The Interim Results of a Randomized Control Trial of the SPARK Early Literacy Program/Milwaukee Community Literacy Project

Curtis J. Jones
Rachel Lander
Rachel Westrum
Socially Responsible Evaluation in Education
University of Wisconsin, Milwaukee

Presented at the Annual Meeting of the Society for Research in Educational Effectiveness, Washington, DC

March 3, 2015

Funding by a Department of Education Investing in Innovations (i3) grant.
Contract number U396C100694

The MCLP/SPARK Program

SPARK was created by Boys & Girls Clubs of Greater Milwaukee in 2005 and piloted at one site in Milwaukee Public Schools (MPS) financed by the United Way. It was then launched in three MPS schools in 2006 with funding from the United Way and AmeriCorps. In 2010, SPARK received a Department of Education investing in innovation (i3) grant award. At this time, SPARK expanded to 10 MPS schools (relying on funds from i3, the United Way and AmeriCorps). Boys & Girls Clubs of America was awarded a 2-year Department of Education Innovative Approaches to Literacy (IAL) award in 2012, for expansion of the SPARK model to 14 schools across six states.

Figure 1 presents the SPARK program model. SPARK students are tutored during the school day for 30 minutes, up to three times per week, for two years. At each site, a program manager, who is also a certified teacher, oversees the tutors. The tutoring component of the Milwaukee Community Literacy Project (MCLP) / SPARK program is loosely based on the Reading Recovery program. Reading Recovery focuses on in-school tutoring with lesson plans written, and assessments analyzed, by the tutors themselves. In a comparison of early literacy intervention programs, Pinnell et al found that Reading Recovery subjects performed significantly better than any other treatment and comparison group on all measures (Pinnell, Lyons, DeFord, Bryk, & Seltzer, 1994).

The MCLP also contains a strong family engagement component. Involving families in tutoring programs can improve children's academic knowledge, skills and confidence (Bryan, 2005; Harvard Family Research Project, 2009). Encouraging family involvement in educational programs traditionally focuses on families attending events, receiving information from staff, volunteering (Epstein, 2001), and generally exhibiting "good parent" behaviors (Li, 2010).

Getting to know families and the ways that their lives are structured outside of the educational setting may lead to a reciprocal relationship that can increase involvement (Graue & Hawkins, 2010). The family component of a program is not only to make families aware of the program's mission and goals but also to empower families in their child's learning both in the program and at home.

To execute the family engagement component, each site has a parent partner who works with each participating student's family. Their work is designed to bridge the divide between school and home by translating literacy concepts, educating families about a variety of literacy activities, and validating the literacy practices already happening in the home. Parent partners help families see how they already are incorporating literacy into their children's lives and show parents how to promote literacy more effectively. Parent partners stay connected with families through a monthly newsletter, monthly family events at each site, and phone calls or emails. These communications are designed to keep families aware of student progress in MCLP, help families promote literacy at home, and address any attendance issues that arise during the program. Parent partners also conduct home visits for all students twice during the summer between their first and second year of participation. These visits are viewed as opportunities to connect with the family in their own space and learn about the literacy activities already taking place in the home.

Assessment is also an important component of the MCLP/SPARK program. All students are assessed with the PALS (Phonological Awareness Literacy Screening) at the beginning and end of the school year. This assessment is used to determine each student's needs and help create individual lesson plans. Tutors participate in trainings at the beginning of the year, as well as follow-up training several times throughout the year. In addition, program managers provide ongoing support and training to tutors throughout the year.

In-School Tutoring Lesson Plan

Each lesson includes six sections: Familiar Activity, Running Record, Word Play, Reading at Instructional Level, Writing, and a Tutor Read Aloud.

- The *Familiar Activity* is a brief element that gets the student ready for learning by reviewing a skill they have recently learned.
- *Running Record* is the assessment students take to monitor their reading level progress. *Word play* is a key element in the lesson where students receive targeted, differentiated instruction on foundational reading skills including phonics and phonemic awareness. Word plan is individualized to focus on students' needs. It is centered on two main activities: Word Sorts and Making Words. These activities combine both constructivist learning and structured instruction. Each of these activities focus on specific skills and tutors are explicit with students about the lesson's foci. Word Sorts involve students

sorting words into various categories to increase their understanding of the structure of sounds and letters. Making Words involves students using different letters to make words and provides a structured way for students to learn how the sounds of language are put together. Students also read phonics-based books during Word Play time and do enrichment activities to cement their understanding of the focus skills.

- During every SPARK lesson, students spend time R*eading a book at their instructional level*. Before reading the book, they do a book walk to familiarize themselves with the content and vocabulary of the book. Students read both fiction and non-fiction books. As students read, tutors use a variety of strategies to help students decode and make meaning of text. Students use graphic organizers to build comprehension skills.
- Students spend time each lesson *Writing* sentences connected to their Word Play skill or their instructional reading book. Tutors help students correctly spell the words in their sentence(s). Elkonin boxes are a central piece of SPARK writing and used to help students encode words. Elkonin boxes are an instructional method used in early elementary grades to build phonological awareness by segmenting words into individual sounds/boxes.
- The lesson ends with a brief opportunity for students to hear their tutor *Read a book*.

Setting

The Milwaukee Community Literacy Project is a collaboration between the Boys & Girls Clubs of Greater Milwaukee and Milwaukee Public Schools (MPS) to provide SPARK to kindergarten through 3rd grade students in seven predominantly low-income and minority elementary schools.

MPS, a district serving over 80,000 students, faces a significant challenge to teach its students how to read and write. According to the Wisconsin Knowledge and Concepts Examination (WKCE) only 15% of MPS students were proficient in reading (2011). This is in comparison to 35% statewide. The results of the WKCE are consistent with results of the National Assessment of Educational Progress (NAEP) and the ACT, which show that MPS students struggle with literacy throughout their education. 15% of 4th grade MPS students are proficient in reading (NAEP, 2011) while 14% of MPS 11th graders scored at least 21 on the ACT Reading Test, the benchmark identified for college readiness (Independent Analysis).

The results of the NAEP further show that there are significant achievement gaps for minority and low-income students. 39% of 4th grade, White MPS students are proficient in reading, compared to 7% of Black and 15% of Hispanic students. 7% of 4th grade low-income (free/reduced lunch participants) MPS students are proficient in reading, compared to 48% of non-low-income students. These data demonstrate that the need for increased literacy

opportunities in the Milwaukee area is urgent, and that this need is even more pronounced for low-income and minority students.

The Evaluation of the MCLP/SPARK Program

The objective of the evaluation is to measure the implementation and impact of the MCLP on reading achievement in two cohorts of participants. One cohort participated for two years in the 2011-2012 and 2012-2013 school years. The other began participation in the 2013-2014 school year and is continuing to participate in 2014-2015.

Participation and Attrition

Cohort 1:

In 2011, 496 students consented to participate in the first cohort. Of these, 245 were randomly selected to participate in the MCLP and 251 were selected as control students. Of the 496 students, 165 (33%) were in kindergarten, 177 (36%) in first, and 154 (31%) in second. 480 (97%) were eligible for free or reduced lunch, 404 (81%) were African American, and 57 (12%) were Hispanic.

Attrition was a problem during the two years that students participated. 222 students were excluded from the final analysis due to attrition, which represents a 44.7% overall attrition rate. However, it is important to note that students were excluded for exogenous reasons, like not taking the pre-test (1 students), moving away (189), being identified for a reading disability (30), and not taking the post-test (2). Regarding differential attrition, 110 (44.9%) of participants and 112 (44.6%) of control students dropped out. The 0.3% differential attrition rate, along with the 44.7% overall attrition, and the exogenous nature of why students were dropped, suggests that the internal validity of the cohort 1 evaluation remained intact and the study should qualify for inclusion in the What Works Clearinghouse without reservation.

Cohort 2:

A total of 576 students across seven schools consented to participate in the second MCLP cohort. 286 students were randomly selected as MCLP participants and 290 as control students. Of the 576 students, 205 (36%) are in kindergarten, 214 (37%) in first, and 157 (27%) in second. 549 (95%) are eligible for free or reduced lunch, 459 (80%) African American, 71 (12%) Hispanic, and 51 (9%) have an IEP for speech/language. During the first year of participation, only 47 (8.1%) students dropped, all because they moved to another school. This included 20 (7%) participants and 27 (9%) control students. A lower percentage of students dropped by this point in the second cohort than did in the first (10.5%). Second cohort selection processes were

changed in an attempt to mitigate attrition. One strategy used to reduce attrition was to wait until after the baseline assessment was administered to select students so that all students would have baseline assessment scores. Also, many students move during the first month of the year, so waiting until after the assessment prevented these students from being enrolled in the evaluation and then dropped. We also identified students from the previous year that had changed schools, and excluded them from the evaluation. These students were still allowed to be tutored outside of the evaluation.

Measuring Fidelity of Implementation

Implementation is measured according to the outputs and key elements contained within the MCLP/SPARK logic model (Figure 1). Students are tutored three times per week for approximately 60 weeks across two school years. However, due to holidays, teacher planning days, field trips, etc, 180 sessions is not possible. Thus, students tutored fewer than 90 times received a low level of SPARK tutoring, students tutored 90 up to 120 times received a moderate level, and students tutored at least 120 times are considered to have received a high level of tutoring. Implementation at a site is rated as high when at least 70% of students receive a high level, and moderate when at least 70% receive at least a moderate level. When more than 30% of students receive a low level of tutoring, the entire in-school tutoring component for that site is rated as a low. Across sites, the SPARK in-school tutoring component is rated as high when at least 70% of sites are rated as at least moderate. Anything less than that and the initiative is rated as low.

For the family engagement component, each family should receive two home visits, 12 additional contacts, and at least 14 newsletters. These data were not consistently tracked in the first cohort so levels of family component implementation could not be determined. However, these data are currently being tracked for the second cohort.

Cohort 1 Implementation Results

Participation data were available for 130 students across the six sites that completed two years in SPARK.[*] These students averaged 120.7 tutoring sessions, with 64.5 in the first year, and 56.3 in the second (Table 1). Although kindergarten students received more tutoring on average than students in other grade levels, the difference was not significant ($p = .223$).

Further, the distribution of total tutoring suggests that most participants received a high amount of tutoring (Figure 2). 76 (58%) participants received a high level (120 or more tutoring sessions) of tutoring, while 45 (34%) received a moderate level (90 to 119), and only nine (7%) received a low level (lower than 90).

[*] Data for eight students that had moved away and had returned to the school are included in the impact analysis but not in the implementation analysis.

Sites ranged from an average of 114 sessions per student to 129 sessions (Table 2). These differences were statistically significant ($F = 2.29$, $p = .049$), suggesting that students in different schools received differing levels of tutoring.

Further, the breakdown within each site of the percent of students receiving different levels of tutoring (Figure 3) suggest that two sites provided a high level of tutoring, with more than 70% of students receiving at least 120 tutoring sessions (81st and Brown), while the other four provided a moderate level. Since all six sites provided at least a moderate level of tutoring, the entire in-school tutoring component across the SPARK initiative is rated as high.

Impact Evaluation Design

The evaluation utilizes a randomized control trial to isolate the impact of MCLP on reading achievement. A random selection of kindergarten, 1st, and 2nd grade students in six MPS schools were selected to participate in Cohort 1 during the 2011-2012 school year. A second cohort was selected to participate in seven schools during the 2013-2014 school year. Students were randomly selected in the fall of each year and stratified by school and grade level within school. Students with a reading-related IEP or who were English Language Learners, were not eligible for participation in the evaluation but were eligible to receive tutoring.

The primary outcome used to evaluate the MCLP is the Measures of Academic Progress (MAP) reading assessment published by the Northwest Evaluation Association (NWEA). The MAP is an on-line, adaptive assessment that has been shown to be both valid and reliable. MPS administers the MAP to all students three times each year, in the fall, winter, and spring. As such, to measure the impact of the MCLP on the first cohort, the evaluation was able to use the fall of 2011 MAP results as the pre-test and the spring of 2013 results as the post-test. To measure the impact of the MCLP on the second cohort, the evaluation uses the fall of 2013 MAP results as the baseline and will use the spring of 2015 results as the post-test. The spring of 2014 MAP results are used to measure interim results.

MPS uses the MAP for Primary Grades (MPG) for kindergarten, 1st grade, and 2nd grade students. 3rd grade students take the MAP. The MPG is vertically aligned to the MAP, so that the same score on both assessments suggests the same reading achievement level. The main difference between the two is that the MPG includes auditory technology to help students complete the assessment.

The statistical analysis of the MCLP follows an Intent to Treat Model, where students selected to participate in the MCLP are included in the analysis regardless of how much tutoring they received. This is done to maximize the internal validity of the study. However, due to missing data, both participant and control students who moved away or were identified as having a disability that prevented them from receiving literacy tutoring were excluded from the analysis.

Data Collection and Analysis

The project has a data sharing agreement with the Milwaukee Public Schools (MPS) that allows it to access achievement results (MPG, MAP, PALS) and student demographics. The project itself maintains participation records and fidelity of implementation results.

To analyze the impact of the MCLP, the evaluation used separate generalized linear statistical models with robust standard error estimators to compare the reading achievement growth of participants and controls for kindergarten, first, and second grade students. The results of the three models were then pooled to estimate the overall impact of the MCLP for each cohort. Post reading achievement scores were standardized to improve interpretability.

For cohort 1, all three grade-level models predicting post MAP/MPG scores controlled for the separate fixed interactions of school effects with both baseline fall 2011 MAP/MPG reading and MAP/MPG math results. The fixed effects of gender, race, disability status, and free/reduced lunch eligibility were also tested in the model. Ultimately, only the main effect of race was found to uniquely predict MAP/MPG reading results and was included in the model. The rest were excluded because they were not found to uniquely predict post-test reading achievement.

For cohort 2, all three grade-level models predicting post MPG scores controlled for the fixed interaction of baseline MPG reading and school effects. The kindergarten and first grade models also controlled for the fall results of the Phonological Awareness Literacy Screener (PALS), used for the first time by MPS with its kindergarten and first grade students. MPG math baseline scores was considered as a covariate, but its inclusion in the kindergarten and first grade models did not reduce standard errors nor improve model fit, so it was not included. However, MPG math baseline scores did improve model fit and reduce standard errors for second grade students, so it was included in that model. Spring 2014 PALS results were also tested for first and second grade students using the same model.

Results of Impact Evaluation

After two years of participation, the MCLP was found to have a small but significant impact on the reading achievement of the first cohort (0.12 standard deviations) (Table 3). However, after only one year of participation, the MCLP was found to have a similar impact on the MAP/MPG scores of the second cohort (0.12 standard deviations) (Table 4). Second cohort MCLP participants also scored higher on the PALS (Table 5); scores were .438 and .372 standardized units higher than control students ($p < .0001$). Pooled together, these results indicate that the MCLP had a large impact (.40 standardized effect) on spring PALS scores ($p < .0001$).

Conclusions

The results of the evaluation of the MCLP are promising. The implementation results for the first SPARK cohort (2011 to 2013) suggest that SPARK was provided at a high intensity to the majority of students and that each site provided at least a moderate level of tutoring to nearly all participants. That the average student received 120.7 tutoring sessions suggests that SPARK tutoring was effectively integrated into the schools. Implementation data are continuing to be monitored for the second cohort of SPARK participants for both the family and the tutoring components.

In relation to effectiveness, the MCLP was found to have a statistically significant impact on the reading development of participants. Although attrition was a problem for the first cohort, there is no evidence that it affected the internal validity of these findings. There is also evidence that changes to the delivery of the MCLP from the first to the second cohort, improved its effectiveness. Specifically, the program revised its approach to working with kindergarten students to account for the evaluation results.

These results suggest that the MCLP/SPARK is a promising program for helping to address the serious challenge facing the Milwaukee Public Schools of teaching students to read.

References

Bryan, J. (2005). Fostering educational resilience and achievement in urban schools through school-family-community partnerships. *Professional School Counseling*, *8*(3), 219.

Graue, E. and Hawkins, M. (2010).”I always feel like they don’t know anything about us”: Diverse families talk about their relations with school. In Miller-Marsh, M. and Turner-Vorbeck, T. (Eds.) *(Mis)Understanding Families: Learning from real families in our schools* (pp.109-125) New York, NY: Teachers College Press.

Harvard Family Research Project. (2009a). *Supporting student outcomes through expanded learning opportunities*. Cambridge, MA: Little, P. M. D.

Epstein, J.L. (2001). *School family, and community partnerships: Preparing educators and improving schools.* Boulder, CO: Westview Press.

Li, G. (2010). Social class, culture and “good parenting”: Voices of low-SES families. In Miller-Marsh, M. and Turner-Vorbeck, T. (Eds). *(Mis)Understanding Families: Learning from real families in our schools* (pp.162-178). New York, NY: Teachers College Press.

Lopez, G.R. & Stoeling (2010). Disarticulating parent involvment in Lation-impacted schools in the Midwest. In Miller-Marsh, M. and Turner-Vorbeck, T. (Eds). *(Mis)Understanding Families: Learning from real families in our schools* (pp.19-36). New York, NY: Teachers College Press.

Nieto, S. (2012). Honoring the lives of all children: Identity, culture, and language. In B. Falk (ed.) *Defending Childhood: Keeping the Promise of Early Education* (pp. 48-62). New York, NY: Teachers College Press.

Pinnell, G. S., Lyons, C. A., DeFord, D. E., Bryk, A., & Seltzer, N. (1994). *Reading Research Quarterly, 29*, 8-39.

Tables and Figures

Table 1: Descriptive statistics of tutoring sessions by grade level

	Mean	SD	Min	Max	N
K to 1st	124.2	14.7	87	155	45
1st to 2nd	119.7	20.6	61	151	44
2nd to 3rd	118.0	16.9	72	155	41
Total	120.7	17.6	61	155	130

Table 2: Descriptive statistics of tutoring sessions by school site

	Mean	SD	Min	Max	N
81st	125.1	15.1	88	146	19
Brown	129.3	15.1	102	155	24
Cass	118.6	20.4	65	140	18
Clarke	114.3	15.7	83	147	18
Rogers	116.7	19.2	61	144	22
Sherman	119.2	17.0	84	151	29

Table 3: MCLP 1st Cohort Final MAP/MPG Results

	Standardized Effect	Robust Standard Errors	p-value
MCLP Kindergarten	0.012	0.123	
MCLP First	0.118	0.143	
MCLP Second	0.288	0.138	
Overall Impact (Weighted Pooled Results)	0.122	0.061	<.05

Table 4: MCLP 2ND Cohort Interim MPG Results

	Standardized Effect	Robust Standard Errors	p-value
MCLP Kindergarten	0.117	0.0968	0.227
MCLP First	0.177	0.0641	0.006
MCLP Second	-0.047	0.1114	0.664
Overall Impact (Weighted Pooled Results)	0.12	0.0482	<.01

Table 5. MCLP 2ND Cohort Interim PALS Results

	Standardized Effect	Robust Standard Errors	p-value
MCLP Kindergarten	0.438	0.0895	<0.0001
MCLP First	0.372	0.0775	<0.0001
Overall Impact (Weighted Pooled Results)	0.400	0.0443	<.0001

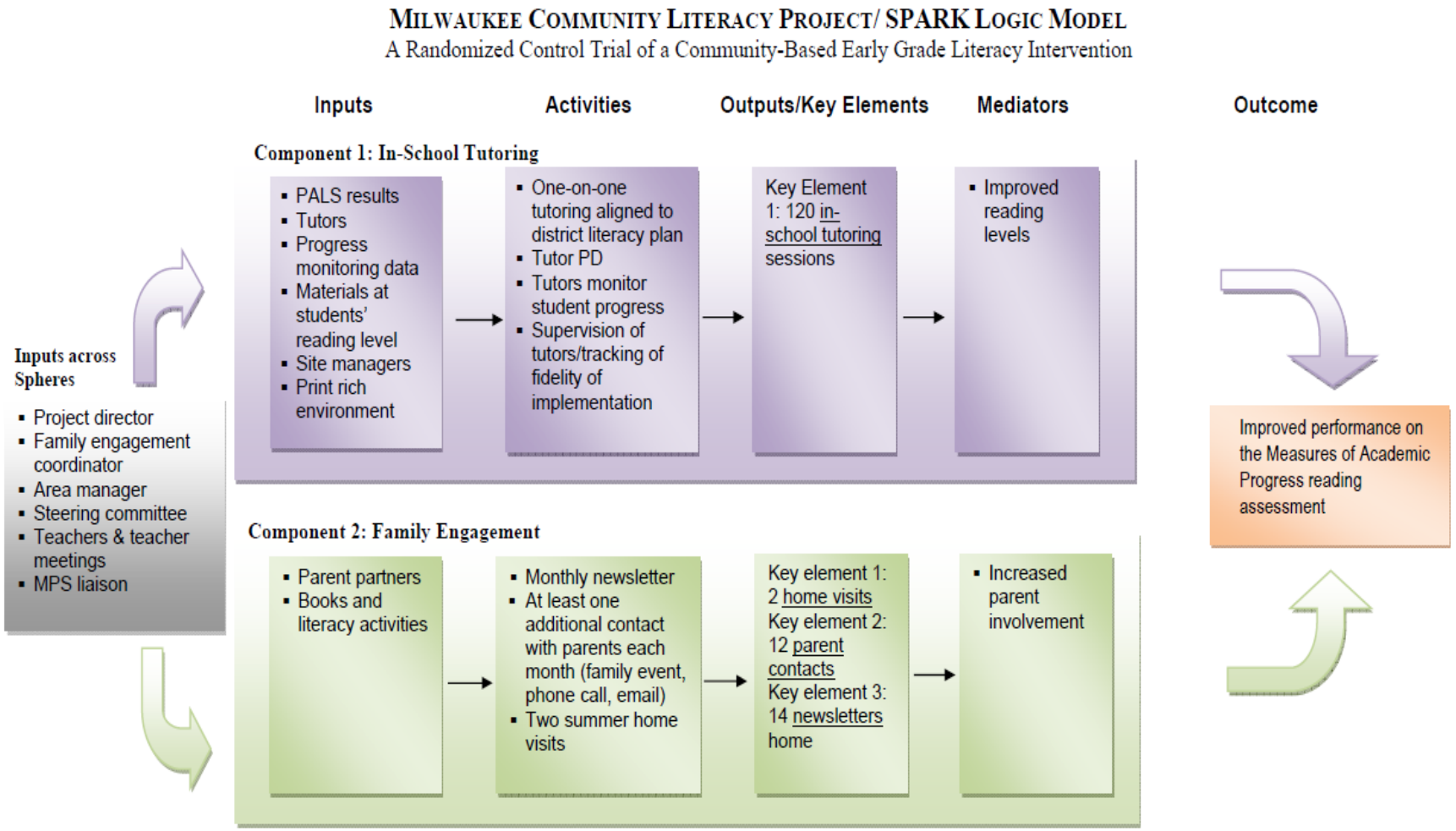

Figure 1: MCLP/SPARK logic model

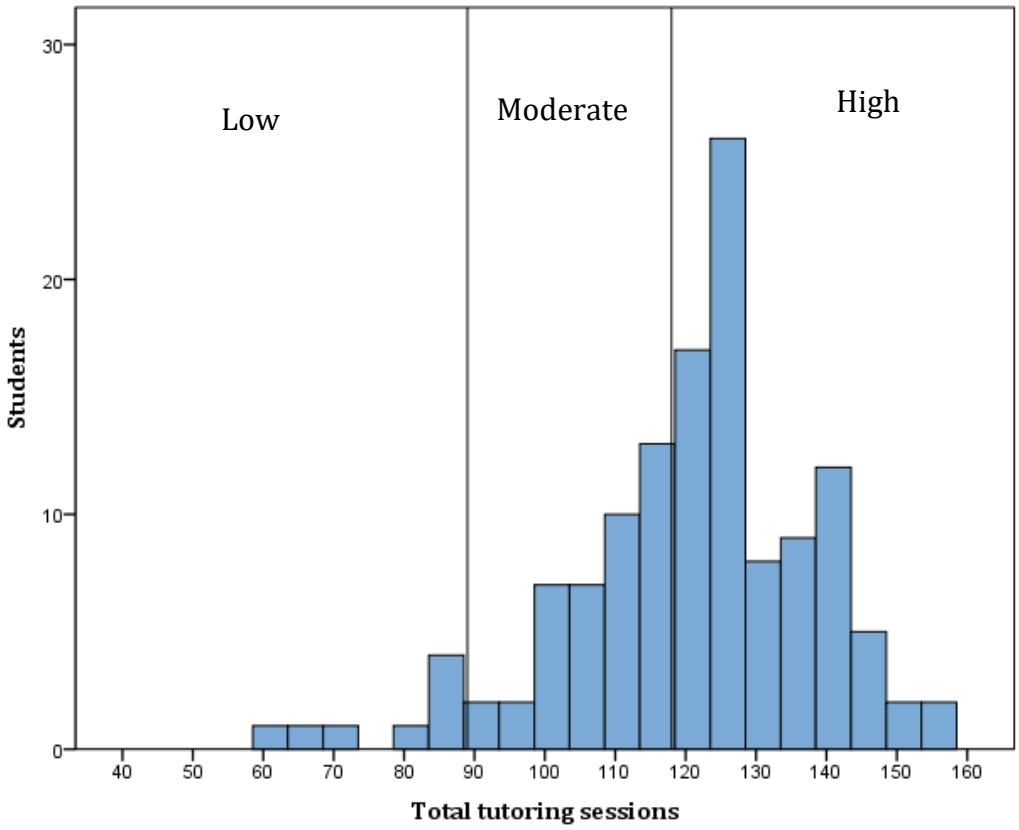

Figure 2: Distribution of total tutoring sessions received for each SPARK participant in the first cohort

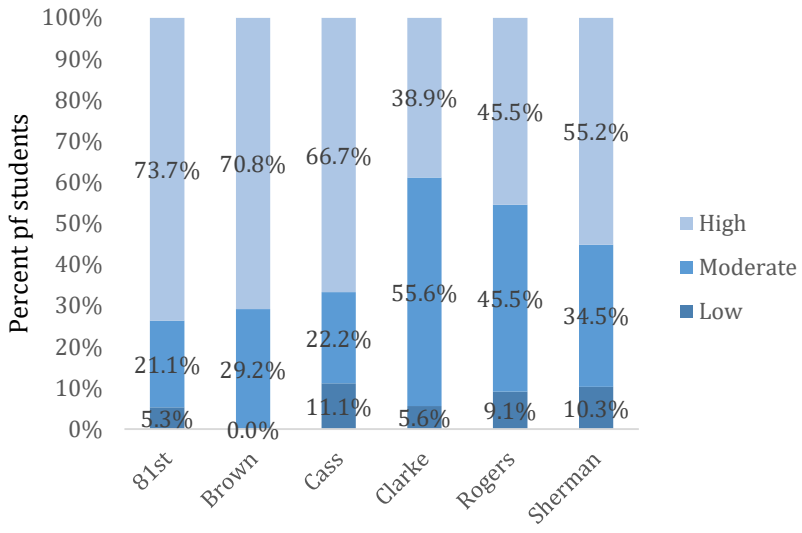

Figure 3: Tutoring implementation levels by site in the first cohort