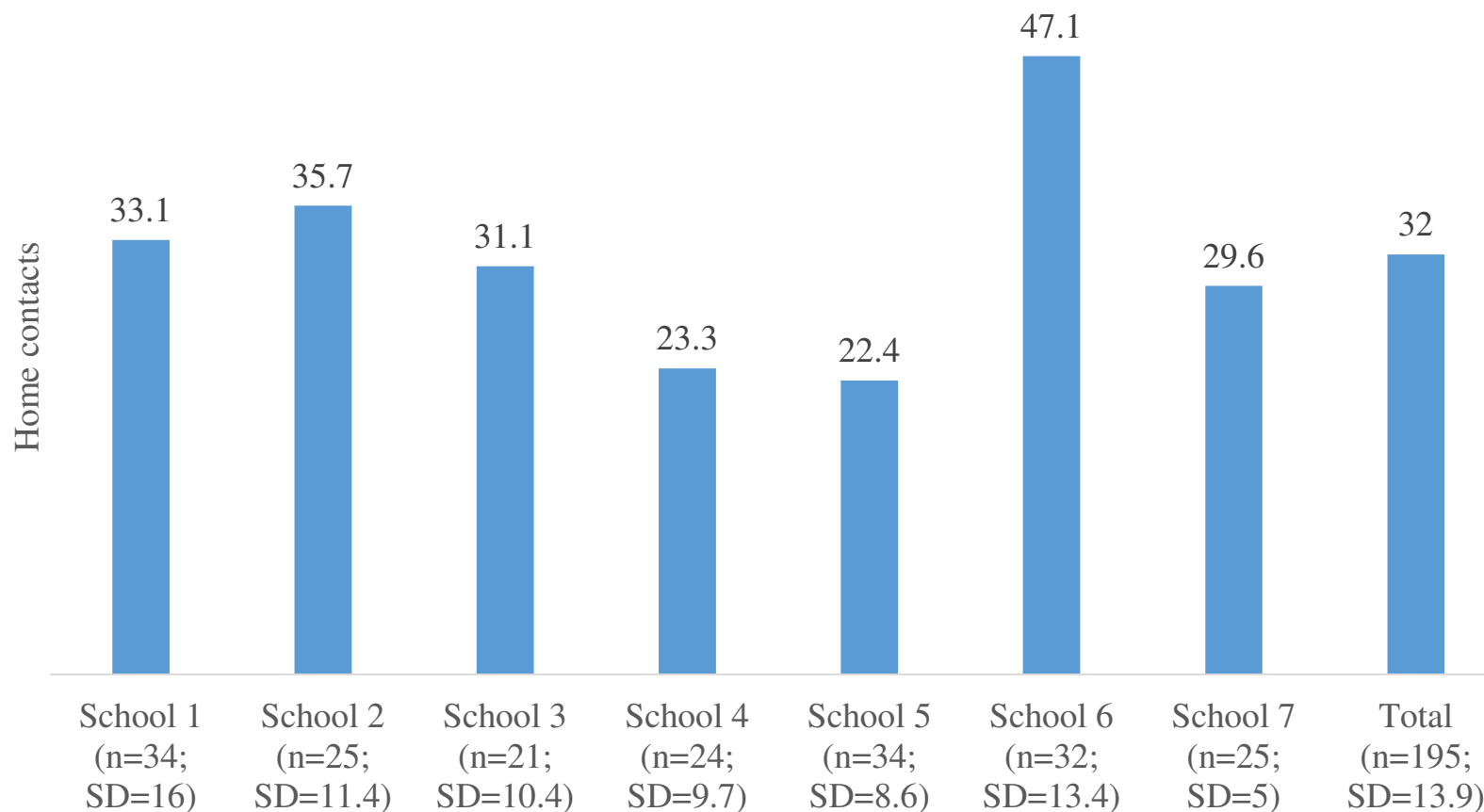
- All schools sent home monthly newsletters to all SPARK students.
- All schools held monthly family SPARK events.
- Parents of SPARK students attended an average of three family events (Range 0 to 10, SD = 2.2), received 32 parent contacts (Range 0 to 69, SD = 13.9) and had 2.4 home visits (Range 0 to 8, SD = 1.4).
- 47 out of 53 interviewed parents indicated that they received regular communications from SPARK.
- All 53 indicated that they understood the components of the SPARK lesson plan.

MCLP/SPARK Implementation - Family

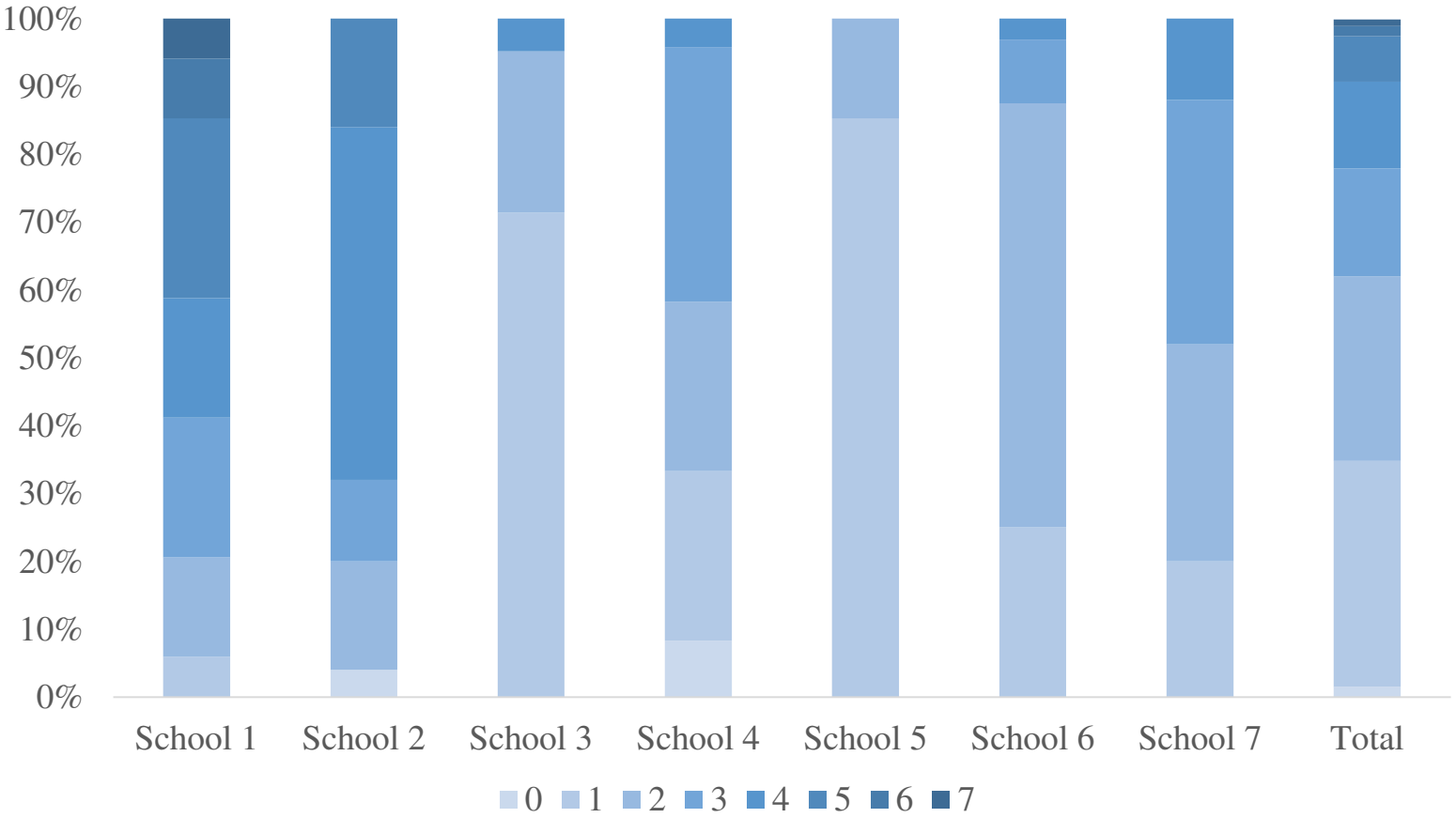
Number of Home Visits



MCLP/SPARK Implementation – Family Average Home Contacts per Participant by School



MCLP/SPARK Implementation – Family Number of Home Visits per Participant by School



Evaluation Design

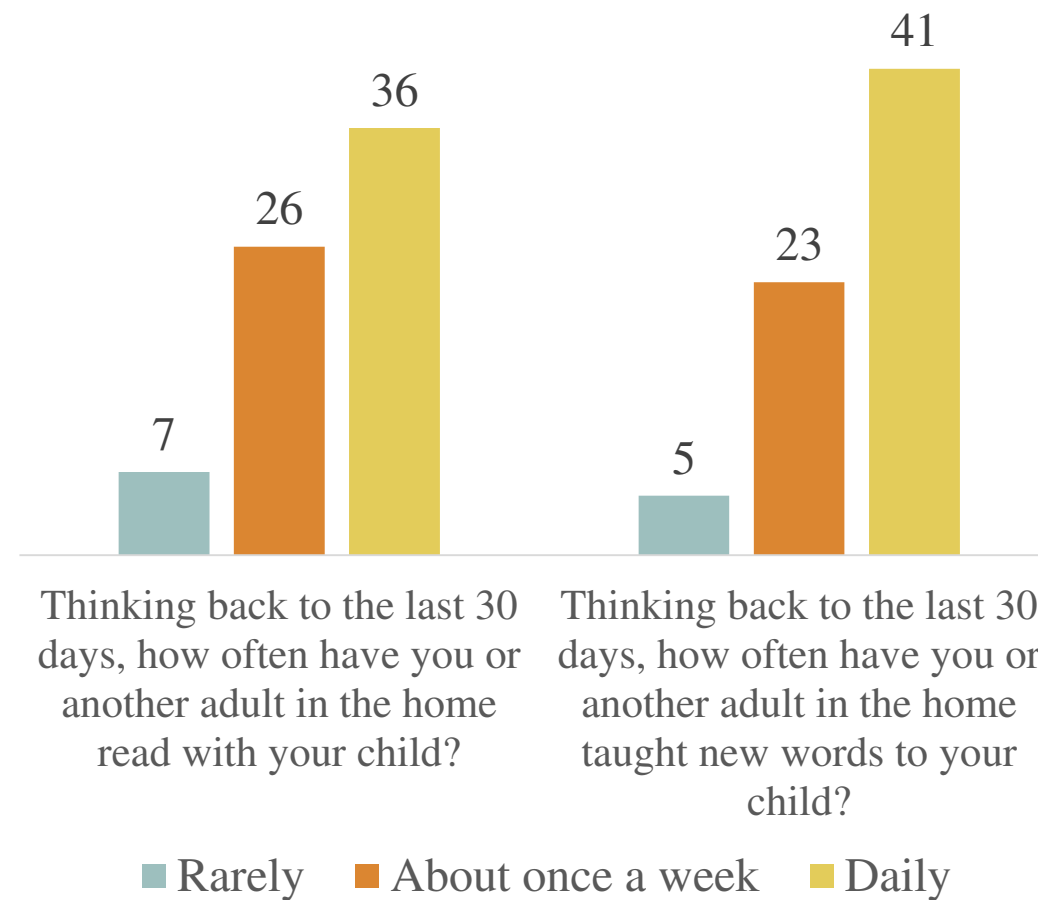
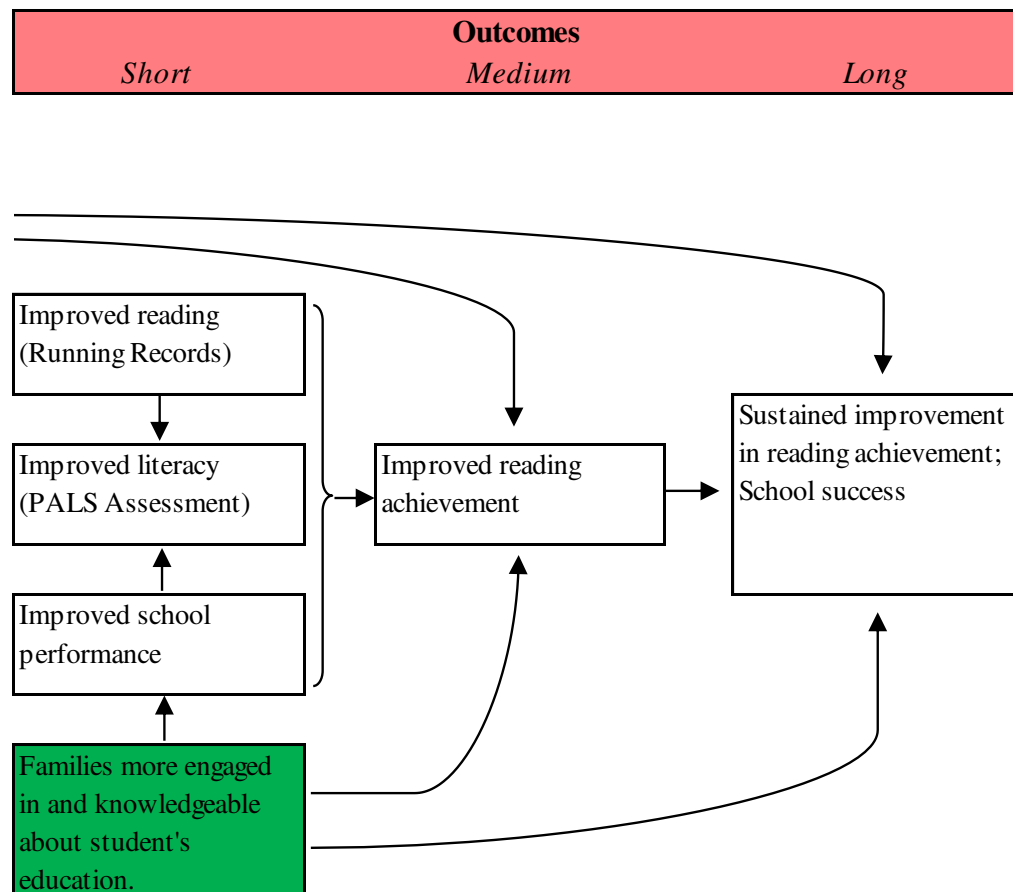
The primary goal of the evaluation was to design a study that would meet What Works Clearinghouse (WWC) design standards without reservations.

- Studies have to use a randomized-control framework to qualify.
 - ✓ 576 students were randomly assigned to either the control (290) or SPARK (286) group by the evaluation team. Assignments were stratified by school and grade level within schools. ELL and IEP students were not eligible for study.
- Studies have to use valid and reliable outcome measures.
 - ✓ NWEA Measures of Academic Progress (MAP) Reading Benchmark Assessment,
 - ✓ Phonological Literacy Screening (PALS),
 - ✓ Regular-school-day attendance.
- Studies have to meet WWC attrition and differential attrition standards.
 - ✓ Study attrition results were within acceptable ranges. In addition, attrition was exogenous to the program. Thus, the evaluation should meet WWC attrition standards.
- The program has to be implemented as intended.
 - ✓ Nearly all students and families received an intensive amount of tutoring and family engagement.

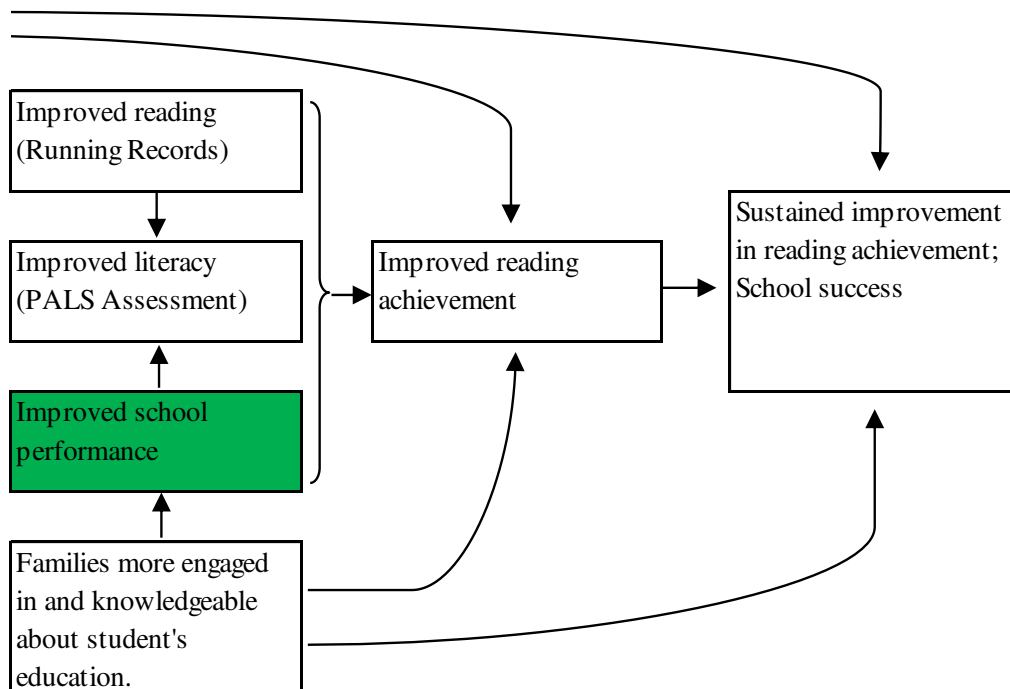
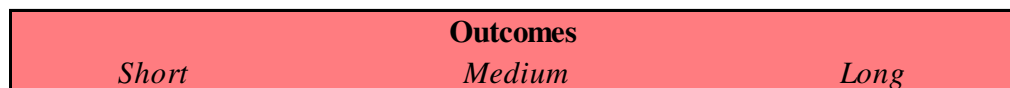
Outcomes/Impact Analysis



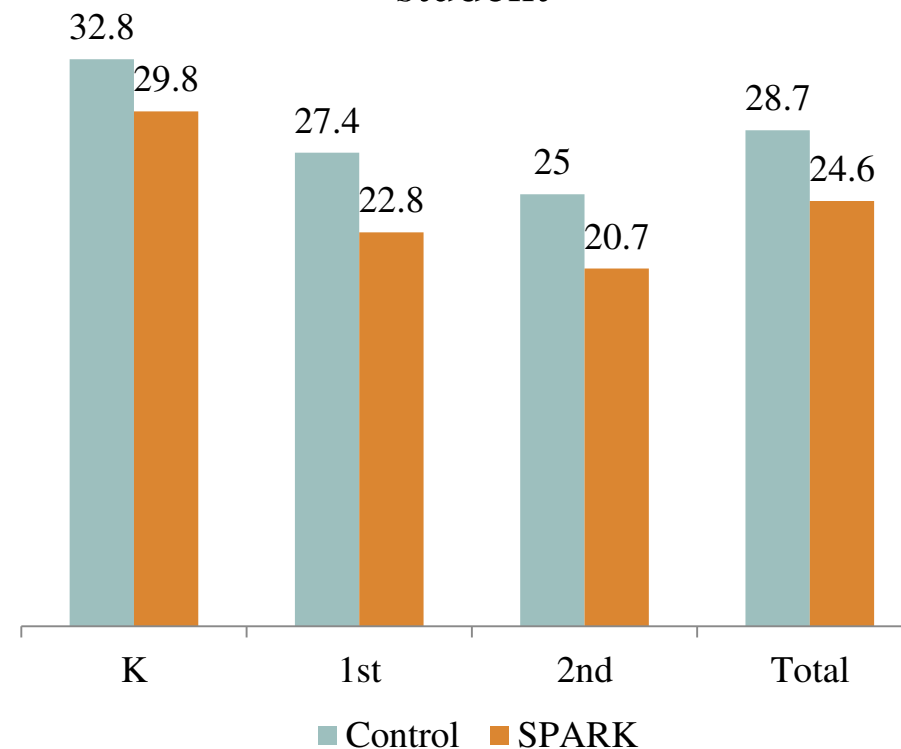
Short-term Outcome - Engaged Families



Short-term Outcome – School Performance (Attendance)



Unadjusted number of absences per student



Short-term Outcome – School Performance (Attendance)

Adjusted Impact of SPARK on Attendance (Number of Absences)

		β	SE	t-stat	p-value
Unstandardized	K-1 st	-5.6	4.5	1.23	
	1 st -2 nd	-5.3	3.9	1.36	
	2 nd -3 rd	-6.5	4.2	1.55	
	Overall	-5.8	2.4	2.39	<.05
Standardized	K-1 st	-0.20	0.16	1.23	
	1 st -2 nd	-0.26	0.19	1.36	
	2 nd -3 rd	-0.36	0.23	1.55	
	Overall	-0.25	0.11	2.34	<.05

K – 1st and 1st – 2nd models

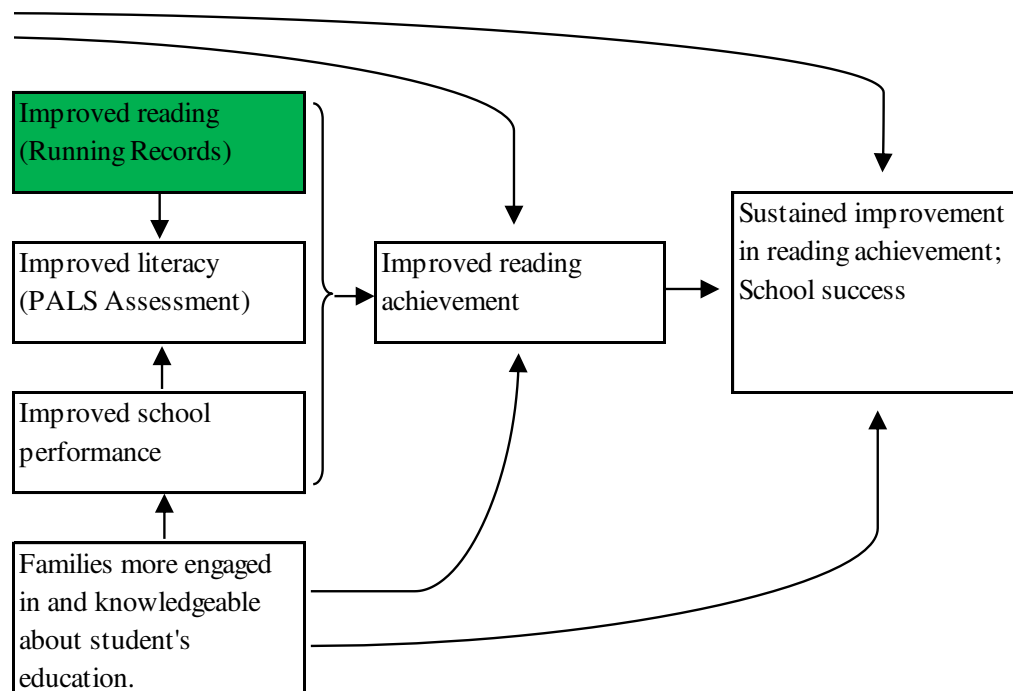
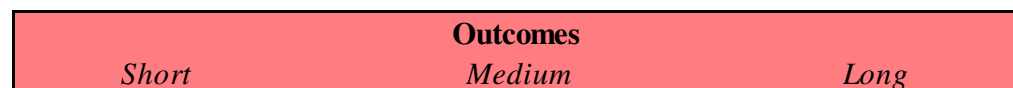
$$Absences_{ij} = \beta_{1ij}SPARK_{ij} + \beta_{2ij}2013PALS_{ij} + \beta_{3ij}Freelunch_{ij} + \sum_{j=1}^J(\beta_{4j}school_j + \beta_{5ij}(2013PALS_{ij})(school_j)) + \varepsilon_{ij}$$

2nd – 3rd model

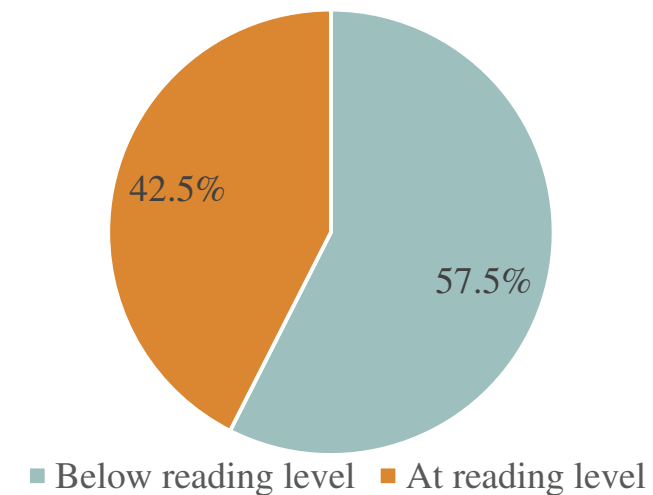
$$Absences_{ij} = \beta_{1ij}SPARK_{ij} + \beta_{2ij}2013MAPMath_{ij} + \beta_{3ij}MAPRead_{ij} + \sum_{j=1}^J\beta_{4j}school_j + \varepsilon_{ij}$$

The pooled interaction term of SPARK participation and baseline PALS scores was not significant ($t = 1.04$, $p > .05$) suggesting that SPARK did not differentially impact students with different starting literacy levels.

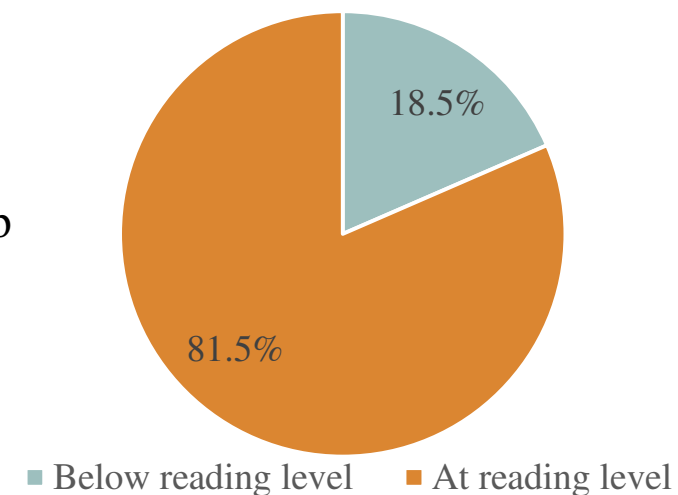
Short-term Outcome – Reading Levels



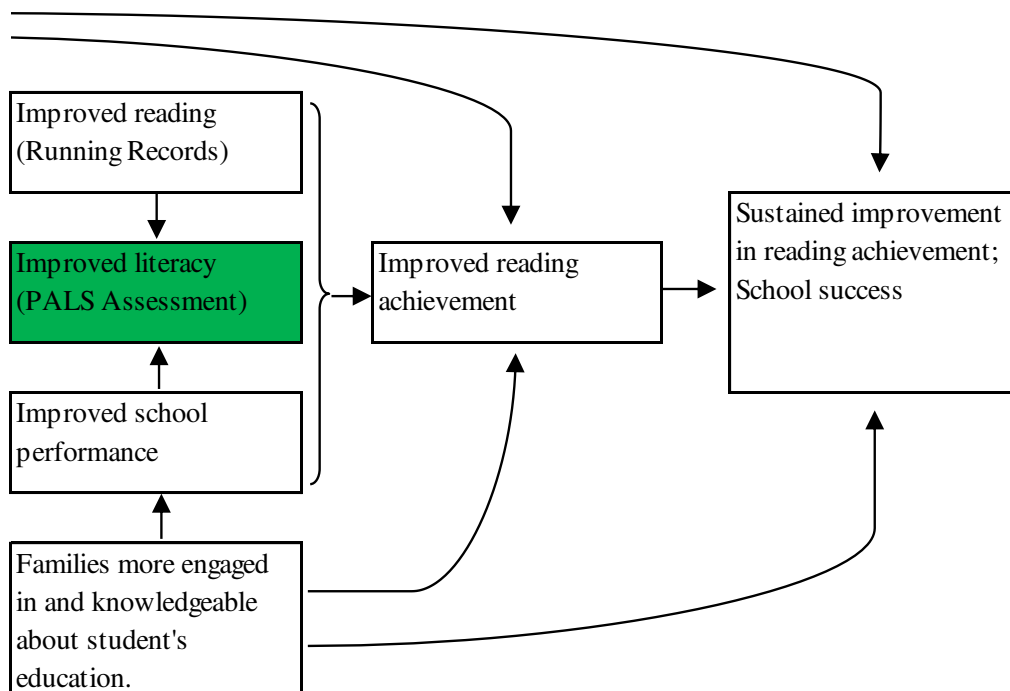
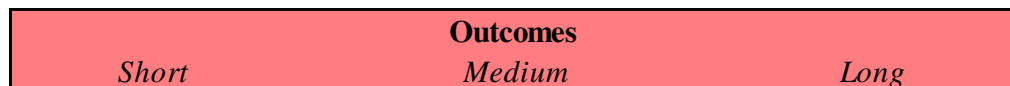
Baseline



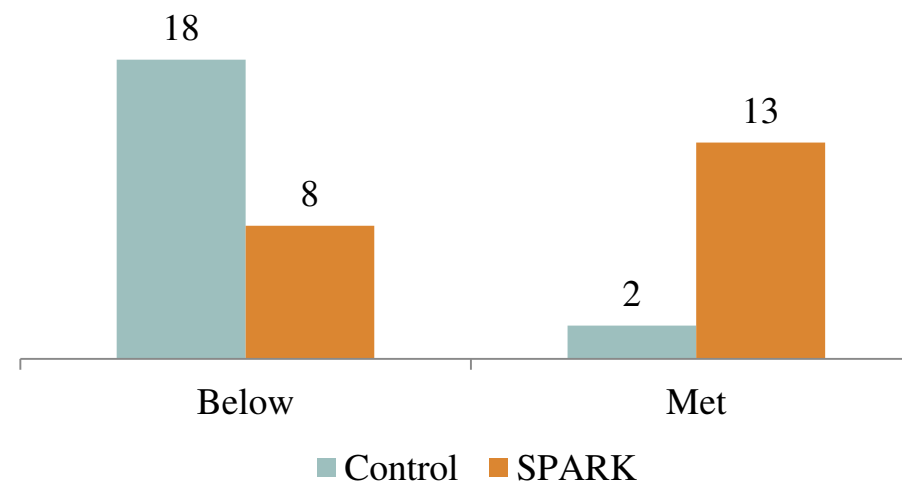
Follow-up



Short-term Impact – PALS



Post Benchmark Status of Students who Started Below Benchmark



Short-term Impact - PALS

Adjusted Standardized Impact of SPARK on PALS scores

		β	SE	t -stat	p -value
Overall	K-1 st	0.34	0.14	2.45	
	1 st -2 nd	0.55	0.16	3.54	
	2 nd -3 rd	0.16	0.16	0.95	
	Overall	0.36	0.09	4.05	<.001
Students in lower half of score distribution at baseline	K-1 st	0.48	0.23	2.08	
	1 st -2 nd	1.14	0.27	4.24	
	2 nd -3 rd	0.14	0.40	0.34	
	Overall	0.66	0.16	4.11	< .001
Students in higher half of score distribution at baseline	K-1 st	-0.05	0.17	-0.29	
	1 st -2 nd	0.06	0.16	0.38	
	2 nd -3 rd	0.23	0.11	2.02	
	Overall	0.13	0.08	1.49	> .05

K – 1st and 1st – 2nd models

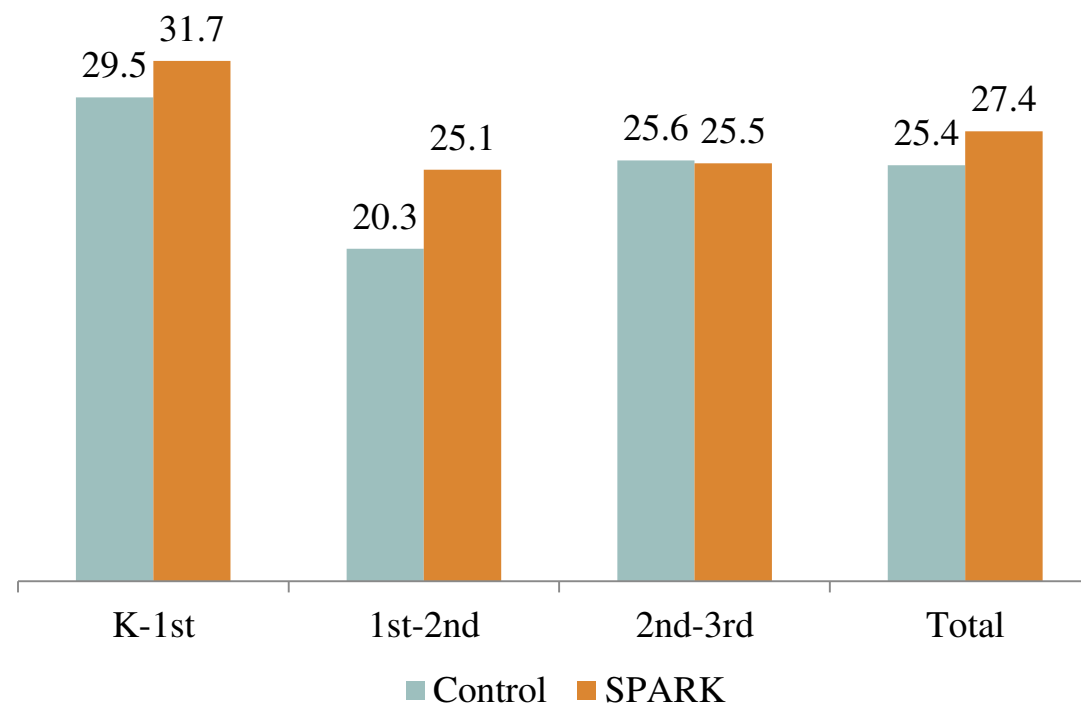
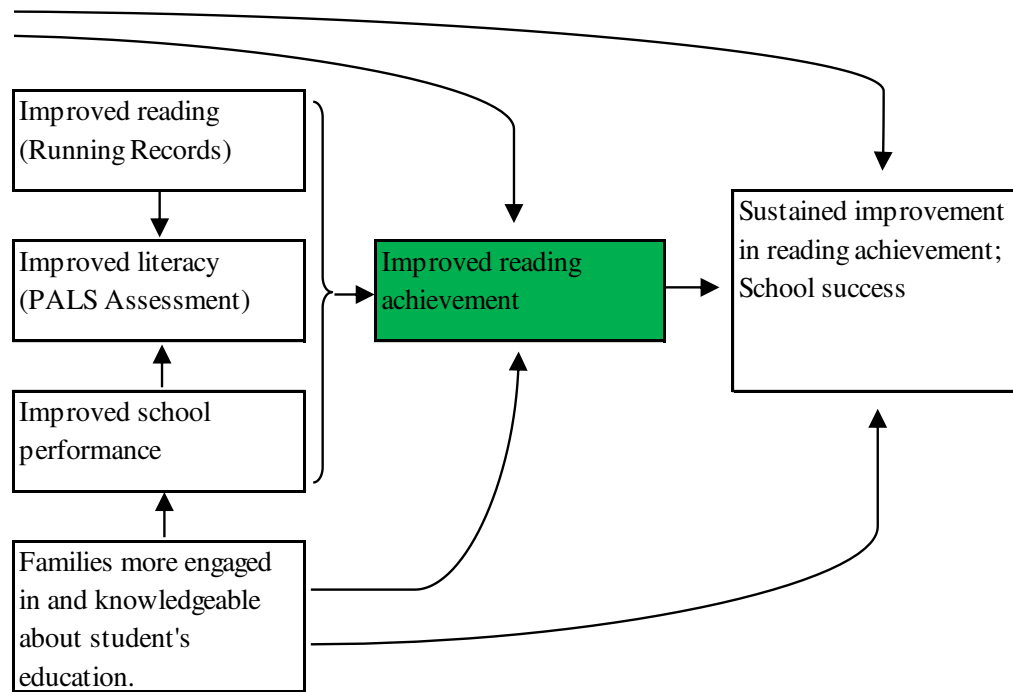
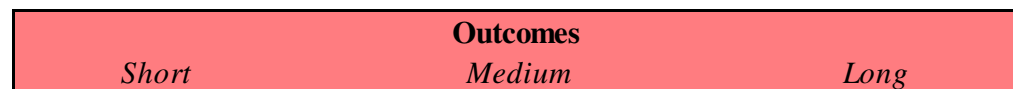
$$2015PALS_{ij} = \beta_{1ij}SPARK_{ij} + \beta_{2ij}2013PALS_{ij} + \beta_{3ij}2013MAPRead_{ij} + \sum_{j=1}^J(\beta_{4j} school_j + \beta_{5ij}(2013PALS_{ij})(school_j) + \beta_{6ij}(2013MAPRead_{ij})(school_j)) + \varepsilon_{ij}$$

2nd – 3rd model

$$2015PALS_{ij} = \beta_{1ij}SPARK_{ij} + \beta_{2ij}2013MAPMath_{ij} + \beta_{3ij}2013MAPRead_{ij} + \sum_{j=1}^J(\beta_{4j} school_j + \beta_{5ij}(2013MAPMath_{ij})(school_j) + \beta_{6ij}(2013MAPRead_{ij})(school_j)) + \varepsilon_{ij}$$

The pooled interaction term of SPARK participation and baseline PALS scores was significant ($t = 5.61$, $p < .001$) suggesting that SPARK differentially impacted PALS scores for students with different starting literacy levels.

Medium-term Impact – Reading Achievement (MAP)



Medium-term Impact - MAP

Adjusted Standardized Impact of SPARK on 2015 MAP RIT Reading scores

		β	SE	t-stat	p-value
Overall	K-1 st	0.18	0.13	1.34	
	1 st -2 nd	0.39	0.13	3.04	
	2 nd -3 rd	0.09	0.14	0.65	
	Overall	0.23	0.08	2.95	<.01
Students in lower half of score distribution	K-1 st	0.23	0.24	0.94	
	1 st -2 nd	0.58	0.23	2.54	
	2 nd -3 rd	0.19	0.28	0.68	
	Overall	0.36	0.14	2.49	< .05
Students in higher half of score distribution	K-1 st	0.11	0.18	0.60	
	1 st -2 nd	0.43	0.20	2.15	
	2 nd -3 rd	0.14	0.26	0.55	
	Overall	0.23	0.12	1.92	< .10

K – 1st and 1st – 2nd models

$$2015MAPRead_{ij} = \beta_{1ij}SPARK_{ij} + \beta_{2ij}2013PALS_{ij} + \beta_{3ij}2013MAPRead_{ij} + \sum_{j=1}^J(\beta_{4j} school_j + \beta_{5ij}(2013PALS_{ij})(school_j) + \beta_{6ij}(2013MAPRead_{ij})(school_j)) + \varepsilon_{ij}$$

2nd – 3rd model

$$2015MAPRead_{ij} = \beta_{1ij}SPARK_{ij} + \beta_{2ij}2013MAPMath_{ij} + \beta_{3ij}2013MAPRead_{ij} + \sum_{j=1}^J(\beta_{4j} school_j + \beta_{5ij}(2013MAPMath_{ij})(school_j) + \beta_{6ij}(2013MAPRead_{ij})(school_j)) + \varepsilon_{ij}$$

The pooled interaction term of SPARK participation and baseline MAP scores was significant ($t = 2.17, p < .05$) suggesting that SPARK differentially impacted MAP scores for students with different starting achievement levels.

Robustness of results: Impact on absences

		<i>B</i>	<i>SE</i>	<i>t</i> -stat	<i>p</i> -value
Number of absences	K-1 st	-5.6	4.5	1.23	
	1 st -2 nd	-5.3	3.9	1.36	
	2 nd -3 rd	-6.5	4.2	1.55	
	Overall	-5.8	2.4	2.39	<.05
Chronic absence (≥10% absent)	K-1 st	-.10	.07	1.32	
	1 st -2 nd	-.11	.08	1.39	
	2 nd -3 rd	-.08	.08	1.04	
	Overall	-.10	.05	2.17	<.05
Severe chronic absence (≥20% absent)	K-1 st	-.04	.06	0.71	
	1 st -2 nd	-.05	.04	1.30	
	2 nd -3 rd	-.08	.04	1.88	
	Overall	-.06	.02	2.29	<.05

Measured impact of SPARK on absences is robust to alternative absence metrics

- Linear probability model with the LHS variable an indicator for chronic (≥10%) or severe chronic (≥20%) absence
- Indicates that SPARK has a specific effect at margins of chronic, severe chronic absence
- Results robust to probit
- While average effect is statistically significant, joint hypothesis that individual grade effects are all zero is not rejected regardless of LHS variable

Robustness of results: Impact on assessments

Model	MAP		PALS	
	<i>B</i>	<i>SE</i>	<i>B</i>	<i>SE</i>
Baseline	0.23	0.08	0.36	0.09
Baseline, assessments as rank-based z-scores	0.25	0.08	0.34	0.08
Single pretest and treatment as only RHS variables	0.13	0.09	0.31	0.08
Two pretests, school FEs, treatment as only RHS	0.15	0.07	0.31	0.08
Baseline plus math MAP pretest, cubic pretest terms (incl. interactions with each other, school FEs)	0.18	0.08	0.28	0.09

Measured impact of SPARK on assessments is robust to alternative specifications, especially in the case of PALS literacy assessment

- Transforming scale of assessment to rank-based z-scores, forcing normality
- Simpler specification with only the pretest and treatment as RHS variables
- More complex specification including math pretest, cubic pretest terms, interactions among linear and quadratic pretest terms, interactions between cubic pretest terms and school fixed effects

Joint hypothesis testing between MAP and PALS

Estimating models for MAP and PALS jointly as seemingly unrelated regressions (SUR) allows testing of joint hypotheses across MAP and PALS

MAP and PALS models use same RHS variables, so joint estimation yields the same results as equation-by-equation

- Results still slightly different from baseline because analysis only included students with both MAP and PALS scores, model assumes i.i.d. standard errors

Substantially rejects null that effect of SPARK across MAP and PALS is zero

- $\beta_{(\text{MAP,overall})} = \beta_{(\text{PALS,overall})} = 0$ rejected at .0001 level
- $\beta_{(\text{MAP,K1})} = \beta_{(\text{PALS,K1})} = \beta_{(\text{MAP,12})} = \beta_{(\text{PALS,12})} = \beta_{(\text{MAP,23})} = \beta_{(\text{PALS,23})} = 0$ rejected at .002 level

Joint hypothesis testing between MAP and PALS (ct'd)

Also substantially rejects null hypothesis that overall effect of SPARK across MAP and PALS is the same for high-achieving and low-achieving students

- $\beta_{(\text{MAP, high, overall})} - \beta_{(\text{MAP, low, overall})} = \beta_{(\text{PALS, high, overall})} - \beta_{(\text{PALS, low, overall})} = 0$ rejected at .01 level
- Analogous null hypothesis for individual grades rejected at .005 level

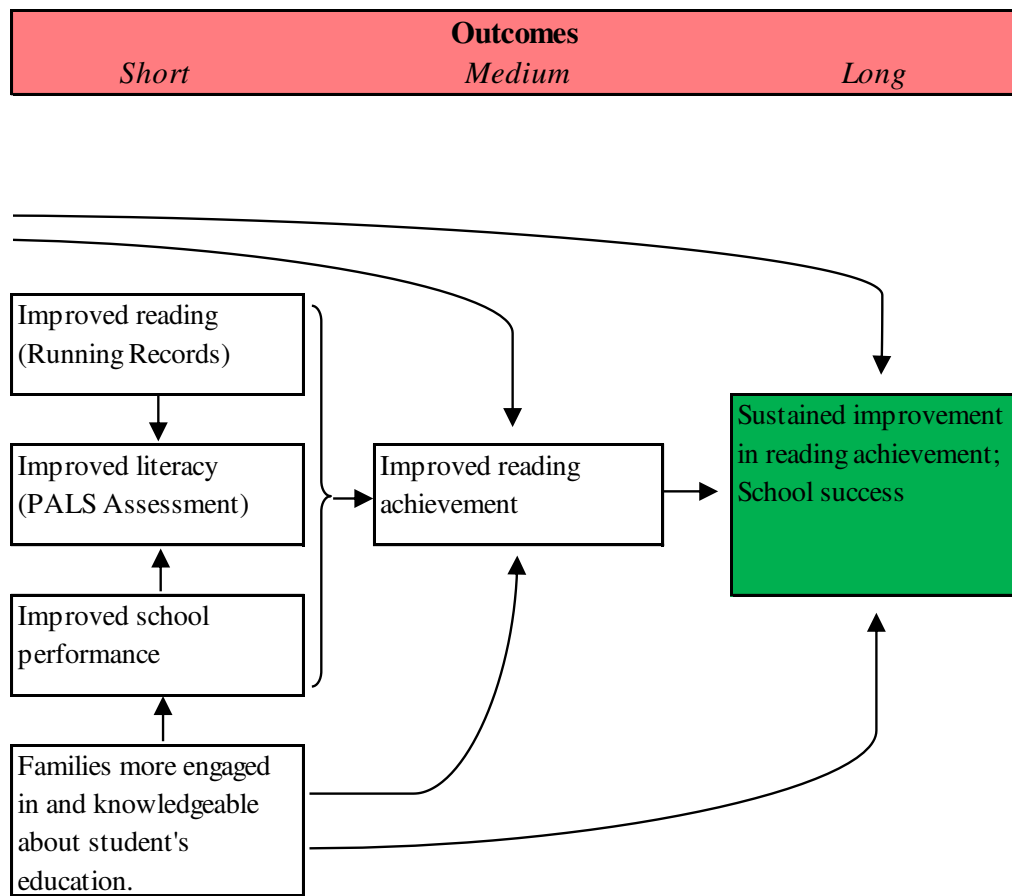
Interpreting standardized effects

Standard deviation of fall MAP scale scores from SPARK sample are 15% to 25% smaller than national norms

- Effect of SPARK in nationwide standard deviation units is approximately 75% to 85% the effect in SPARK sample standard deviation units

Standard deviation of scale scores				
	Fall MAP pretest		Spring MAP posttest	
	SPARK Sample	MAP Norm	SPARK Sample	MAP Norm
K-1 st	10.1	13.5	11.3	14.5
1 st -2 nd	10.9	13.1	15.4	15.2
2 nd -3 rd	11.9	15.5	13.6	15.1

Long-term Outcome – Sustained Reading and School Success



Typical literacy interventions focus mostly on skill acquisition and ignore the reasons why students fall behind.

- SPARK supports the whole child and family and thus is designed to address some of the reasons why students are struggling to learn to read.

The benefits observed in literacy interventions typically fade over time (Suggate, 2016).

- By engaging the whole family and by continuing to work with students and families beyond SPARK, BGCGM is in a unique position to maintain and expand the impact of SPARK.

The evaluation will continue to work with BGCGM to follow SPARK students as they progress in their education and will measure the sustained impact of SPARK over time.

Thank you!

For more information about the SPARK evaluation contact Curtis Jones at:

jones554@uwm.edu

312.421.0277