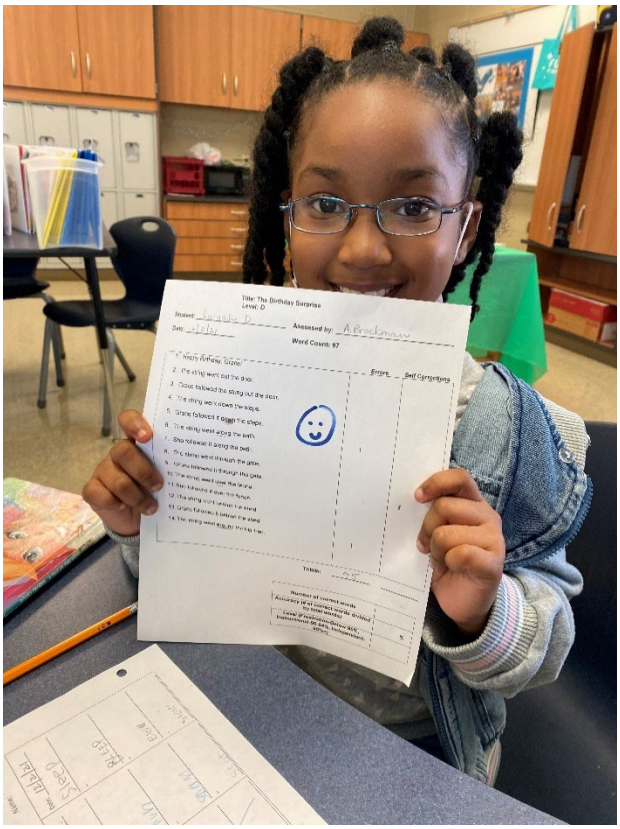
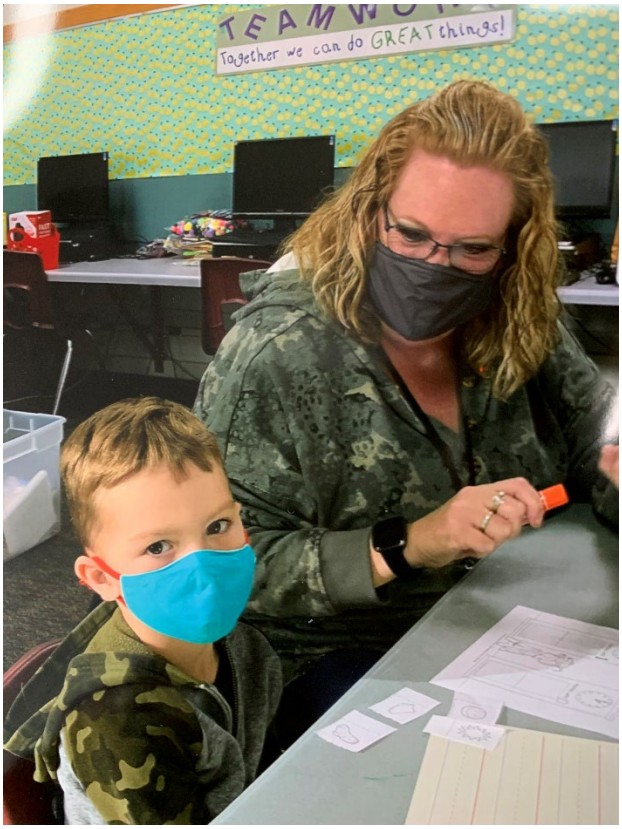




A Switched–Replication Study of the Future Forward Literacy Program: Interim Results from the 2021-22 Evaluation

Curtis J. Jones and Dongmei Li

Socially Responsible Evaluation in Education
University of Wisconsin Milwaukee

July 2022

This study was supported through a federal grant from the US Department of Education. We appreciate the support of Pamela Anderson, Kate Bauer-Jones, Kris Bischoff, Angie Brockman, and Tia Hatchett. Thanks to Cristofer Price, Jessica Cordaro, and Alexandria Sedar for their helpful feedback on draft of the report.

Curtis J. Jones is the Director of the Office of Socially Responsible Evaluation in Education at UWM. DongMei Li is a Senior Research/Evaluation Associate in the Office of Socially Responsible Evaluation in Education at UWM. For more information about this report or about the evaluation of the Future Forward, please contact Curtis Jones at jones554@uwm.edu or visit www.uwm.edu/sreed

Abstract

Future Forward is an early primary literacy program that pairs one-on-one tutoring with family engagement. As part of an Education Interventions and Research (EIR) grant, the 2021-22 evaluation of Future Forward leveraged a switched-replication (Shadish, Cook, & Campbell, 2002) randomized design to examine its implementation and impact on 127 students in two Alabama schools and one Wisconsin school. Roughly half of students were assigned to receive Future Forward in the fall semester, with the other half assigned to receive Business-as-Usual (BAU) reading instruction in the fall. BAU students will receive Future Forward in the spring. Focusing on the fall semester, the results of this interim report suggest that all but two students received the intended amount of tutoring of at least two sessions per week. Regarding family engagement, 32.3% of families were contacted at least twice per month and 64.5% at least once per month. Even considering the small sample, Future Forward students benefitted from their participation, with a statistically significant impact on reading achievement (0.29 standard deviations, p =.005). Future Forward was not found to have statistically significant impacts on school attendance or social-emotional learning (SEL). The overall impact of Future Forward on reading achievement was driven by its large impact on students in the two Alabama schools (0.47 standard deviations, p =.001). Future Forward also had a statistically significant impact on the local reading assessment in Alabama (0.31 standard deviations, p = .012). The evaluation will continue in the spring semester as BAU participants transition to Future Forward and Future Forward participants transition to BAU.

A Switched –Replication Study of the Future Forward Literacy Program:

Interim Results from the 2021-22 Evaluation

Future Forward (FF) is an early elementary literacy program, administered by Education Analytics (EA), which combines one-on-one tutoring with parent engagement to promote student literacy development, both at school and at home. In 2011, FF was funded by an Investing in Innovations (i3) grant to develop the program and test its impact in seven Milwaukee schools. Two randomized control trial (RCT) studies found the program had positive impacts on literacy, reading achievement, and school attendance (Jones, 2018; Jones & Christian 2021). Further, five-years after the i3 study ended, FF was found to have a significant sustained impact, equal to approximately one-half year of academic growth on reading achievement (Jones, Reeves, Li, & Gilman, 2021). Further, former FF participants were less likely to be receiving special education services than students assigned to Business-as-Usual (BAU) literacy instruction.

In 2017, EA received an Education Innovation and Research (EIR) Mid-Phase grant to expand FF and test its impact on students in 14 schools across three states. Although the program and its evaluation have been severely affected by COVID-19, three partial studies have been completed examining its impact on reading achievement (Jones, Bower, Pyatigorsky, & Price, 2021; Jones & Li, 2022) and school attendance (Jones & Li, 2021). These studies have shown significant impacts on school attendance and achievement, with differential positive impacts on Black students.

The Future Forward Program

FF employs a school, community, family partnership approach (Epstein, 2001) to support student literacy development of early primary grade students. The site manager is typically a certified teacher who manages personalized one-on-one tutoring provided by paraprofessionals. The site manager works with the school and tutors to develop a tutoring schedule. This involves identifying times students are pulled out of class to receive tutoring and finding tutors who can work during those times. Students are tutored by the same tutor throughout their time in FF. The site manager provides ongoing support, development, and supervision to the tutors. Each FF student is scheduled for 90 minutes of tutoring each week for one semester.

The site manager also works closely with the family engagement (FE) coordinator, who is responsible for family outreach and communication. These typically involve monthly family events and ongoing contacts through in-person meetings, phone call conversations, or email or text conversations. A fuller description of FF has been published elsewhere (Jones & Christian, 2021).

Current Study of Future Forward

The 2021-22 FF study was limited to three of the 14 schools originally included in the EIR-funded study. Interruptions to schooling during the 2020-21 school year caused by COVID-19 resulted in excess funds to continue FF in the 2021-22 school year. During 2020-21, fewer schools implemented Future Forward than was planned or budgeted for. In 2021-22, these leftover funds allowed Future Forward to be implemented in one Wisconsin and two Alabama schools (Table 1). These three schools partnered with two local Boys & Girls Clubs to implement FF.

The two Alabama schools are located in an urban community. Both schools are relatively high performing, with 51% and 57% reading proficiency, and do not receive Title I funding. They do not provide students Tier II intervention services. Both schools mostly served White students and families before a 2015 consent order of United States District Court for the Northern District of Alabama, Northeastern Division changed their attendance areas to include students living in a segregated Black community. Staff at both schools shared that they have struggled to support these new students.

The Wisconsin school is in a small town/rural community. The school is low performing, with 32% reading proficiency and receives Title I funding. It serves mostly White (73%), low-income (67%) students. Students do receive Tier II interventions as part of its Response to Intervention process.

Table 1: Study schools

	Community Type	Reading Proficiency	Percent White	Percent Low-income	Grades of Participating Students
AL school 1	Urban	51%	43%	43%	Grades KG-3
AL school 2	Urban	57%	54%	28%	Grades KG-3
WI school 1	Rural	32%	73%	67%	Grades KG-2

Partially due to the increasing need for support in these schools caused by the COVID-19 disruptions to schooling, the program length was changed from one year to one semester. The evaluation then used a switched-replication randomized design (Shadish, Cook, & Campbell, 2002), so that all families who requested support would receive FF by the end of the school year, half in the fall and half in the spring. Given the continued impact of COVID-19 on students, we felt that the switched-replication design was the most morally defensible, still allowing for a rigorous impact study but not excluding any students from receiving services. This interim impact study presents the results of FF implementation and measures the impact of the fall semester of FF on students attending these three schools.

Interim Research Questions

How was Future Forward implemented in the fall of 2021-22?

What was the impact of one semester of the FF intervention on the regular-school-day attendance, SEL, and reading achievement of K-3rd students as compared to BAU reading instruction?

Research Design

We used a switched-replication design to randomly assign students to either FF or BAU reading instruction in the fall; fall BAU students would then receive FF in the spring. This design respects participants by not subjecting anyone to the study without directly benefiting them. The design has methodological advantages as well. Since all students ultimately are served there is less chance that school staff will try to provide additional supports to BAU students, thereby contaminating the study. The design also allows tracking student outcomes as they move in and out of programming and estimating the sustained impact of FF after students switch to receiving BAU instruction.

Outcomes

School attendance was measured twice, prior to the start of the program and during the program. Attendance rates were computed by dividing the total attended days by the total days of school during that time.

School teachers completed an ***SEL assessment*** for each student at the beginning and end of the fall semester. SEL was assessed via two scales from the Devereux Students Strengths Assessment (DESSA), Optimistic Thinking (OT) and Goal Directed Behavior (GD) (LeBuffe, Shapiro, & Robitaille, 2017). The OT scale has a reported internal consistency of 0.89 when teachers complete it. The GD has an internal consistency of 0.93. Both the OT and GD scales have strong concurrent validity with teacher assessments of students using the Behavioral and Emotional Rating Scales-2 (Nickerson & Fishman, 2009).

The Star Reading assessment was administered to all students in the fall and at the end of participation in January. Star Reading is a short, online adaptive assessment with high internal reliability (0.95) and concurrent validity with other reading assessments (Renaissance Learning, 2021).

Local Reading/Literacy Assessments included the Phonological Awareness Literacy Screening (PALS) in Wisconsin and the Dynamic Indicators of Basic Early Literacy Skills (DIBELS) in Alabama. The Wisconsin schools administers the PALS in the fall and spring. The Alabama schools administer the DIBELS in the fall, winter, and spring.

PALS is a criterion-referenced, teacher-administered assessment of foundational literacy (Invernizzi et al., 2003). The assessment's internal reliabilities range from 0.76 to 0.83, inter-rater reliabilities are 0.92, and test-retest reliabilities are between 0.92 and .96 (Invernizzi et al., 2015). The assessment also has strong evidence of predictive validity for student academic performance (Invernizzi et al., 2004).

DIBELS is a teacher-administered assessment of reading skills (University of Oregon, 2018). The composite score, which was used in the current study, combines the results from five measures of reading development. Composite score test-retest reliability is high, ranging from 0.70 to 0.93 depending on the grade level and form used. It also has high concurrent and predictive validity with the Iowa Test of Basic Skills (University of Oregon, 2018-2020).

Random Assignment

One hundred twenty-seven student families consented to be in the study. Assignment was made within 36 regular-school-day classroom blocks. Sixty-five students were assigned to receive FF in the fall. The other 62 were assigned to receive BAU instruction. In the spring, the 62 BAU students received FF and the 65 FF students received BAU instruction. The size of each block ranged from 1 to 8 with 33 out of 36 classrooms having 2-5 students and an average of 3-4 students per block[1].

Study Participants

Eligible study participants included Kindergarten, first, second, or third grade students without an Individualized Education Plan (IEP) and who were not English language learners. Most of the 127 students were eligible for free or reduced lunch (88%), roughly one-third were Black, and half were White (Table 2). FF participants started the study with slightly better school attendance and SEL (Table 3). Star Reading was used to calibrate assignment and to obtain balance between assignment groups. Local reading assessment results were provided after assignment. These results indicate that students assigned to receive FF in the fall scored higher on local reading assessments than did students assigned to BAU instruction ($F = 3.89$, $p = .051$).

[1] Assuming a fixed program effect and 27% of the variance in reading explained by covariates, 41% of the variance in attendance explained by covariates, and 49% of the variance in SEL explained by covariates, the current study, prior to attrition, had an 80% likelihood of detecting an impact of 0.396 standardized units on reading achievement, an impact of 0.405 standardized units on student attendance, and an impact of 0.338 standardized units on SEL.

Table 2: Characteristics of study participants

		BAU	FF	Total
Grade Level	KG	20 (32.3%)	23 (35.4%)	43 (33.9%)
	1st	16 (25.8%)	16 (24.6%)	32 (25.2%)
	2nd	19 (30.6%)	21 (32.3%)	40 (31.5%)
	3rd	7 (11.3%)	5 (7.7%)	12 (9.4%)
School	AL school 1	20 (32.3%)	20 (30.8%)	40 (31.5%)
	AL school 2	22 (35.5%)	21 (32.3%)	43 (33.9%)
	WI school 1	20 (32.3%)	24 (36.9%)	44 (34.6%)
Race/ Ethnicity	Black	22 (35.5%)	26 (40.0%)	48 (37.8%)
	White	37 (59.7%)	30 (46.2%)	67 (52.8%)
	Other	3 (4.8%)	9 (13.8%)	12 (9.4%)
Gender	Female	32 (51.6%)	36 (55.4%)	68 (53.5%)
	Male	30 (48.4%)	29 (44.6%)	59 (46.5%)
F/R Lunch	No	8 (12.9%)	7 (10.8%)	15 (11.8%)
	Yes	54 (87.1%)	58 (89.2%)	112 (88.2%)
Total		62	65	127

Table 3: Baseline attendance, SEL, and reading achievement of study participants

		BAU	FF	Total
Attendance Rate	Mean	91.64	92.84	92.25
	SD	9.92	8.95	9.42
	n	61	64	125
SEL: DESSA				
Optimistic Thinking	Mean	-0.04	0.05	0.01
	SD	1.00	1.02	1.01
	n	59	63	122
Goal Directed Behavior	Mean	-0.03	0.09	0.03
	SD	1.01	1.00	1.01
	n	59	63	122
Local Reading Assessment	Mean	-0.17	0.17	0.00
	SD	0.83	1.07	0.97
	n	59	64	123
Star Reading	Mean	-0.01	0.06	0.02
	SD	0.83	1.08	0.98
	n	62	65	127

Attrition and Students Included in the Final Analysis

The final analytic sample included 59 BAU and 63 FF students. Only five students attrited (3.9%), three BAU (4.8%) and two FF (3.1%). All five students transferred out of their school. Four of 83 Alabama students (4.8%) and one of 44 Wisconsin students (2.3%) left the study. Differential attrition in Alabama (4.7%) and Wisconsin (4.2%) were small. The overall (3.9%) and differential attrition (1.7%) are within the conservative levels of acceptability as established by the What Works Clearinghouse (2020).

Analysis Strategy

We used general linear models (GLM) to estimate the impact of FF using the following equation:

$$(1)\quad Y_{ij} = \beta_0 + \beta_1(T_{ij}) + \beta_2(OT_{ij}) + \beta_3(GD_{ij}) + \beta_4(A_{ij}) + \beta_5(Star_{ij}) + \beta_{6a}(PALS_{ij}) + \beta_{6b}(DIBELS_{ij}) + \sum_{m=1}^{M} \beta_{7.m} X_{mij} + \sum_{j=1}^{J-1} \beta_{8.j} Block_j + \varepsilon_i$$

Y_{ij} is the *winter regular-school-day attendance, winter SEL, or winter Star Reading* for the i^{th} student in the j^{th} block; β_1 is the impact of FF; β_2 and β_3 are the effects of baseline SEL scores, β_4 is the effect of attendance in the months before the study started, β_5 is the effect of baseline Star Reading, β_6 is the effect of fall local baseline reading assessment, $\beta_{7.m}$ is the effects of student demographics, and $\beta_{8.j}$ is the effect of block. Dummy replacement method was used for missing baseline data (Puma et al, 2009).

We used fixed block effects (classroom) to control for any unobserved block-specific factors. We also conducted exploratory analyses of the impact of FF on outcomes by race and state.

Future Forward Implementation

Tutoring

All tutoring was done in person. Sixteen tutors supported FF (Table 4). All were female and all but one was White. The number of students each tutor worked with ranged from one to six. Training was provided on-site. Site managers, who are certified teachers, observed each tutor's practice twice each month and provided ongoing feedback to each tutor.

Table 4: Future Forward tutor characteristics (fall of 2021)

	Tutors	White	Other Race	Female	College Graduate	Have Teaching Experience	Students Served per Tutor
AL school 1	4	4	0	4	2	0	5
AL school 2	4	4	0	4	4	3	4-6
WI school 1	8	7	1	8	6	4	1-5

Each tutoring session was scheduled for 30 minutes. Students were scheduled to receive three tutoring sessions per week (90 total minutes). Because of a variety of interruptions to schooling such as in service days, sick days, and field trips, FF expects students to receive an average of only two sessions per week, or 60 minutes of tutoring. Tutoring spanned from October to January. The average student received a total of 25.5 sessions and 2.6 sessions per week (Table 5). All but two students received the targeted amount of tutoring (at least two sessions each week).

Table 5: Tutoring implementation

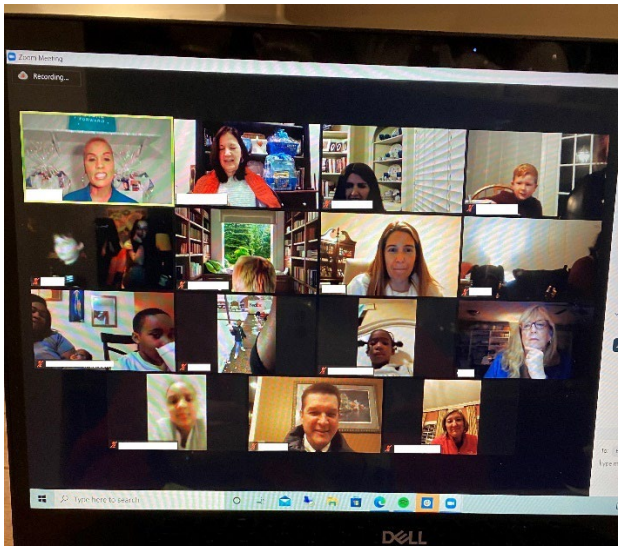
	Average Total Sessions (*SD*)	Average Sessions Per Week (*SD*)	Students Receiving 2+ Sessions Per Week (%)	FF Students
AL school 1	20.3 (5.1)	2.2 (0.6)	18 (90.0%)	20
AL school 2	22.9 (1.5)	2.5 (0.2)	20 (100%)	20
WI school 1	32.3 (3.6)	3.0 (0.3)	23 (100%)	23
Overall	25.5 (6.4)	2.6 (0.5)	61 (96.8%)	63

Family Engagement (FE)

FE involves substantive interactions or communications with participant family members. Successful contacts are defined as an exchange between FF staff and a student's family member. Family engagement efforts in Alabama were mostly done virtually, while in Wisconsin family engagement was mostly done in person. The average student's family was engaged successfully 1.5 times per month. Nineteen (30.2%) student families were engaged at least twice per month while 40 (63.5%) were engaged at least once per month (Table 6).

Table 6: FE implementation

	Average total contacts (*SD*)	Average contacts per month (*SD*)	Students' families contacted 2+ times each month (%)	Students' families contacts 1+ times each month (%)	FF students
AL school 1	5.1 (4.8)	1.7 (1.6)	7 (35.0%)	13 (65.0%)	20
AL school 2	5.5 (3.5)	1.8 (1.2)	8 (41.0%)	15 (75.0%)	20
WI school 1	3.3 (2.7)	1.1 (0.9)	4 (17.4%)	12 (52.2%)	23
Overall	4.6 (3.8)	1.5 (1.3)	19 (30.2%)	40 (63.5%)	63



Virtual family event

The impact of one semester of the FF

Descriptive results

In terms of student demographics, after attrition FF and BAU students were similar (Table 7). FF participants continued to demonstrate better baseline school attendance, SEL, and reading assessment results after attrition (Table 8). There was very little change in school attendance and SEL from fall to winter. However, regarding Star Reading, FF participants showed greater separation from BAU students in the winter assessment (0.29 standard deviations) than they did in the fall (0.08 standard deviations). Winter local reading assessment results are not presented because only Alabama schools administered a winter reading assessment.

Table 7: Characteristics of students included in the analytic sample

		BAU	FF	Total
Grade Level	KG	20 (33.9%)	21 (33.3%)	41 (33.6%)
	1st	15 (25.4%)	16 (25.4%)	31 (25.4%)
	2nd	17 (28.8%)	21 (33.3%)	38(31.1%)
	3rd	7 (11.9%)	5 (7.9%)	12 (9.8%)
School	AL school 1	19 (32.2%)	20 (31.7%)	39 (32.0%)
	AL school 2	20 (33.9%)	20 (31.7%)	40 (32.8%)
	WI school 1	20 (33.9%)	23 (36.5%)	43 (35.2%)
Race/ Ethnicity	Black	20 (33.9%)	25 (39.7%)	45 (36.9%)
	White	36 (61.0%)	30 (47.6%)	66 (54.1%)
	Other	3 (5.1%)	8 (12.7%)	11 (9.0%)
Gender	Female	31 (52.5%)	35 (55.6%)	66 (54.1%)
	Male	28 (47.5%)	28 (44.4%)	56 (45.9%)
F/R Lunch	No	8 (13.6%)	7 (11.1%)	15 (12.3%)
	Yes	51 (86.4%)	56 (88.9%)	107 (87.7%)
Total		59	63	122

Table 8: Baseline (fall) and post (winter) attendance and assessment results for students in the analytic sample.

		Fall			Winter		
		BAU	FF	Total	BAU	FF	Total
Attendance Rate	Mean	91.62	92.97	92.32	91.81	92.75	92.29
	SD	9.98	8.96	9.45	6.64	7.96	7.34
	n	59	63	122	59	63	122
OT DESSA	Mean	-0.05	0.05	0.00	-0.10	0.10	0.00
	SD	1.00	1.00	1.00	1.02	0.98	1.00
	n	59	63	122	59	63	122
GD DESSA	Mean	-0.04	0.04	0.00	-0.06	0.06	0.00
	SD	1.01	1.00	1.00	1.09	0.91	1.00
	n	59	63	122	59	63	122
Local Reading Assessment*	Mean	-0.16	0.21	0.03			
	SD	0.83	1.08	0.98			
	n	59	63	122			
Star Reading	Mean	-0.01	0.08	0.04	-0.15	0.14	0.00
	SD	0.95	1.03	0.99	0.96	1.02	1.00
	n	59	63	122	59	62	121**

* The Winter local reading assessment was only available as an outcome in the 2 Alabama schools.

**One FF student in Alabama moved to another school in Oct 2021 after the FF tutoring started and moved back in January 2022 before the FF tutoring ended. This student was not assessed with Star Reading in the winter.

Statistical modeling results

After adjusting for baseline attendance and achievement, block effects, and student demographics (equation 1), FF was not found to have statistically significant impacts on school attendance or SEL (Table 9). Regarding reading achievement, FF had a statistically significant impact on Star Reading (B = 0.289, p =.005). We also examined the impact of FF on local reading assessments (DIBELS). These were only available for the two Alabama schools. Local reading assessments were not administered in the winter in the Wisconsin school. Consistent

with Star Reading results, FF had a statistically significant impact on DIBELS (0.31 standard deviations, $t = 2.61$, $p = .012$).

Table 9: Impact estimates of FF

	B	Standard error	*t*	*p*
School attendance	0.033	1.232	0.026	0.979
DESSA – Goal Directed Behaviors	-0.034	0.147	0.233	0.816
DESSA – Optimistic Thinking	0.082	0.116	0.710	0.480
Star Reading	0.289	0.100	2.899	0.005
DIBELS*	0.305	0.117	2.610	0.012

* DIBELS is the Winter local reading assessment that was only available as an outcome in the 2 Alabama schools

FF instilling a love of reading

One of FF's goals is to instill a love of reading in students. A student's attitude toward reading should change as they gain confidence as a reader. Classroom teachers were asked to report in the fall and winter how often each student showed an interest in reading books, *frequently*, *occasionally*, or *never/rarely*. Figure 1 presents the number of students who *frequently* showed an interest in reading books. In the fall, before FF, the same number of students assigned to FF and BAU were reported as frequently showing an interest in reading books (32). In the winter, more FF students (32) were reported as frequently showing an interest in reading books than BAU students (22). Interestingly, the change in interest in reading books was entirely driven by the reduction in the number of Alabama BAU students from the fall to the winter who *frequently* showed an interest in reading (Figure 2).

Figure 1: Number of students *frequently* showing an interest in reading books.

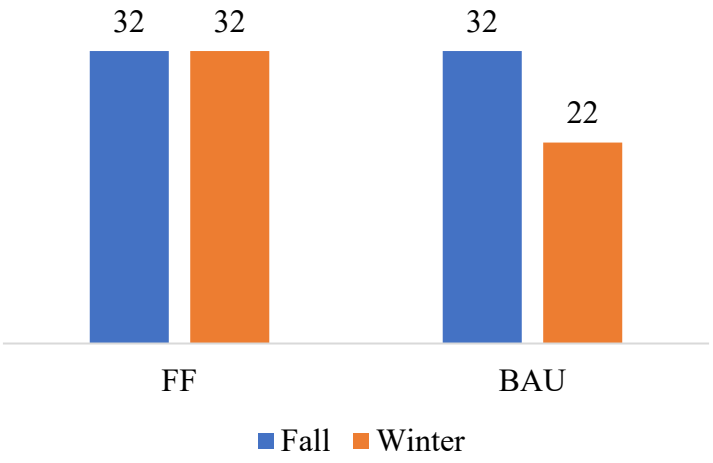
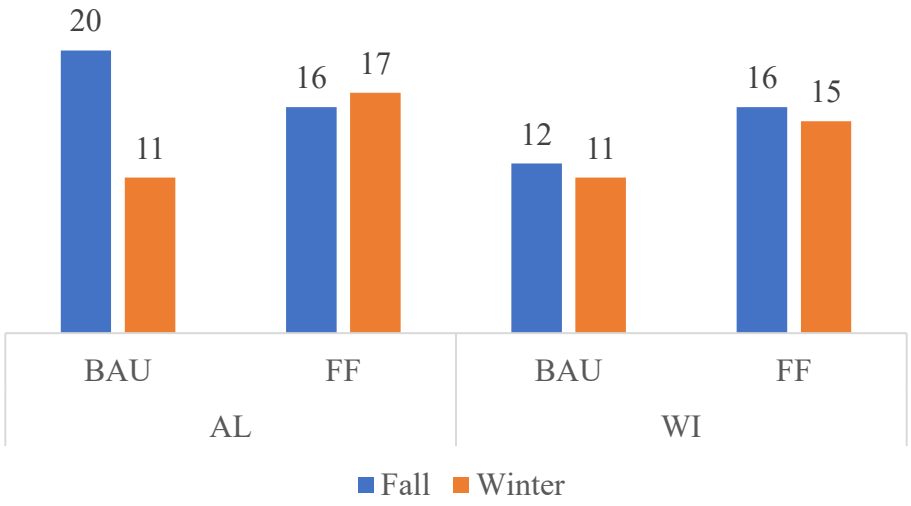


Figure 2: Number of students *frequently* showing an interest in reading books by state.



Who benefitted from Future Forward?

We explored two possibilities regarding who might have benefited from their participation in Future Forward. The first was that the overall impact was driven by differences in implementation between the Alabama and Wisconsin schools. The two Alabama schools were managed by the same staff and had similar systems, so it made sense to combine them in a comparison with students in the Wisconsin school. The second was that the impact was driven by

FF's impact on Black students. In previous studies of FF, we found evidence that Black students differentially benefit from their participation (Jones & Li, 2022; Jones & Li, 2021).

We conducted four separate statistical models (using equation 1) for Black students, White students, Wisconsin students, and Alabama students for each of the four outcomes (Table 10). The results of these models show much larger impacts of FF on Black students across all four outcomes, with a sizable impact on Star Reading ($B = 0.671$, $p = .005$). The results also show much larger impacts of FF on Alabama students, again with a sizable impact on Star Reading ($B = 0.473$, $p = .001$). Conversely, the results show negative or null effects on White students and students in Wisconsin.

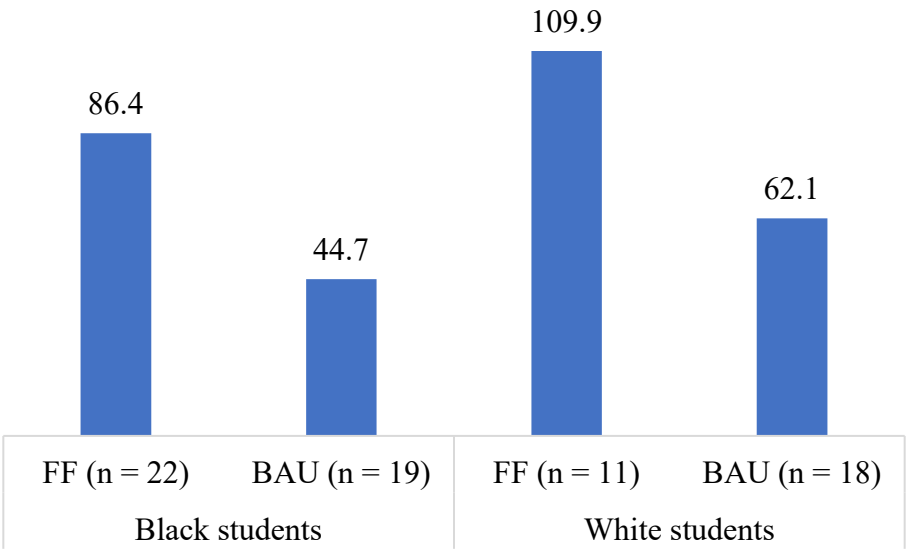
Table 10: Differential impact estimates of FF on…

	B	Standard error	*t*	*p*
Star Reading				
Black students	0.671	0.199	3.363	0.005
White students	0.067	0.170	0.394	0.696
WI students	-0.192	0.160	-1.198	0.243
AL Students	0.473	0.127	3.712	0.001
Attendance				
Black students	1.54%	2.388	0.645	.530
White students	0.85%	2.162	0.393	0.696
WI students	-1.34%	2.517	-0.530	.601
AL Students	1.59%	1.260	1.260	.214
OT DESSA				
Black students	0.050	0.250	0.199	.845
White students	-0.151	0.190	-0.799	.431
WI students	-0.107	0.160	-0.667	.511
AL Students	0.155	0.164	0.943	.351
GD DESSA				
Black students	0.249	0.308	0.807	.433
White students	-0.267	0.253	1.056	.299
WI students	-0.139	0.222	-0.625	.538
AL Students	0.061	0.202	.303	.764

Did Black students benefit from FF because they were mostly in Alabama?

Forty-two of the 45 Black students enrolled in the study attended the two Alabama schools. These 42 students accounted for 53% of all Alabama students. Because Black students were not evenly distributed across both Wisconsin and Alabama, it is difficult to disentangle the findings that FF differentially impacted Black and Alabama students. To unpack these findings, within Alabama, we compared the Star Reading growth of Black and White FF students with BAU students (Figure 3). These show that both Black and White Alabama students assigned to receive FF in the fall demonstrated nearly twice as much growth compared to students assigned to BAU in the fall. Based off these results, we conclude that the overall effect of FF and the impact on Black students were driven by the differential effect of the Alabama FF program.

Figure 3: Star Reading scale score growth from fall to winter in Alabama



Understanding the null or negative impact of FF in Wisconsin

It is concerning that Wisconsin students assigned to FF performed worse on the winter Star Reading assessment than students assigned to BAU (B = -0.192). Although this difference was not statistically significant, the magnitude of the negative effective is somewhat large. There were two circumstances that we believe help explain the negative measured effect of FF in Wisconsin. First, in Wisconsin the Winter Star assessment for FF students was completed by

early February, while for BAU students it was not completed until mid-March.[2] Thus, the Winter assessment of some BAU students accounted for over a month of additional reading development. Further, switching from BAU to FF occurred at the beginning of February. So the additional time between assessments also reflected a partial impact of FF on students assigned to BAU instruction in the fall. A second possible explanation was that Wisconsin program assignment was somewhat contaminated by local implementation of Tier II interventions. In the fall, three students assigned to BAU instruction received tier II intervention services from the school. No students assigned to FF did. Even though one student assigned to FF in the fall was identified for possible tier II intervention services, the school decided to wait until the spring, when the student was switched into BAU instruction, before providing it. Conversations with the site reveal a complication with using the switched-replication design. The site misunderstood the BAU condition, believing BAU meant students needed to receive all "usual" supports by the end of the year. It was difficult for the school to schedule FF and Tier II intervention simultaneously, so they decided to avoid placing students in both at the same time.[3] In the current study as implemented in the Wisconsin school, rather than preventing study contamination the switched-replication design contributed to it.

Summary

This switched-replication study of FF in the 2021-22 program adds to the growing body of evidence of the scalability and effectiveness of the FF program. Because of budget limitations, FF was only implemented in three schools. To improve the scalability of FF and to increase the number of students who could receive services, participation was limited to one semester. A switched-replication randomized study design was used so that all interested families could participate in FF by the end of the school year. Students were randomly assigned to receive FF in the fall or the spring.

Regarding implementation, tutoring was implemented as intended. All tutoring occurred in person. FF students received an intensive amount of supplemental reading instruction. The average student received 2.6 tutoring sessions each week. Family engagement was implemented

[2] BAU and FF Alabama students were all assessed within the same two weeks.
[3] This was not a problem in Alabama. Alabama sites did not provide Tier II interventions.

with a combination of in-person and virtual events. Most families were contacted at least once per month.

Due to our small sample size, our ability to detect program impacts was greatly diminished. FF was not found to impact school attendance or SEL. The study only had the ability to detect large program effects. That is exactly what we detected however, regarding reading achievement. FF had a large, statistically significant, positive impact on reading achievement (0.29 standard deviations). This impact was driven by a large impact on students in the two Alabama schools (0.47 standard deviations). We also found that FF had a statistically significant positive impact on local reading assessment results in Alabama (0.31 standard deviations). Considering the differential positive impact on Alabama students, we decided to include a table of baseline assessments and attendance as an appendix to this report (Table 11). These show Alabama FF students started the study with significantly higher baseline DIBELS results than BAU students (0.38 standard deviations). So, although our DIBELS analysis controlled for baseline assessments, BAU and FF students were different enough at baseline to still question the validity of the modelling results.

There is evidence that FF impacted student attitudes toward reading. While fall classroom teachers reported the same number of FF and BAU students frequently showed an interest in reading books, in the winter more FF students continued to frequently show an interest in reading books than BAU students.

The negative effect of FF measured in Wisconsin was surprising. We believe this was likely caused by a breakdown in the implementation of the evaluation and possibly reflects a challenge in implementing a switched-replication evaluation study. More up-front guidance to program and school staff and oversight may have prevented this breakdown.

The switched-replication study will continue in the spring semester, as former BAU students finish their participation in FF. The evaluation will examine if students assigned to FF in the spring catch up to the students who received FF in the fall. We will also examine the sustained impact of FF one semester after participation ended in the winter. This will involve several analyses. Mainly though, finding that the assessment results of students assigned to receive FF in the fall are roughly the same as the assessment results of students assigned to receive FF in the

spring would suggest the impact of fall semester FF participation has been sustained through the end of the year.

Appendix

Table 11: Baseline (fall) and post (winter) attendance and assessment results for Alabama students in the analytic sample.

		Fall			Winter		
		BAU	FF	Total	BAU	FF	Total
Attendance Rate	Mean	93.24	94.54	93.90	92.45	94.12	93.29
	SD	7.12	7.45	7.28	7.03	7.09	7.07
	n	39	40	79	39	40	79
OT DESSA	Mean	-0.04	0.02	-0.01	-0.27	-0.01	-0.13
	SD	0.93	0.99	0.95	1.10	1.00	1.05
	n	39	40	79	39	40	79
GD DESSA	Mean	-0.18	0.01	0.09	-0.27	0.00	-0.13
	SD	1.00	0.98	0.99	1.13	0.90	1.02
	n	39	40	79	39	40	79
Local Reading	Mean	-0.19	0.19	0.01	-0.22	0.21	0.00
Assessment-DIBELS	*SD*	0.82	1.10	1.00	0.90	1.05	1.00
	n	38	40	78	38	40	78
Star Reading	Mean	-0.15	-0.01	-0.08	-0.25	0.25	-0.00
	SD	1.01	0.91	0.96	1.03	1.09	1.08
	n	39	39	78	39	39	78

References

DIBELS. (2021). *What are DIBELS?* https://dibels.uoregon.edu/assessment/dibels#:~:text=The%20Dynamic%20Indicators%20of%20

Epstein, J.L. (2001). *School, Family, and Community Partnerships: Preparing Educators and Improving Schools*. Boulder, CO: Westview Press.

Invernizzi, M., Swank, L., Juel, C., & Meier, J. (2003). *Phonological Awareness Literacy Screening-Kindergarten.* Charlottesville, VA: University Printing.

Invernizzi, M., Justice, L., Landrum, T. J., & Booker, K. (2004). Early literacy screening in kindergarten: Widespread implementation in Virginia. *Journal of Literacy Research, 36*(4), 479-500.

Invernizzi, M., Juel, C., Swank, L., & Meier, J. (2015). *Phonological Awareness Literacy Screening.* Charlottesville, VA: University of Virginia.

Jones, C. J. (2018). SPARK 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 spark literacy program: An Innovative approach that pairs one-on-one tutoring with family engagement. *Journal of Education for Students Placed at Risk* (*JESPAR), 26*, 185-209.

Jones, C. J., Bowser, J., Pyatigorsky, M., & Price, C. (2020, September). *The Results from a Multi-site Regression Discontinuity Impact Study of the Future Forward Literacy Program*. https://uwm.edu/sreed/wp-content/uploads/sites/502/2020/10/Regression-Discontinuity-Study-of-Future-Forward.pdf

Jones, C. J., & Li, D. (2021, March). *Testing the Impact and Scalability of the Future Forward Literacy Program.* https://uwm.edu/sreed/wp-content/uploads/sites/502/2021/03/FF-attendance-impact-evaluation-19-20.pdf

Jones, C. J., Reeves, M., Li, D., & Gilman, L. (2021, April). What is the sustained impact of future forward on reading achievement, attendance, and special education placement five years after participation? https://uwm.edu/sreed/wp-content/uploads/sites/502/2021/04/The-Sustained-Impact-of-Future-Forward.pdf

Jones, C. J., & Li, D. (2022, January). *The 2020-21 Future Forward Literacy Program: Implementation and Impact during the COVID-19 Pandemic*. https://uwm.edu/sreed/wp-content/uploads/sites/502/2022/02/FF-2020-21-report.pdf

LeBuffe, P.A., Shapiro, V.B., & Robitaille, J.L. (2018). The Devereux Student Strengths Assessment (DESSA) Comprehensive System: Screening, assessing, planning, and monitoring. *Journal of Applied Developmental Psychology, 55*, 62-70.

Nickerson, A. B. & Fishman, C. (2009). Convergent and divergent validity of the Devereux Student Strengths Assessment. *School Psychology Quarterly*, 24, 48-59.

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. https://files.eric.ed.gov/fulltext/ED511781.pdf

Renaissance Learning. (2021). *Star Assessments™ for Early Literacy Technical Manual*. https://help.renaissance.com/US/PDF/SEL/SELRPTechnicalManual.pdf

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference.* Houghton, Mifflin and Company.

Shapiro, V. B., Kim, B. K. E., Robitaille, J. L., & LeBuffe, P. A. (2016, October 13). Protective factor screening for prevention practice: Sensitivity and specificity of the DESSA-Mini. *School Psychology Quarterly*. http://dx.doi.org/10.1037/spq0000181

University of Oregon. (2018). *Understanding the Research Behind DIBELS® 8th Edition (Technical* Report 1801). Eugene, OR: Author. https://dibels.uoregon.edu/sites/dibels1.uoregon.edu/files/DIBELS8thEdition_TechRpt1801_ResearchBrief.pdf

University of Oregon. (2018-2020). *8th Edition of Dynamic Indicators of Basic Early Literacy Skills (DIBELS®): Technical Manual.* Eugene, OR: University of Oregon. Available: https://dibels.uoregon.edu

University of Oregon (2020). *Understanding DIBELS® 8th Edition Composite and Measure Scores*. https://dibels.uoregon.edu/docs/UnderstandingDIBELS8CompositeScores.pdf

What Works Clearinghouse. (2020). *Standards Handbook, Version 4.1.* https://ies.ed.gov/ncee/wwc/Docs/referenceresources/WWC-Standards-Handbook-v4-1-508.pdf