

What is the Sustained Impact of Future Forward on Reading Achievement, Attendance, and
Special Education Placement Five Years after Participation?

Curtis J. Jones, Marlo Reeves, Dongmei Li, & Leon Gilman

Socially Responsible Evaluation in Education

University of Wisconsin Milwaukee

Curtis Jones, Ph.D. is the Research Program Director at the Office of Socially Responsible Evaluation in Education at the University of Wisconsin Milwaukee / 17 West Main Street Madison, Wisconsin 53703 / jones554@uwm.edu / His research focuses on the ecology of education and the ethics of research.

Marlo Reeves, Ph.D. is a former Senior Research Associate at the Office of Socially Responsible Evaluation in Education at the University of Wisconsin Milwaukee and current Researcher in Youth, Family, and Community Development at the American Institutes for Research (AIR) / 10 S. Riverside Plaza, Suite 600, Chicago, IL, 60606 / mreeves@air.org. Her research uses sociological frameworks to study community-based youth organizations as sites of resistance and submission amidst growing social inequities.

Dongmei Li, Ph.D. was a senior research associate at the Office of Socially Responsible Evaluation in Education at the University of Wisconsin Milwaukee / 17 West Main Street Madison, Wisconsin 53703 / dongmeili0414@gmail.com / Her research interests include K-16 education policy studies associated with access and equity, accountability, educational reform, and international education.

Lean Gilman, M.S. is a former Research Associate at the Office of Socially Responsible Evaluation in Education at the University of Wisconsin Milwaukee and current Data Analytics and Research Specialist at Medical College of Wisconsin / 8701 Watertown Plank Rd. Milwaukee, WI 53226 / ljgilman@me.com

Thanks to...

Andrew Muriuki, Marc Sanders, and Melanie Stewart of the Milwaukee Public Schools;
Former director of Future Forward Patricia Marcus and current director Kate Bauer-Jones;
Future Forward administrative team members Kris Bischoff and Tia Hatchett;
Thomas Johnson, Cris Price, Denise Ross, Alexandria Sedar, and Erica Turner for their input in
the development of the manuscript.

Dedicated to the memory of my friend, mentor, and collaborator Edison J. Trickett.

What is the Sustained Impact of Future Forward on Reading Achievement, Attendance, and Special Education Placement Five Years after Participation?

In 2010, Boys & Girls Clubs of Greater Milwaukee (BGCGM) was awarded an Investing in Innovations (i3) grant to develop and test the impact of Future Forward, which was then called SPARK, in seven Milwaukee Public Schools (MPS). MPS, a district serving over 80,000 students, faces a significant challenge in teaching its students how to read and write. According to Wisconsin Forward Exam results, only 18.6% of MPS students were proficient in reading in 2019, compared to 40.8% statewide. The overall proficiency of MPS masks deep inequities between White, Black, and Latinx students. While 45.4% of White MPS students are proficient in reading, only 19.2% of Latinx and 10.6% of Black students are proficient.

Future Forward was developed to help address the needs of struggling readers and their families in MPS. Future Forward leverages a school-family-community partnership approach (Epstein, Sanders, Simon, Salinas, Jansorn, & Van Voorhis, 2002) to address the literacy needs of students and families. Through this partnership approach, Future Forward views literacy supports through a systems lens (Bronfenbrenner, 1979) (Figure 1). At its center, students receive 30 minutes of phonics-focused, one-on-one tutoring from a paraprofessional or volunteer three times each week for up to two years. Learning opportunities are also embedded in the community through club-based programming. Family engagement involves ongoing communications to families, from a family engagement coordinator, regarding their student's progress. The family engagement coordinator is typically from the community and a parent of a student in the school. Communications are focused on student successes and meant to be positive (Love, 1996). Families are also provided development opportunities for supporting their student's literacy outside of school. These occur during home visits and monthly family events

held at the school or a community center. Typical MPS school instruction occurs in a reading block that involves small group work, classroom instruction, and station activities. Teachers work with Future Forward staff to align instruction and supports to individual students' strengths, interests, and needs. Future Forward maintains a communication log or binder to facilitate communications with teachers about student progress. Teachers also often help plan and attend family events, leveraging the resources of Future Forward to effectively engage families and connect with students. The collaborative work between teachers, Future Forward staff, and families helps develop a learning team and builds trust between the three partners (Graham-Clay, 2005) that continues past a student's direct participation in Future Forward.

Insert Figure 1 Here

The i3 grant resulted in two randomized studies that measured the impact of Future Forward. The first study, conducted while Future Forward was still in development during the 2011-12 and 2012-13 school years, found a small, positive, but significant impact on reading achievement (Jones, 2018); students who received Future Forward support for two years had slightly greater reading achievement than students receiving only business-as-usual (BAU) reading instruction (0.12 standard deviations). The second study tested the impact of the fully developed Future Forward program in the 2013-14 and 2014-15 school years. This study found Future Forward had significant, positive impacts on literacy development (0.23 standard deviations), reading achievement (0.10 standard deviations), and school attendance (4.5 fewer days absent from school). Further, Future Forward had a differential positive impact on the literacy development of students with the least developed literacy skills (0.46 standard deviations). Although students were provided Future Forward for two years, most, if not all, of

the benefit was realized after just one year (0.34 standard deviations impact on literacy development; 0.11 standard deviation impact on reading achievement).

The current study follows the students who were assigned to receive two years of Future Forward and those assigned to receive BAU reading instruction for five years past the conclusion of the study in the spring of 2015. Data provided by MPS allows for the unique opportunity to determine if Future Forward participants' reading and attendance gains were sustained through middle school. These data also allow for an analysis of Future Forward's impact on referrals to special education, expanding the potential benefits from individual students to the entire district. The average cost per pupil in MPS is \$14,897 (Lisowski, January 31, 2019). Annual costs for a Wisconsin student in special education are approximately double that of a student in regular education, approximately \$15,000 extra per year for MPS students (Wisconsin Policy Forum, 2019). The state only covers about 25% of the additional cost for educating a student in special education. Thus, \$11,250 of the added cost is paid from MPS general funds. A MPS student who begins receiving special education services in elementary school can cost this amount for ten years (or more) as they progress through middle and high school, ultimately costing upwards of \$112,500. If Future Forward prevents students from being referred for specialized services, districts partnering with Future Forward stand to benefit fiscally.

The Dissipating Impact of Tutoring Programs

Although a great deal of research exists “proving” the effectiveness of various tutoring programs, few programs have proven effective under rigorous evaluation scrutiny, such as with a randomized control trial (RCT). Much of this research can be found summarized on websites like the What Works Clearinghouse (WWC) (<https://ies.ed.gov/ncee/wwc/>) and Evidence for ESSA (<https://www.evidencefoessa.org/>). These two websites help filter much of the noise about the

effectiveness of the hundreds of tutoring programs a Google search will uncover. A search on the Evidence for ESSA website for early elementary literacy tutoring programs with “strong” evidence for effectiveness, i.e., with at least one positive effect in an RCT, whittles the number down to only 13 programs, one of which is Future Forward.

While applying rigorous evidence standards thins the field of literacy tutoring programs, applying a standard that programs have evidence of a sustained impact after the program ends nearly eliminates their numbers. It is difficult to conduct long-term follow-up studies. Students move and fewer can be tracked the further out the follow-up study reaches. There is also less motivation to do them, i.e., for programs to define success by what students take away from their participation. Funding opportunities like Education Innovations and Research (EIR) only ask for interventions to prove their impact within the period of student participation. Further, the WWC has not developed review standards for follow-up studies. The field is not defining program success by how students benefit from their participation, if you define benefit by what students keep after the program ends. Predictably, few programs are intentionally designed to have a sustained impact, in that they do not address any reasons for why students fall behind in their reading development and they do not embed any supports in the natural environment around students, which can continue beyond their direct participation in the program. Any follow-up studies conducted by educational interventions typically represent a form of wishful thinking that although their program does not address the reasons why students are struggling, they hope they can still find evidence of a sustained impact.

Of the 13 early primary tutoring programs included in the Evidence for ESSA website, only three have publicly available studies measuring their sustained impact, Sound Partners, Reading Recovery, and Success for All. Predictably, the relatively small number of follow-up

studies attempted on these tutoring programs have shown program impacts mostly dissipate after students finish their participation (Jesson & Limbrick, 2014; Hurry & Sylva, 2007). Regarding Sound Partners, a one-on-one phonics-focused tutoring program, program impacts are at least partially sustained two years past program participation (Vadasy & Sanders, 2012; Vadasy & Sanders, 2013). Although program effects in both studies diminished, decreasing from 0.47 to 0.33 and 0.84 to 0.23, they remained statistically significant. Although Sound Partners deserves credit for publishing these studies, it is worth considering that both studies were small, with fewer than 100 students.

Success for All, which has also successfully scaled up across the United States (Quint, Zhu, Balu, Rappaport, & DeLaurentis, 2015), is a whole-school intervention developed in Baltimore with a rich history of rigorous research proving its effectiveness in a variety of contexts and with a variety of student groups (Cheung, Xie, Zhang, Neitzel, & Slavin, 2021). Given that Success for All is a whole-school intervention, expecting it has a sustained impact seems reasonable. Still, follow-up studies of its sustained effect are uncommon. A study by Borman and Hewes (2002) provides the most compelling evidence of its sustained effect. The authors report former Success for All students had higher reading achievement (equal to approximately 0.5 grade levels), were less likely to be in special education, and had faster grade progression than students in matched schools three years after leaving a Success for All school. These findings suggest its impact may be sustained for several years after participation.

Research Recovery, a one-on-one tutoring program, is the most utilized and studied reading program in the United States and perhaps internationally. With so much research conducted on Reading Recovery, even having a dedicated refereed journal, it is not surprising that follow-up studies have been attempted that measure its sustained impact. A follow-up study

of Reading Recovery in Michigan (D'Agostino, Lose, & Kelly, 2017), using propensity score matching, found the effect of first grade Reading Recovery was sustained through third grade, with a 0.16 effect size on the state reading assessment, but was gone by fourth grade. The study did report that the subset of students with the greatest need for support sustained an impact of 0.34 standard deviations on the state reading assessment. A follow-up study of its impact was also done as part of the i3 grant awarded to scale up Reading Recovery in 2010 (May, Sirinides, Gray, & Goldsworthy, 2016). The regression discontinuity study, which compared the follow-up achievement of first-grade students who were eligible to receive Reading Recovery with the achievement of students who were ineligible, included 331 schools and over 5,000 students across the United States. At the end of the program, the overall impact was 0.68 standard deviations on the Observation Survey of Early Literacy (OS), an assessment aligned to and created by Reading Recovery (Clay, 2015). One semester after participation, the impact on the OS was reduced to 0.27 standard deviations. Two years later, at the end of third grade, no effect was measured on state reading assessments. A further three-year follow-up study of students from the i3-funded project found a significant negative impact of Reading Recovery on reading achievement (May, Blakeney, Shrestha, Mazal, & Kennedy, 2022).

Considering the scant follow-up research publicly available on tutoring programs, and the amount of research about tutoring programs available more generally, it is possible that there exist additional, unpublished follow-up studies with less favorable results. The impact of publication bias (Rosenthal, 1979) may be as much about the *willingness* of programs to publish null findings as it is about their *ability* to publish findings. Programs generally influence what is written about them and probably do not want null or negative findings released. Publication bias was confirmed in a meta-analysis of studies reporting long-term effects of reading interventions

more generally (Suggate, 2014). Because of this, the mixed at best findings presented here about the long-term impacts of tutoring programs are likely better than what would be typical.

The current study aims to examine the sustained effect realized in the i3 study of Future Forward (Jones & Christian, 2021), five years after participation ended. By working to scaffold supports for literacy instruction around the larger social context of students, Future Forward is designed to improve student literacy beyond participation in the program. Future Forward does not treat literacy deficits as entirely attributable to a lack of effective literacy instruction or skill deficits. That a student in first grade is behind in their literacy development is due to a myriad of contextual factors in and outside the school. Programs that focus solely on skill development do not address why students lack foundational reading skills. After the program ends, the reasons why a student was behind in their reading remain and begin to pull the student back. Through a school-family-community partnership approach (Epstein, 2001), Future Forward seeks to scaffold supports around students' lives by embedding supports in the home and community and connecting those supports to work effectively together. Once the program ends, the various systems affecting students are better able to support a student's continued development.

Follow-up Study of Future Forward

In the fall of 2013, 576 kindergarten, first, and second grade students across seven low-income elementary schools primarily serving students of color were assigned to receive two years of Future Forward literacy or business-as-usual (BAU) reading instruction. Students assigned to Future Forward were pulled out of noncore classes for tutoring and received all the regular school day reading instruction provided to students assigned to BAU reading. Of the seven schools, six served primarily Black students. The other school served mostly Latinx students. All schools included in the study had a Boys & Girls Club program attached. Only

kindergarten, first, or second grade students without an IEP, except one related to speech-language¹, and who were not English Language Learners (ELL), were eligible to participate in the study.

Regarding implementation, Future Forward students received an intense amount of tutoring (average of 122.5 tutoring sessions ($SD = 27.3$) or 61 hours of tutoring), across two years of participation. Future Forward worked to coordinate its tutoring with teacher instruction. This was done through ongoing communications between Future Forward staff and teachers about the progress of individual students and to identify strategies for supporting student learning in Future Forward and the classroom. Communications involved everything from frequent meetings to occasional emails, depending on the preferences and availability of individual teachers. Further, Future Forward maintained a communication log between Future Forward staff and teachers. Teachers would also often visit the Future Forward room to talk with tutors and observe tutoring with specific students.

All seven schools also implemented strong family engagement supports. All seven schools organized 10 family events and sent home newsletters every month, which were also provided to teachers. All but one family attended at least one family event with the average family attending three ($SD = 2.2$). Family events were often developed in collaboration with teachers, with teachers attending and participating. Home visits occurred in the summer of 2014, between the first and second year of Future Forward. During that summer, family engagement coordinators visited the homes of 117 families at least once. Fifty-five homes were visited more than once, when the family expressed interest in receiving additional supports. All but one family also received at least one home visit during the 2014-15 school year. The family engagement

¹ At the start of the study, 51 students had a speech-language IEP.

coordinator also maintained an open line of communication with all families of students in Future Forward. Exclusive of the monthly newsletters, the typical family was engaged 32 times across the two study years ($SD = 13.9$). Communications were two-way, typically focused on updating families about their student's literacy development within Future Forward but also often involved brokering communications between teachers and families, supporting literacy resources in the home, and supporting the involvement of families in school more generally.

One hundred thirty-two students were also engaged in some amount of additional club-based, after-school programming. These typically involved additional literacy-focused activities. Generally though, students only participated a few times. There was less interest from families in students staying after school and extending the school day. The feedback Future Forward received was that the school day was already very long for young students and that students mostly wanted to go home at the end of it. In fact, prior to this study, the Boys & Girls Club did not offer programming to students until third grade. Sixty-nine students did attend a week-long Future Forward summer camp in the summer of 2015. Ultimately, what perhaps mattered the most about the community-based programming offered as part of Future Forward was that it connected students and families to their local Boys & Girls Club; when students were older, their families would be familiar with the people at the Club and the programming available to them.

As mentioned previously, the results of the i3 study were positive, with a significant impact on literacy development ($B = 0.23, p < .01$), a significant impact on reading achievement after the first year of participation ($B = 0.11, p < .05$), and a significant impact on school attendance ($B = 4.53$ fewer absences, $p < .05$) (Jones & Christian, 2021). The current study follows the educational development of study participants, as assigned in the fall of 2013. This follow-up study applies an intent-to-treat approach for testing the sustained impact of Future

Forward, including all assigned participants regardless of the amount of Future Forward they received. One-hundred ninety-eight participants were attrited before the end of the original i3 study. The great majority of attrition occurred during the summer of 2014, between the first and second program years. Attrition was due to students changing schools, moving away, or not taking outcome assessments. The hope was that some of the students who were lost to the original study would have follow-up assessment results available. Adding the 198 students who were originally attrited back into the study helps strengthen our power to detect a follow-up effect. One concern with adding the attrited students back into the study was that the impact would be watered down by students who only received part of the Future Forward program, instead of the intended two years. However, most, if not all, of the impact of Future Forward was realized during the first year of participation, which had very low attrition. No families purposefully dropped out of the program or study.

Study Questions

We followed the 286 students assigned to Future Forward and 290 to BAU for five years after the program ended in the spring of 2015, through the winter of 2020, to answer the following questions:

- *To what extent did students assigned to Future Forward continue to demonstrate greater reading achievement each of the five years after the program ended, compared to students assigned to business-as-usual reading instruction?*
- *To what extent did students assigned to Future Forward continue to demonstrate greater regular-school-day attendance each of the five years after the program ended, compared to students assigned to business-as-usual reading instruction?*

- *Were fewer students assigned to Future Forward referred for special education services, compared to students assigned to business-as-usual reading instruction?*

Methods

The Milwaukee Public Schools provided the data necessary to complete this study. Their research office provided student attendance, achievement, and demographic information for all available students from the original i3 study.

Instruments

NWEA Measures of Academic Progress (MAP). As part of the original i3 study, MPS provided Fall 2013 MAP Reading and Math scores. The MAP (Northwest Evaluation Association, 2009) is a norm-referenced, adaptive assessment of reading achievement. The technical reference manual reports an internal marginal reliability of .95. Test-retest reliabilities are reported as between .76 and .89. The MAP is also reported to have high concurrent validity with various other reading assessments, including the Iowa Test of Basic Skills and the Stanford 9 achievement test.

Phonological Awareness Literacy Screening (PALS). As part of the original i3 study, MPS provided Fall 2013 PALS scores. The PALS (Invernizzi, Swank, Juel, & Meier, 2003) is a criterion-referenced, teacher-administered literacy assessment. Although the PALS assesses foundational literacy across several areas including spelling, word recognition, fluency, and comprehension, only the overall literacy score was used in this study as a measure. The technical reference manuals report internal reliabilities of between .76 and .83, inter-rater reliabilities of .92, and test-retest reliabilities of between .92 and .96. It is also reported having both concurrent and predictive validity with various other reading assessments. MPS began administering the PALS to all kindergarten and first grade students in the 2013-14 school year. MPS began

administering the PALS to second grade students in the 2014-2015 school year. Thus, baseline (2012-13) PALS results were not available for second-grade students.

Renaissance Star Reading. MPS switched from the MAP to the Star Reading (Renaissance Learning Inc, 2019) assessment in the 2015-16 school year. MPS provided Star Reading scores covering the spring of 2016, 2017, 2018, 2019, and the winter of 2020. MPS students did not take the spring 2020 Star assessment because schools were closed due to COVID-19. Star Reading is a short, online adaptive assessment. It is reported to have high internal reliability (0.95) and concurrent validity with other reading assessments such as AIMSweb, the Iowa Test of Basic Skills, and state reading tests more generally. Star Reading scores are normed nationally and can be converted to grade levels equivalencies.

School Absences. MPS provided attendance data for the 2013-14 and 2014-15 school years during the original i3 study. Later, MPS provided attendance data covering the 2016-17 through 2019-20 school years. MPS was unable to provide attendance data covering 2015-16 because their electronic systems do not maintain attendance data that far back. Students with only partial attendance data for a school year were excluded from follow-up analysis for that year.

Attendance data were used to calculate the number of days absent students were each year from school and the total number of days absent from the 2016-17 to 2019-20 school years.

IEP Status. MPS provided the 2018, 2019, and 2020 IEP status of study participants. No other years of IEP status were provided to the authors.

Modeling Strategies

General linear models (GLM) were used to estimate the impact of 2013 assignment to Future Forward on spring 2016, spring 2017, spring 2018, spring 2019, and winter 2020 Star Reading scores. Star Reading scores were standardized to local norms, so model coefficients of

the Future Forward impact would approximate effect sizes. GLM was also used to estimate the impact of assignment to Future Forward on 2017 absences, 2018 absences, 2019 absences, 2020 absences (on a shortened school year because of COVID-19), and total absences from 2017 to 2020. Absences data were positively skewed so were square root transformed. Across all models, errors terms were clustered within assignment blocks (school x grade level) (Athey & Imbens, 2017),

Star Reading and attendance (absences) were modeled using the same linear regression equation, or a variant of it, employed in the original i3 study of Future Forward:

$$\begin{aligned}
 Y_{ij} = & \beta_0 + \beta_1(FF_{ij}) + \beta_2(PALS_{ij}) + \beta_3(PALS\ Missing_{ij}) + \beta_4(MAP\ Reading_{ij}) \\
 & + \beta_5(MAP\ Math_{ij}) + \beta_6(MAP\ Math\ Missing_{ij}) + \sum_{c=1}^c \beta_{7,c}X_{cij} + \sum_{j=1}^{j-1} \beta_{8,j}Block_j \\
 & + \varepsilon_{ij}
 \end{aligned} \tag{1}$$

where Y_i is the *outcome (Star, or absences)* for the i^{th} student in the j^{th} block; β_0 is the intercept; β_1 is the impact of Future Forward; FF_i is a binary indicator for Future Forward participation; $PALS_{ij}$ is baseline PALS score set to 0 if missing; $PALS\ Missing_{ij}$ is a binary indicator for not having a baseline PALS score; $MAP\ Reading_{ij}$ is baseline MAP Reading score. No students were missing a baseline MAP Reading score; $MAP\ Math_i$ is baseline MAP Math score set to 0 if missing; $MAP\ Math\ Missing_{ij}$ is a binary indicator for not having a baseline MAP Math score; X_{cij} = the c^{th} of c additional covariates representing demographic characteristics (gender, race, baseline IEP status, free/reduced lunch eligibility); $Block_j$ is the fixed effect of assignment group; and ε_{ij} is the error term.

The impact of assignment to Future Forward on 2020 IEP status was tested using logistic regression. The logistic regression model presented below predicting η_i (the log odds of a student i having an IEP in 2020), includes the same predictor variables as equation 1.

$$\ln \left[\frac{\phi_i}{1 - \phi_i} \right] = (\eta_i) \quad (2)$$

$$\begin{aligned} \eta_{ij} = & \beta_0 + \beta_1(FF_{ij}) + \beta_2(PALS_{ij}) + \beta_3(PALS\ Missing_{ij}) + \beta_4(MAP\ Reading_{ij}) \\ & + \beta_5(MAP\ Math_{ij}) + \beta_6(MAP\ Math\ Missing_{ij}) + \sum_{c=1}^C \beta_{7,c} X_{cij} + \sum_{j=1}^{j-1} \beta_{8,j} Block_j \end{aligned}$$

Dummy variable replacement was used for missing baseline PALS and MAP Math scores.

(Puma et al., 2009). No other baseline data were missing. With dummy variable replacement, missing data are set to 0. A dummy variable is then added to the model, set to 1 if missing and 0 if not missing. The drawback of using this method is that it biases the estimates of variables with missing data. This was not a concern for this study as we were not focused on the relationships of covariates with outcomes.

In the original study, Future Forward had a differential impact on students in the lower half of the PALS (for K and first grade) or MAP Reading (for second grade) baseline distribution (Jones & Christian, 2020). The current study again tests the sustained impact of Future Forward on students in the upper and lower halves of the baseline literacy distribution. Fall 2013 PALS scores were used to determine this for kindergarten and first grade students. Fall 2013 MAP Reading was used for second grade students and for other students missing PALS scores.

How Many Participants Remained in Each Follow-up Analysis?

Assignment occurred in the fall of 2013, with 286 kindergarten, first, and second grade students randomly assigned to receive two years of Future Forward and 290 to receive BAU reading instruction. Assignment was done within grade levels within schools. With seven schools

and three grade levels included in the study, there were 21 assignment blocks. Student grade levels at follow up (Table 1) suggest few study participants had been held back in the years since the study ended. By 2020, nearly all assigned students were in middle school grades.

Insert Table 1 here

Over the course of the five follow-up years from 2016 to 2020, overall attrition ranged from 28.1% to 45.5% (Figure 2). A student was considered attrited if their outcome data were not provided for a specific year. Because a small number of students did not participate in the Star Reading assessment each year, and because some students unenrolled and reenrolled during the follow-up period, a student could be attrited one year and not attrited in a later year. Differential attrition (Future Forward attrition – BAU attrition) ranged from -5.8% in 2016 to 1.2% in 2017. Although it is unclear to what extent the What Works Clearinghouse (WWC) review standards apply to follow-up studies, the combination of overall and differential attrition meets WWC liberal attrition standards in 2016 (28.1% overall and 5.8% differential), 2018 (39.4% overall and 4.0% differential), and 2019 (42.9% overall and 2.5% differential). It meets conservative standards in 2017 (36.1% overall and 1.2% differential) and 2020 (45.5% overall and 2.0% differential) (What Works Clearinghouse, 2020).

Additional analyses suggest students were missing at random. There was no consistent difference in baseline or outcome reading achievement between missing cases assigned to Future Forward or BAU. In some years, missing students assigned to Future Forward had slightly lower baseline achievement or follow-up achievement for years they had Star Reading results, compared to students assigned to BAU reading. In other years, missing Future Forward students had slightly higher scores. In most cases though, Future Forward and BAU students with missing data had roughly equivalent baseline and follow-up scores.

Put Figure 2 about here

Follow-up analyses included between 153 and 214 Future Forward participants and 161 and 200 BAU students, depending on the year (Table 3). A comparison of baseline MAP and PALS results for students included in each year's analysis suggests the analytic samples had similar baseline reading skills as the students excluded each year (Table 2). General Linear Modeling (GLM), controlling for block fixed effects, was used to compare baseline assessment scores between students assigned to Future Forward or BAU each year. The results suggest there was little variation in baseline reading levels between these groups.

Regarding measuring the sustained impact on reading achievement, assuming 50% of the variance explained by covariates, the study had an 80% likelihood of detecting a standardized effect of between 0.20 and 0.22. Regarding attendance (absences), assuming 20% of the variance explained by covariates, the study has an 80% likelihood of detecting a standardized effect of between 0.26 and 0.28.

Insert table 2 here

Results

What Was the Sustained Impact of Future Forward on Reading Achievement?

Table 3 presents Star Reading assessment results for each follow-up analysis, both overall and for students who started the study in the lower and upper half of the reading distribution. These unadjusted results suggest students assigned to Future Forward, overall, continued to score higher on the Star Reading assessment at follow-up, compared to students assigned to BAU reading. However, looking closer at the results suggests this difference was due to its impact on students with *more* developed reading skills at the start of the study. This result

is contrary to the findings from the original study, which found that students with the greatest need for reading support benefitted the most from their participation.

Adjusted Star Reading assessment results, using equation 1, are consistent with the unadjusted results. Overall, assignment to Future Forward continued to impact students' reading achievement three years past their involvement in the program (Table 3). The effect ranged in magnitude from 0.15 standard deviations in 2016 ($p = 0.003$) to 0.20 in 2018 ($p = 0.024$). Although the 2019 and 2020 adjusted impact estimates were similar in magnitude to the previous years' estimates, they were not statistically significant. Regarding students who started Future Forward with greater literacy instructional needs, assignment to Future Forward was not associated with improved follow-up reading assessment results. The overall sustained impact of assignment to Future Forward was driven by its differential impact on students who started the study with more developed literacy skills. The impact on this group was sustained for five years after participation ended, ranging in magnitude from 0.24 standard deviations in 2016 ($p = 0.004$) to 0.40 standard deviations in 2019 ($p = 0.011$).

Insert Table 3 here

We next examined the practical significance of the sustained impact of Future Forward by comparing the grade level equivalencies of students assigned to BAU and Future Forward. Figure 3 presents the years ahead or behind grade level students were at the end of each academic year after Future Forward ended. One year after the study ended, students assigned to Future Forward were reading right at grade level (0.03), while students assigned to BAU were reading 0.29 years below grade level. This difference of about one-third of a school year was maintained across all five follow-up years. Even considering the positive impact of Future Forward, over time, students still fell far behind grade level norms. By 2020, the typical former

Future Forward student was reading 2.03 years below grade level. Even students assigned to Future Forward who started the study with more developed literacy skills (in the upper 50%) fell behind as they moved into middle school. By 2020, the typical student with more developed literacy skills assigned to Future Forward was reading 1.17 years below grade level norms.

Insert Figure 3 here

It is worth noting though that MPS students did not participate in the Star Reading spring assessment in 2020 because of the disruption to schooling caused by COVID-19. Thus, the 2020 assessment is the winter assessment. That this assessment was administered earlier in the year may help explain the larger drop off in reading grade levels seen between 2019 and 2020. While it might be reasonable to expect spring assessment results would have shown students with higher grade level equivalencies, MPS schools were shut down in March and stayed closed for the remainder of the school year. Had there been an assessment at the end of the school year, with schools closed since early March, it seems likely the results would not have changed much from the winter assessment that was administered in January. Regardless, students assigned to Future Forward ended the follow-up study reading two years behind grade level, despite its positive sustained impact.

What Was the Sustained Impact of Future Forward on Absences?

Overall, students assigned to Future Forward continued to have fewer absences from school every year after the study ended (Table 4). Among students who had complete attendance data across all four years (2017 to 2020), students assigned to Future Forward had a total of 5.9 fewer absences from school than students assigned to BAU reading. As was the case with reading achievement, this positive difference was driven by students who started the study with

more developed literacy skills. Among these students, those assigned to Future Forward were absent from school 14.9 fewer days than students assigned to BAU reading.

Adjusted absences results, using equation 1, suggests Future Forward did not have an overall impact on school absences or a differential impact on the absences of students who started the study with less developed reading skills. Consistent with its impact on achievement, assignment to Future Forward was associated with fewer absences for students who started the study with more developed reading skills. The 3.8 fewer absences students assigned to Future Forward had in 2019 and the cumulative 14.9 fewer absences they had across all four years both reflect statistically significant differences (Table 4).

Insert Table 4 here

What Was the Impact of Future Forward on Referral for Specialized Services?

In 2020, five years after the end of the program, 13 of 145 students (9.0%) assigned to Future Forward were receiving specialized services compared to 23 of 154 (14.9%) assigned to BAU (Figure 4). Regardless of assignment condition, roughly the same percent of the 141 students who started the study with greater needs for literacy supports were receiving specialized services. However, regarding 158 students who started the study with more developed literacy skills, only one of 82 students assigned to Future Forward (1.2%) was receiving specialized services compared to 10 of 76 BAU students (13.2%).

Insert Figure 4 here

Adjusting these results with equation 2, suggests the overall impact of Future Forward on referral for specialized services was statistically significant (Table 4). Logistic regression analyses provide an impact estimate called an odds ratio. Odds ratios range from 0 to an unlimited number. Odds ratios below one indicate the group has lower odds for the outcome, in

this case referral for specialized services. Odds ratios above one indicate greater odds. Students assigned to Future Forward had 0.32 the odds ($Exp(B) = 0.320$) of referral for specialized services compared to students assigned to BAU reading. Assignment was not related to referral for specialized services with students who started the study with greater needs for reading supports. However, students assigned to Future Forward with more developed literacy skills were much less likely to be referred for specialized services, with only 0.05 the odds ($Exp(B) = 0.05$) compared to students assigned to BAU reading (Table 5).

Insert Table 5 here

Summary Findings

We followed participants in the i3-funded impact study of Future Forward (Jones & Christian, 2020) for five years after it ended in 2015, to determine if the positive impacts of Future Forward on reading achievement and school attendance measured in that study were sustained. Across all study participants, regarding reading achievement, the impact of Future Forward was sustained. At the end of the study in 2015, its impact on reading achievement was 0.10 standard deviations. In subsequent years, Future Forward demonstrated overall, sustained impacts of between 0.14 and 0.20 standard deviations. As the sample size decreased due to attrition, the 2019 and 2020 sustained impact estimates were no longer statistically significant. The impact on reading achievement is equivalent to approximately one-third of a school year's growth, as defined by national norms. However, in the context of study schools, where reading achievement typically improved by only 0.5 to 0.6 grade levels each year, the impact of Future Forward was roughly equivalent to one-half of a year of academic growth. Regarding school attendance, the results were less clear. Although overall, Future Forward students continued to have fewer school absences than BAU students, these differences were not statistically

significant. This may be attributable to the reduced sample of students included in follow-up analyses and the lack of strong covariate controls.² We also examined the impact of Future Forward on referrals to special education services. Fewer Future Forward than BAU students were receiving specialized services in 2020. Overall, Future Forward students had 0.32 the odds of referral to special education.

We also explored whether the differential impact of Future Forward on students who started the program with lower literacy compared to other students in their school was sustained beyond their participation. We did not find that these students continued to benefit from their previous participation. Surprisingly, the sustained impact observed overall was driven by a differential positive impact on students who started Future Forward with above median literacy.

While the i3 study of Future Forward (Jones & Christian, 2020) concluded that students with below average literacy benefitted more, with an effect of 0.46 standard deviations on foundational literacy skills, that was not the case for all measures. The two-year impact of Future Forward on the reading achievement (as measured by the MAP) of students who started their participation with more (0.09 standard deviations) or less (0.12 standard deviations) developed literacy skills was about the same. Further, the one-year impact of Future Forward on reading achievement was considerably greater for students starting their participation with more developed literacy skills (0.18 standard deviations) than with less developed literacy skills (0.02 standard deviations). Consistent with this, in subsequent years, Future Forward demonstrated a large, sustained impact of between 0.24 standard deviations and 0.40 standard deviations on students who started the program with above median literacy. The reading achievement for these students was approximately one-third of a school year ahead that of BAU students. It is worth

² Absences models only accounted for about 10% of the variance.

noting that, as a group, MPS students starting with more developed reading skills were still at considerable risk of failing to reach grade-level standards.

In addition to the impact on reading achievement, students who started the program with above median reading development were absent from school less often and were much less likely to be referred for specialized services than BAU students. The sustained impact of Future Forward on the absences of students who started the study with more developed reading skills suggests students assigned to Future Forward were in school for nearly three additional weeks over the four years for which we had follow-up attendance data (2017 to 2020). These findings validate the school-family-community partnership approach of Future Forward. Other programs employing similar partnership approaches have also found significant impacts on school attendance (Sheldon, 2007; Childs & Grooms, 2018). Given the clear connection between school attendance, school grades (Gottfried, 2010; Gershenson, Jacknowitz, & Brannegan, 2017; Morrissey, Hutchison, & Winsler, 2014), and achievement (Gottfried, 2010; Gottfried, 2019) it seems likely that the sustained impact on reading achievement and reduced take up rate of former Future Forward students for special education was partially due to students being in school more often.

Beyond the long-term academic benefits to students, that Future Forward reduced the need for specialized services speaks to a potential fiscal benefit for districts. At follow-up, ten fewer former Future Forward participants were receiving specialized services than students who received just BAU reading instruction. Nine of the ten students started the study with more developed reading skills, a group of students for which we found a stronger and clearer impact of Future Forward on their later referral for specialized services. Because the summative local

added cost of specialized services for a third-grade student³ is approximately \$112,500, each student diverted away from specialized services could save the district that amount. Multiply that by nine, and Future Forward could ultimately save the district \$1,012,500 in the additional local costs associated with educating these students. The cost per student of Future Forward was \$2,700. Two hundred fifty-nine students completed at least one year of Future Forward, totaling \$699,300. This represents a possible savings of \$313,200 to the district. Of course, this cost analysis does not speak to the lifetime benefit to a student who is not tracked into special education, and is therefore more likely to graduate high school on time, go to college, stay out of prison, and be gainfully employed (Wager & Blackorby, 1996)

Together, these results suggests that perhaps it is not entirely accurate to conclude the Future Forward program impacted students with less advanced literacy skills more than students with more developed skills. It may be more accurate to say that it impacted each group differently. Students with greater academic needs benefitted more while they were in Future Forward. Students with more developed literacy skills, and their families, were able to better build off what they learned.

Discussion

This study stands out among other longitudinal studies of tutoring programs. This study is unique in the length of time it tracks participants. Students ended their participation as first, second, or third graders in 2015. Five years later, in 2020, students were in middle school. Following students this long provides robust evidence regarding the long-term stability of the impact of Future Forward. The previously reviewed studies of Reading Recovery followed

³ We do not have data for when study participants were specifically referred for special education. We have their IEP status from 2018 to 2020 though. All students with an IEP in 2020 also had an IEP in 2018. Since we know that no students were in special education when the study ended in 2015, we also know that all 36 students who had an IEP in 2020 were referred for specialized services during the 2016, 2017, or 2018 school years.

students two (May et al., 2016) and three (May et al., 2022) years past program participation. The follow-up studies of Sound Partners followed students for two years (Vadasy & Sanders, 2012; Vadasy & Sanders, 2013) while the follow-up study of Success for All followed students for three to four years after participation (Borman, & Hewes, 2002). Estimating the program impact each year after participation, as this study does, also provides a unique look at how the sustained impact changes over time. The authors are not aware of another literacy program evaluation that has done this.

This study demonstrates that contextually relevant and system-focused extended learning programs can have a sustained impact and *meaningfully* benefit students. Given this, it is disappointing that so few other literacy-focused tutoring programs are intentionally designed to have a sustained impact on students. Future Forward's focus on family engagement seems to at least partially explain its sustained impacts. Tutoring programs that involve families have been shown to improve students' 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). Future Forward goes much further, honoring what the family is already doing at home that supports their student's academic skills (Nieto, 2012) and empowering them to be more effective in doing so. Future Forward also helps families overcome challenges to their school engagement, such as when there is a mismatch between schools and families in terms of language, schedules, and expectations (Lopez & Stoeling, 2010). Schools 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 their involvement over time (Graue & Hawkins, 2010). That only one student who

started the study with more developed reading skills was receiving special education services is important. We are not suggesting that Future Forward prevented students from having a disability. Instead, we suggest it likely prevented students from being *identified* for services. MPS overidentifies Black and Latinx students for special education, 22% compared to 17% of White students.⁴ We argue that few students in this group, BAU or Future Forward, truly have a disability (Annamma, Ferri, & Connor, 2018; Sullivan, 2011). By building lines of communication between families, teachers, and Boys & Girls Club staff, through its partnership approach, Future Forward may have given schools other strategies for addressing student deficits rather than referring student for special education. Schools better understood the individual, familial, and community strengths and resources available to students and how to access them.

As discussed, the current research milieu does not emphasize the importance of students truly benefitting from their participation, i.e., the program making a difference in their education. The current research and program evaluation enterprise, through public policies, research funding like EIR, and the Institute for Education Sciences WWC, is focused on elevating programs, not students or communities. There is little to no expectation that projects be designed to have a sustained impact and certainly not to measure it. In some cases, there may be an assumption that benefits last beyond the program. But any assumption that a program will have a lasting impact on students would go against a large body of research that shows this to be rare (Jesson & Limbrick, 2014; Hurry & Sylva, 2007). Considering this, schools may be better off not forced to spend Title I money on expensive interventions that do not benefit students. With little evidence that the existing array of tutoring programs truly make a difference for students, requiring underserved schools to spend money on unproven programs, is, at best, paternalistic.

⁴ <https://mps.milwaukee.k12.wi.us/en/District/About-MPS/School-Board/Office-of-Accountability-Efficiency/Public-Items-Emjay/District-Enrollment.htm>

We are telling schools they are incapable of addressing their students' needs and taking money away from them to give to someone else. At worst, it reflects a form of segonomics (Rooks, 2017), where money is siphoned away from low-income schools for the enrichment of external education agencies and researchers.

This study's findings also show the limitations of extended learning programs applied to an education system built on fundamental inequities. Even with Future Forward, the opportunity gaps between Black, Latinx, and White MPS students remain large. After a student leaves the program, their continued development depends on, what is often ineffective, reading instruction provided in their regular classroom (Shanahan & Barr, 1995). Programs like Future Forward exist as pieces of potential on an educational terrain built to create and sustain disparities between students. When programmatic successes are analyzed as a small part of this terrain rather than singular programmatic solutions, we highlight the need for systemic change beyond Future Forward's reach. Examining the microlevel impact of programs like Future Forward emphasizes the capacity for growth and implicit rejection of a deficit-based ideology often conferred on Black and Latinx students. When looking across institutions, however, we see the limits of programmatic solutions, no matter how thoughtfully designed. It is unreasonable to expect time and resource-limited programs to fundamentally transform the educational terrain responsible for creating the disparities. Such expectations unfairly relocate the liability for transformation from system-level to individual schools or programs. Referring again to Figure 1, the greatest systemic determinants of the fact that so few students of color are proficient in reading are not affected by Future Forward or any other literacy programs. Milwaukee is the most racially segregated city in the United States, with fewer economic and educational opportunities available to families of color (Levine, 2020). Schools serving students of color

typically have serious educator retention challenges (Achinstein, Ogawa, Sexton, & Frietas, 2010) that result in less effective and newer teachers, unchecked racial bias, and a weak learning community (Jones, Reeves, & Rainey, 2021; Shanahan & Barr, 1995). Far from being organized to improve student outcomes, the schools serving Black students in Milwaukee are more often organized to stagnate (Jones, Reeves, & Rainey, 2020). No one program can change these things. The level of change necessary to equalize the opportunities available to families and students of color requires a reorganization of economic and education systems.

Likely because of this limitation, after Future Forward ended, the students with the most instructional needs quickly lost whatever benefit they received from participating. Even higher achieving students, whom Future Forward had a long-term impact on, fell significantly behind in their reading development as they progressed through their education. Most programs with a sustained impact equivalent to one-third of a year of academic growth, according to national norms, and one-half of a year, according to local norms, would be considered a huge success. However, given students were still far behind in reading, it is difficult to celebrate this impact. Borman and Hewes (2002) made a similar finding in their follow-up study of Success for All. While Success for All had a sustained impact equivalent to one-half of a year of reading growth, students were still reading far below grade level. The authors suggested Success for All was not the “Great equalizer” in education, but that it could be a part of the solution if paired with an array of validated programs. Similarly, while Future Forward can certainly be part of the solution for the low academic performance of underserved students in Milwaukee and elsewhere, it is not nearly sufficient to overcome the myriad of factors that hold students back.

Future research on Future Forward

In 2017, Future Forward received an Education Innovation and Research (EIR) Mid-Phase grant from the U.S. Department of Education to test its scalability and impact in 14 schools across six school districts and three states. Part of this work involved testing the impacts of one year of Future Forward, rather than the two years employed in the i3 study, on reading achievement and school attendance. Along with everything else, the EIR-funded study of Future Forward was negatively impacted by the COVID-19 pandemic and the associated disruption to education. In spring of 2020, schooling moved online. Plans for student testing were scrapped. This also caused Future Forward to move tutoring and family engagement online during the 2020-21 school year. Even considering the disruption caused by COVID-19, the results of the EIR Mid-Phase grant still found Future Forward had significant impacts on school attendance (Jones & Li, 2021) and reading achievement (Jones & Li, 2023), with greater impacts on underserved students (Jones & Li, 2021; Jones & Li, 2022). The evaluation also found evidence that the impact of one semester of Future Forward on reading achievement was sustained for at least one semester (Jones & Li, 2023).

In 2021, Education Analytics was awarded an EIR expansion grant. The scaled-up Future Forward program will continue its focus on developing parent, family, and community partnerships that align supports and strengthen trust between partners to continue benefitting students beyond their immediate participation. Participation will be shortened from one school year to one semester to support more students and to reduce costs. The associated evaluation will leverage a RCT delayed-intervention design, which provides rigorous impact evidence while also ensuring equitable access to the program. During this study, the evaluation will work to document the program conditions that make Future Forward unique among tutoring programs. What does involvement look like for families of students in Future Forward? To what extent do

families feel agency regarding their involvement in their student's literacy development and schooling more generally? To what extent do teachers view parents as resources or barriers? The evaluation will examine how well these conditions can be realized in a scaled-up iteration of Future Forward. The evaluation will also continue to monitor student participation in specialized services. The exact mechanisms for how Future Forward was able to divert students away from specialized services remains unclear. Greater attention will be given to documenting the specific disabilities that students of varying reading abilities are or are not identified with over time. Comparing the specific disabilities of students who start their participation with more or less developed reading skills could elucidate exactly how Future Forward was able to reduce participation in specialized services for students with more developed reading skills at follow-up. By exploring these and other process questions, the expansion grant provides Future Forward and the field with a unique opportunity to document the program conditions that can result in literacy-focused tutoring programs having a meaningful and sustained impact on students.

References

- Achinstein, B., Ogawa, R. T., Sexton, D., & Freitas, C. (2010). Retaining teachers of color: A pressing problem and a potential strategy for “hard-to-staff” schools. *Review of Educational Research*, 80(1), 71–107. <https://doi.org/10.3102/0034654309355994>
- Annamma, S. A., Ferri, B. A., & Connor, D. J. (2018). Disability Critical Race Theory: Exploring the Intersectional Lineage, Emergence, and Potential Futures of DisCrit in Education. *Review of Research in Education*, 42(1), 46–71.
<https://doi.org/10.3102/0091732X18759041>
- Athey, S. & Imbens, G. W. (2016). The econometrics of randomized experiments. In the *Handbook of Economic Field Experiments*, 73–140.
- Borman, G. D. & Hewes, G. M. (2002). The long-term effects and cost-effectiveness of Success for All. *Educational Evaluation and Policy Analysis*, 24(4), 243–266.
<https://doi.org/10.3102/01623737024004243>
- Bryan, J. (2005). Fostering educational resilience and achievement in urban schools through school-family-community partnerships. *Professional School Counseling*, 8(3), 219.
- Bronfenbrenner, U. (1979). *The Ecology of Human Development: Experiments by Nature and Design*. Cambridge, MA: Harvard University Press.
- Cheung, A., Xie, C., Zhang, T., Neitzel, A., & Slavin, R. E. (2021). Success for All: A quantitative synthesis of evaluations. *Journal of Research on Educational Effectiveness*, 14 (1), 90-115.
- Childs, J. & Grooms, A. A. (2018). Improving school attendance through collaboration: A catalyst for community involvement and change. *Journal of Education for Students Placed at Risk (JESPAR)*, 23 (1-2),122-138.
<https://doi.org/10.1080/10824669.2018.1439751>

- Clay, M. M. (2005). *An Observation Survey of Early Literacy Achievement*, 2nd Edition. Portsmouth, NH: Heinemann.
- D'Agostina, J. V., Lose, M. K., & Kelly, R. H. (2017). Examining the sustained effects of Reading Recovery. *Journal of Education for Students Placed at Risk (JESPAR)*, 22, 116-127.
- Epstein, J.L. (2001). *School, Family, and Community Partnerships: Preparing Educators and Improving Schools*. Boulder, CO: Westview Press.
- Epstein, J. L., Sanders, M. G., Simon, B. S., Salinas, K. C., Jansorn, N. R., & Van Voorhis, F. L. (2002). *School, Family, and Community Partnerships: Your Handbook for Action*, 2nd ed. Thousand Oaks, CA: Corwin Press.
- 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.
- Gershenson, S., Jackowitz, A., Brannegan, A. (2017). Are student absences worth the worry in U.S. primary schools? *Education Finance and Policy*, 12 (2), 137–165.
https://doi.org/10.1162/EDFP_a_00207
- Gottfried, M. A. (2010). Evaluating the relationship between student attendance and achievement in urban elementary and middle schools: An instrumental variables approach. *American Educational Research Journal*, 47(2), 434–465.
<https://doi.org/10.3102/0002831209350494>
- Gottfried, M. A. (2019). Chronic absenteeism in the classroom context: Effects on achievement. *Urban Education*, 54(1), 3–34. <https://doi.org/10.1177/0042085915618709>

- Graham-Clay S. (2005). Communicating with parents: Strategies for teachers. *School Community Journal*, 15, 117-129.
- Harvard Family Research Project. (2009a). *Supporting Student Outcomes Through Expanded Learning Opportunities*. Cambridge, MA: Little, P. M. D.
- Hurry, J. & Sylva, K. (2007). Long-term outcomes of early reading intervention. *Journal of Research in Reading*, 30, 227-248.
- Invernizzi, M., Juel, C., Swank, L., & Meier, J. (2015). *Phonological Awareness Literacy Screening*. University of Virginia.
- Jesson, R. & Limbrick, L. (2014). Can gains from early literacy interventions be sustained? The case of Reading Recovery. *Journal of Research in Reading*, 37(1), 102-117.
- Jones, C. J. (2018). Future Forward Early Literacy: Testing the impact of a family-school-community partnership literacy intervention. *School Community Journal*, 28, 247-264.
- Jones, C. J. & Christian, M. (2021). The results of a randomized control trial evaluation of the Future Forward literacy program: An innovative approach that pairs one-on-one tutoring with family engagement. *Journal of Education for Students Placed at Risk (JESPAR)*, 26(3), 185–209. <https://doi.org/10.1080/10824669.2020.1809419>
- Jones, C. J. & Li, D. (2021). *Testing the Impact and Scalability of the Future Forward Literacy Program*. Retrieved online at: <https://uwm.edu/sreed/wp-content/uploads/sites/502/2021/03/FF-attendance-impact-evaluation-19-20.pdf>
- Jones, C. J. & Li, D. (2022). *The 2020-21 Future Forward Literacy Program: Implementation and Impact during the COVID-19 Pandemic*. Retrieved online at: <https://uwm.edu/sreed/wp-content/uploads/sites/502/2022/02/FF-2020-21-report.pdf>
- Jones, C. J. & Li, D., (2023). *2021-22 Future Forward Switched Replication Study*:

- Final Report*. Retrieved online at: <https://uwm.edu/sreed/wp-content/uploads/sites/502/2023/04/2021-22-FF-report.pdf>
- Jones, C. J., Reeves, M. A., & Rainey, K. M. (2020). *Reorganizing the Professional Culture in Wisconsin Schools to Eliminate Racial Achievement Gaps*. Wisconsin Educator Effectiveness Research Partnership: Milwaukee. Retrieved online at: <https://uwm.edu/sreed/wp-content/uploads/sites/502/2020/10/WEERP-Report-Reorganizing-the-Professional-Culture-in-Wisconsin-Schools.pdf>
- Jones, C. J., Reeves, M. A., & Rainey, K. M. (2021). *Race, Relational Trust, and Teacher Retention in Wisconsin Schools*. Paper presented at the 2021 meeting of the Association for Education Finance and Policy. Retrieved online at: <https://uwm.edu/sreed/wp-content/uploads/sites/502/2019/11/WEERP-Brief-Nov-2019-Race-Relational-Trust-and-Teacher-Retention.pdf>
- Levine, M. V. (2020). The state of Black Milwaukee in national perspective: Racial inequality in the nation's 50 largest metropolitan areas. In 65 charts and tables. *Center for Economic Development Publications*, 56. Retrieved online at: https://dc.uwm.edu/ced_pubs/56
- 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.
- Lisowski, O. (January 31, 2019). Data Release Shows Just 54% of K-12 Funding is Spent on Instruction. Retrieved online at: <https://www.maciverinstitute.com/2019/01/data-release-shows-just-54-of-k-12-funding-is-spent-on-instruction/>
- Love, F. E. (1996). Communicating with parents: What beginning teachers can do. *College Student Journal*, 30(4), 440-444.

May, H., Blakeney, A., Shrestha, P., Mazal, M., & Kennedy, N. (2022, September).

Long-Term Impacts of Reading Recovery Through Third and Fourth Grade: A Regression Discontinuity Study From 2011-12 Through 2016-17. Paper presented at the annual meeting of the Society for Research in Educational Effectiveness (SREE) in Washington, DC.

May, H., Sirinides, P., Gray, A., & Goldsworthy, H. (2016). *Reading Recovery: An Evaluation of the Four-Year i3 Scale-Up*. CPRE Research Reports. Retrieved online at:

http://repository.upenn.edu/cpre_researchreports/81

Morrissey, T. W., Hutchison, L., & Winsler, A. (2014). Family income, school attendance, and academic achievement in elementary school. *Developmental Psychology*, 50(3), 741–753. <https://doi.org/10.1037/a0033848>

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.

Northwest Evaluation Association. (2009). *Technical Manual for Measures of Academic Progress and Measures of Academic Progress for Primary Grades*. Lake Oswego, OR: Author.

Puma, M. J., Olsen, R. B., Bell, S. H., & Price, C. (2009). *What to do When Data are Missing in Group Randomized Controlled Trials* (NCEE 2009-0049). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. Retrieved online at: <https://files.eric.ed.gov/fulltext/ED511781.pdf>

- Quint, J., Zhu, P., Balu, R., Rappaport, S., & DeLaurentis, M. (2015). *Scaling up the Success for All Model of School Reform: Final Report from the investing in innovation (i3) Evaluation*. MDRC, New York.
- Renaissance Learning. (2019). *Star Assessments™ for Reading Technical Manual*. Wisconsin Rapids, WI: Author. Retrieved online at:
<https://help.renaissance.com/US/PDF/SR/SRRPTechnicalManual.pdf>
- Rooks, N. (2017). *Cutting School: The Segronomics of American Education*. The New Press: New York.
- Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological bulletin*, 86(3), 638.
- Shanahan, T. & Barr, R. (1995). Reading Recovery: An independent evaluation of the effects of an early instructional intervention for at-risk learners. *Reading Research Quarterly*, 30, 958–996. doi:10.2307/748206
- Sheldon, S. B. (2007) Improving student attendance with school, family, and community partnerships. *The Journal of Educational Research*, 100(5), 267-275. DOI: 10.3200/JOER.100.5.267-275
- Sirinides P, Gray A, May H. (2018). The impacts of Reading Recovery at scale: Results from the 4-Year i3 external evaluation. *Educational Evaluation and Policy Analysis*, 40(3), 316-335. doi:10.3102/0162373718764828
- Sullivan, A. L. (2011). Disproportionality in special education identification and placement of English Language Learners. *Exceptional Children*, 77(3), 317–334.
<https://doi.org/10.1177/001440291107700304>

- Suggate, S. (2014). A meta-analysis of the long-term effects of phonemic awareness, phonics, fluency, and reading comprehension interventions. *Journal of Learning Disabilities, 49*, 1-20.
DOI:10.1177/0022219414528540.
- Wagner, M. M., & Blackorby, J. (1996). Transition from high school to work or college: How special education students fare. *The Future of Children, 6*(1), 103–120.
<https://doi.org/10.2307/1602496>
- Vadasy, P. F., & Sanders, E. A. (2012). Two-year follow up of a kindergarten phonics intervention for English learners and native English speakers: Contextualizing treatment impacts by classroom literacy instruction. *Journal of Educational Psychology, 104*, 987–1005.
- Vadasy, P. F., & Sanders, E. A. (2013). Two-year follow-up of a code-oriented intervention for lower-skilled first-graders: The influence of language status and word reading skills on third-grade literacy outcomes. *Reading and Writing, 26*, 821–843. doi:10.1007/s11145-012-9393-4
- What Works Clearinghouse. (2020). *What Works Clearinghouse Standards Handbook, Version 4.1*. Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance. Retrieved online at: <https://ies.ed.gov/ncee/wwc/handbooks>.
- Wisconsin Policy Forum. (2019). Special education funding in Wisconsin: How it works and why it matters. *The Wisconsin Taxpayer, 87* (1). Retrieved online at: https://wispolicyforum.org/wp-content/uploads/2019/02/Taxpayer_19_01.pdf

Figure 1: Future Forward systems framework for sustained literacy development

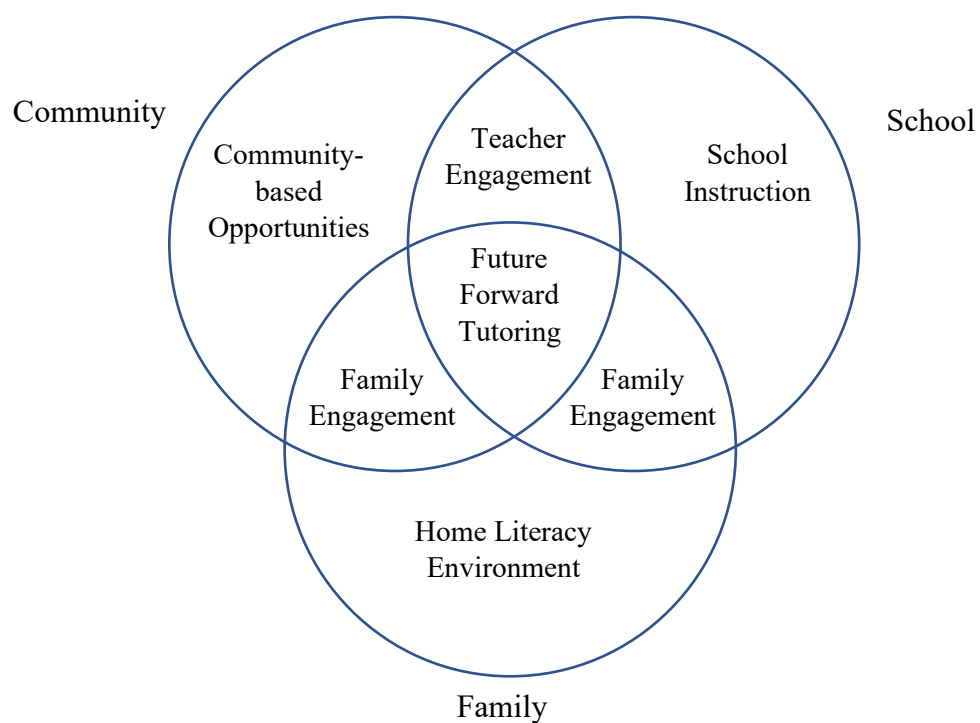


Figure 2: Attrition rates during and after the i3 study

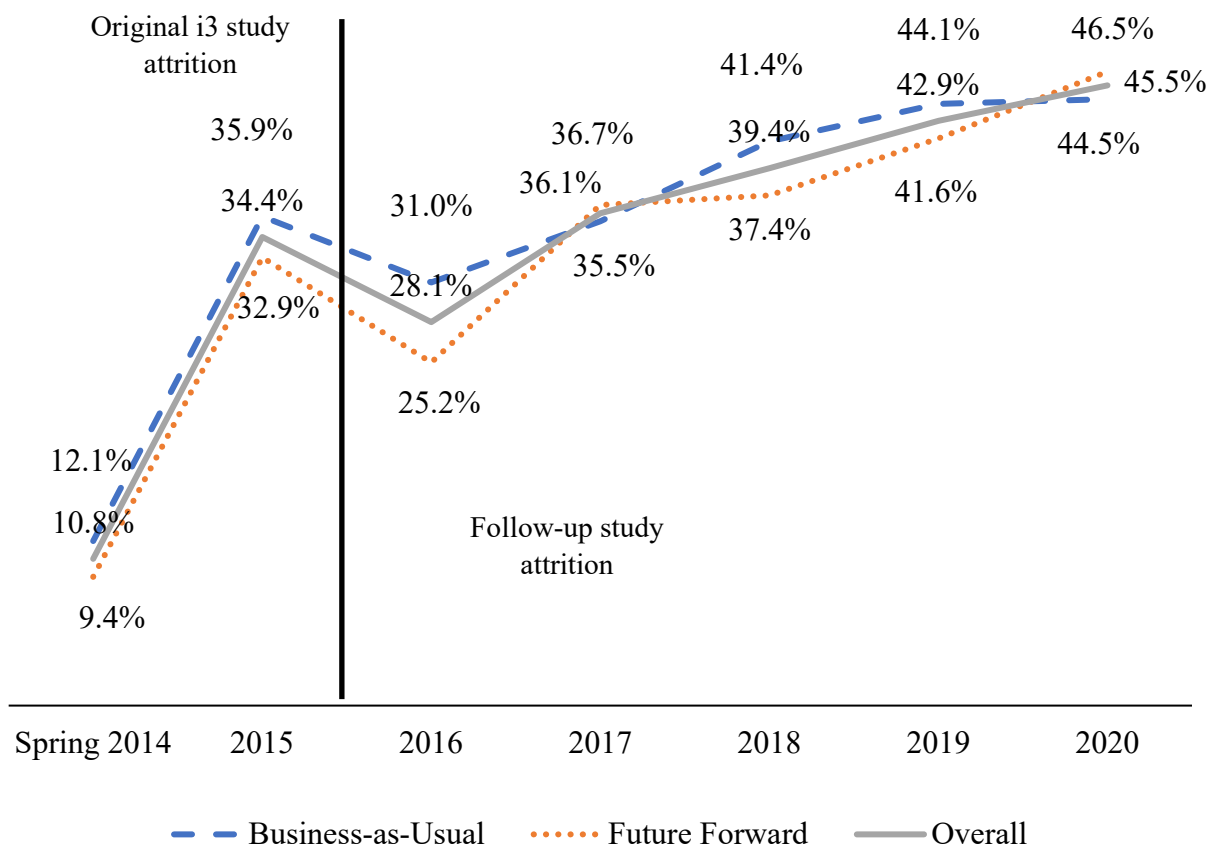


Table 1: Grade levels of students included in each follow-up analysis

	K	1	2	3	4	5	6	7	8
2014	205	214	157						
2016			146	154	114				
2017				131	133	104			
2018				5	121	129	95		
2019					4	128	110	87	
2020						6	104	117	87

Table 2: Standardized baseline achievement for sample included in each follow-up analysis

	Map Reading		Map Math		PALS	
	Leavers	Remaining	Leavers	Remaining	Leavers	Remaining
2016	0.14 (0.98)*	-0.05 (1.00)*	0.14 (1.04)	-0.04 (0.98)	-0.06 (1.03)	0.02 (1.00)
2017	0.14 (0.98)*	-0.08 (1.00)*	0.12 (1.05)*	-0.06 (0.97)*	0.03 (0.97)	-0.02 (1.02)
2018	0.04 (1.01)	-0.02 (0.99)	0.09 (1.07)	-0.05 (0.95)	-0.02 (1.04)	0.01 (0.97)
2019	0.04 (1.03)	-0.03 (0.98)	0.05 (1.06)	-0.04 (0.95)	-0.02 (1.00)	0.02 (1.00)
2020	0.08 (1.04)	-0.07 (0.96)	0.08 (1.04)	-0.06 (0.96)	-0.00 (0.99)	0.00 (1.01)

Notes:

Baseline achievement was standardized to zero with a standard deviation of one for each measure.

* Future Forward and BAU follow-up samples statistically different at baseline ($p < .05$).

Table 3 Achievement impact estimates of Future Forward for each year after participation ended

Adjusted Impact Estimates						Future Forward			BAU		
						Unadjusted Scale Scores			Unadjusted Scale Scores		
<i>Robust</i>						Standard			Standard		
		Impact	SE	<i>z</i>	<i>p</i>	<i>n</i>	Mean	Deviation	<i>n</i>	Mean	Deviation
Overall	2016	0.15*	0.05	3.00	0.003	214	317.0	164.2	200	287.7	133.2
	2017	0.16*	0.07	2.24	0.025	181	387.1	169.8	187	358.4	156.9
	2018	0.20*	0.10	2.25	0.024	179	462.4	209.8	171	418.6	170.3
	2019	0.14	0.08	1.63	0.103	167	516.8	237.8	162	481.5	214.8
	2020	0.15	0.10	1.53	0.127	153	539.5	253.5	161	508.8	224.7
Students with less developed reading skills (in the lower 50% of study participants)	2016	0.05	0.09	0.52	0.605	99	236.1	119.3	108	233.8	108.8
	2017	-0.00	0.12	-0.00	0.997	89	307.1	148.0	99	304.7	131.4
	2018	0.06	0.12	0.52	0.602	85	353.9	151.9	95	347.7	137.3
	2019	-0.05	0.10	-0.50	0.615	78	410.4	191.2	89	411.3	161.7
	2020	-0.02	0.14	-0.15	0.880	70	419.6	194.8	85	428.8	183.6
Students with more developed reading skills (in the upper 50% of study participants)	2016	0.24*	0.09	2.87	0.004	115	386.6	166.0	92	351.0	131.8
	2017	0.33*	0.12	2.84	0.004	92	464.5	153.5	88	418.9	161.9
	2018	0.31*	0.15	2.10	0.036	94	560.5	207.1	76	507.2	167.0
	2019	0.40*	0.16	2.54	0.011	89	610.0	236.4	73	567.0	240.2
	2020	0.33*	0.16	2.13	0.033	83	640.6	254.0	76	598.4	233.6

*Impact is statistically significant $p < .05$

Figure 3: The number of academic years former Future Forward and BAU students are above or below reading grade level

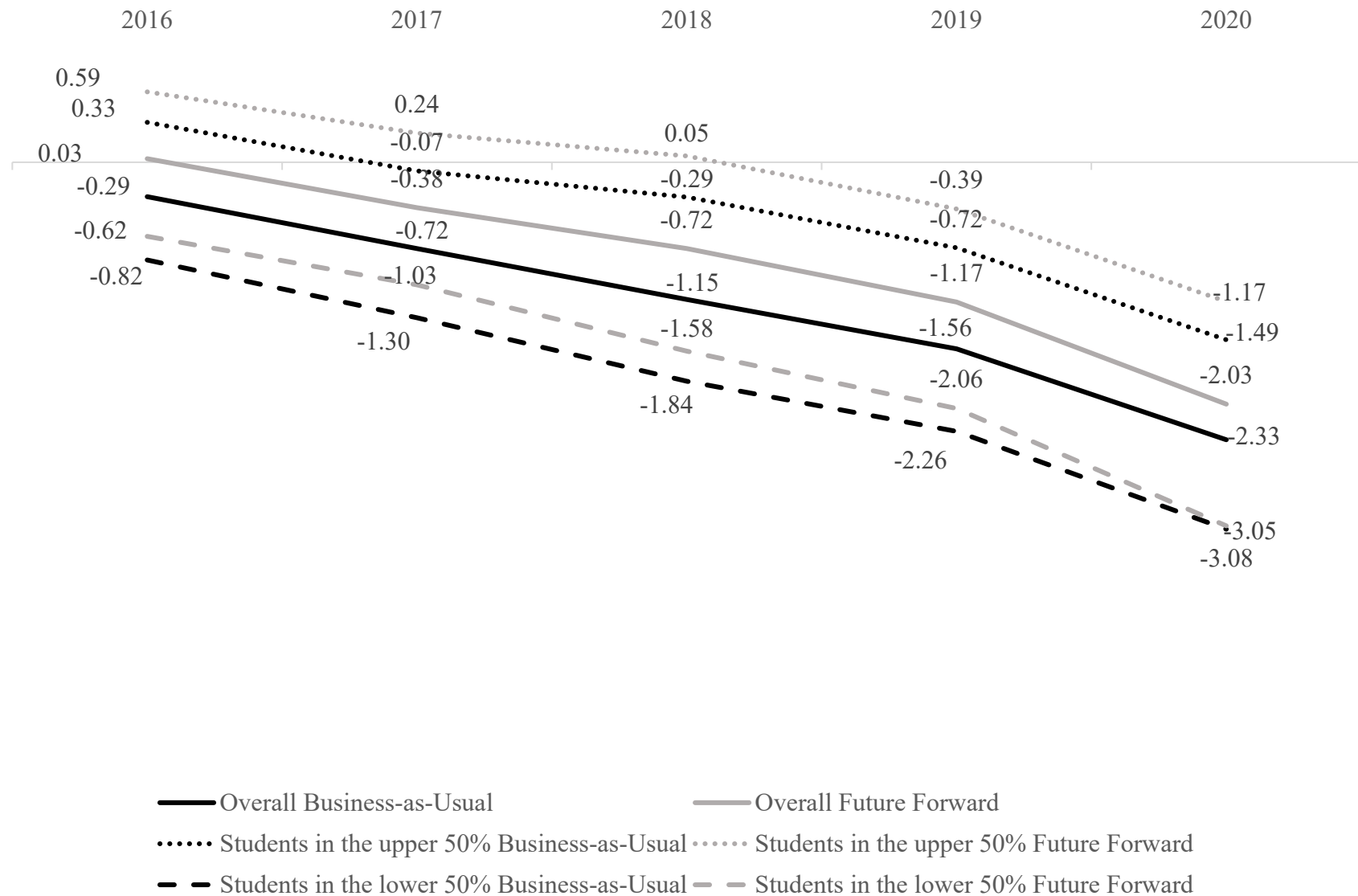


Table 4: Sqrt transformed attendance (absences) impact estimates of Future Forward for each year after participation ended

						Future Forward			BAU		
						Unadjusted Absences			Unadjusted Absences		
						Standard			Standard		
<i>Robust</i>											
		Impact	SE	t	p	n	Mean	Deviation	n	Mean	Deviation
Overall	2017	-0.16	0.20	-0.80	0.422	193	16.8	14.1	195	18.4	16.1
	2018	-0.05	0.19	-0.26	0.796	168	18.8	16.4	155	19.3	17.9
	2019	-0.28	0.19	-1.46	0.145	162	19.1	16.6	153	21.4	18.0
	2020	-0.19	0.17	-1.11	0.268	145	14.4	14.7	154	15.3	14.4
	Cumulative	-0.50	0.32	-1.56	0.119	121	67.1	49.4	124	73.0	55.8
Students with reading skills in the lower 50% of study participants	2017	-0.04	0.21	-0.21	0.833	92	20.2	13.9	100	20.8	16.7
	2018	0.24	0.23	1.07	0.287	75	23.3	16.8	78	20.7	17.8
	2019	0.08	0.32	0.26	0.795	72	23.8	18.0	79	23.7	18.8
	2020	-0.10	0.29	-0.36	0.721	63	15.5	14.7	78	16.6	14.7
	Cumulative	0.51	0.51	1.00	0.316	53	82.7	50.5	59	76.6	51.4
Students with reading skills in the upper 50% of study participants	2017	-0.30	0.30	-1.00	0.318	101	13.7	13.7	95	15.9	15.2
	2018	-0.23	0.20	-1.20	0.232	93	15.1	15.3	77	17.8	18.0
	2019	-0.56*	0.23	-2.46	0.014	90	15.3	14.3	74	19.1	17.0
	2020	0.00	0.31	0.02	0.987	82	13.6	14.8	76	13.9	14.0
	Cumulative	-1.00*	0.46	-2.17	0.030	68	54.9	45.4	65	69.8	59.7

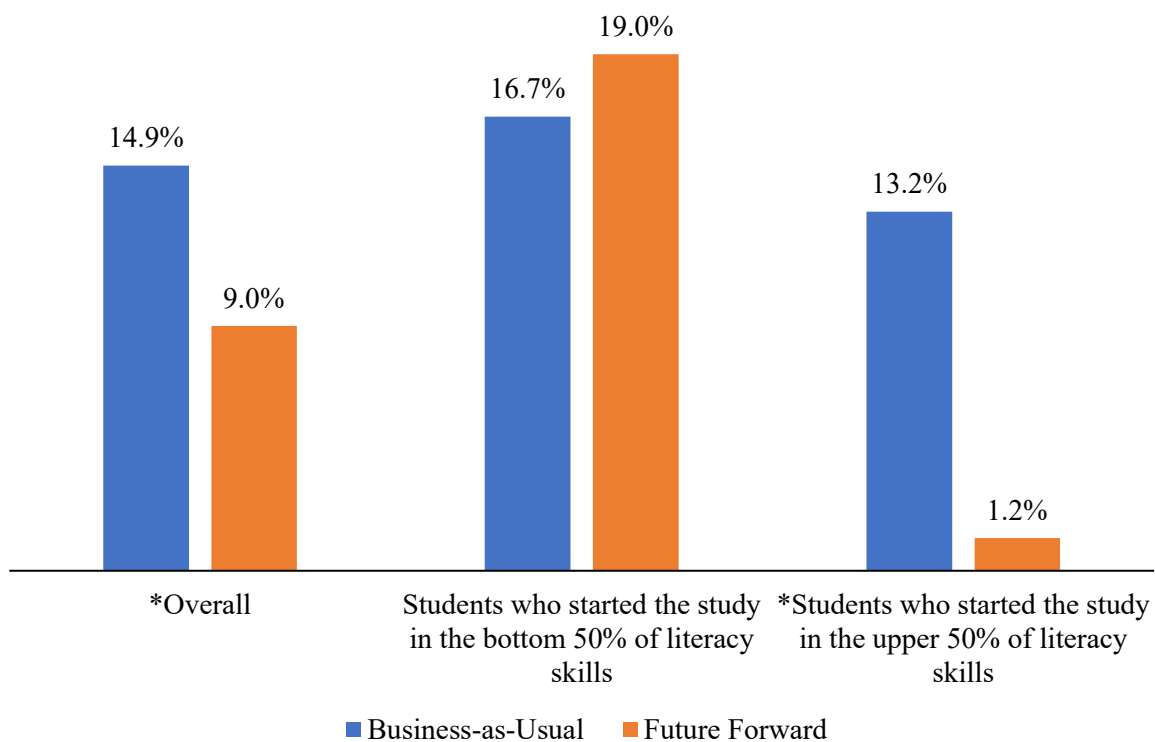
Notes:

*Impact is statistically significant $p < .05$;

2016 attendance data were not available;

Cumulative attendance data only calculated for students who had complete attendance data from 2017 to 2020.

Figure 4: The percentage of students with an IEP in 2020 - five years after Future Forward ended



* Difference indicates a significant impact ($p < .05$) of Future Forward on referral for special education

Table 5: Logistic regression impact estimates of Future Forward on receiving specialized services in 2020

	Impact (Logodds)	<i>Robust</i> <i>SE</i>	<i>z</i>	<i>p</i>	Odds Ratio	Future Forward Students with an IEP	BAU students with an IEP
Overall	-1.14*	0.49	-2.31	0.021	0.32	13	23
Students with reading skills in the lower 50% of study participants	-0.26	0.96	-0.27	0.789	0.77	12	13
Students with reading skills in the upper 50% of study participants	-2.93*	1.12	-2.62	0.009	0.05	1	10

Notes:

*Impact is statistically significant $p < .05$

Overall results estimated with 273 students in 18 blocks.

Lower 50% results estimated with 104 students in 14 blocks.

Upper 50% results estimated with 100 students in 9 blocks.