





Curtis J. Jones and Dongmei Li

January 2023

2021-22 Future Forward Switched Replication Study: Final Report





For more information on this publication, visit <u>www.uwm.edu/sreed</u> Published by the Office of Socially Responsible Evaluation in Education @ the University of Wisconsin-Milwaukee © Copyright 2022



Socially Responsible Evaluation in Education

Socially Responsible Evaluation in Education (SREED) is a research and evaluation group at the University of Wisconsin-Milwaukee that partners with educators to promote system change, capacity building, and improvement for the end of improving educational equity for students from all backgrounds. SREED centers its approach on equity, viewing it as the primary objective for its work.



Acknowledgements

This study was supported through a federal grant from the US Department of Education. There are a number of people who deserve credit for this report. Within the Future Forward team, we would like to thank Pamela Anderson, Kate Bauer-Jones, Kris Bischoff, Angie Brockman, and Tia Hatchett. We would also like to thank Cris Price for his support and guidance throughout the EIR-funded grant.

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 <u>www.uwm.edu/sreed</u> For information about the Future Forward program, please contact Kate Bauer-Jones at <u>kbauerjones@edanalytics.org</u>



Executive Summary

Future Forward (FF) 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 randomly assigned to receive Future Forward in the fall semester and the other half assigned to receive Future Forward in the impact of COVID-19 related school closings on students, the program was changed from one school year to one semester, so that more students could be served.

Across both semesters, Future Forward tutoring was implemented as intended; all but three Future Forward students received the targeted amount of tutoring, at least two sessions each week. The average student received 2.7 tutoring sessions each week. Regarding family engagement, the average student's family was engaged successfully 1.6 times per month with 30.1% of student families engaged at least twice per month and 67.8% at least once per month.

Even considering a small sample size, in the fall semester, Future Forward had a large, statistically significant, positive impact on reading achievement (0.29 standard deviations, p = .005), but it was not found to have statistically significant impacts on school attendance or social-emotional learning (SEL). In the spring, Future Forward was not found to have statistically significant impacts on student outcomes. However, our analysis was complicated by the continued engagement of families of students who received Future Forward in the fall semester. The most important finding is that the end of year outcomes of students who received Future Forward in the fall and spring semesters were the same. This suggests the impact of Future Forward on reading achievement measured in the fall semester was likely sustained through the end of the year.



Contents

2021-22 Future Forward Switched Replication Study: Final Report	1
The Future Forward Program	1
Current Study of Future Forward	2
Research Questions	3
Research Design	4
Outcomes	4
Random Assignment	5
Study Participants	7
Attrition and Students Included in the Final Analysis	8
Analysis Strategy	9
Future Forward Implementation	12
Tutoring	12
Future Forward Impact	15
Overall Results	15
Descriptive results	15
Modeling results	19
FF promoting an interest in reading books	21
Differential Impact Results	22
Winter differential impact results	22
Spring differential impact results	22
Full-year differential impact results	23
FF promoting an interest in reading books in different student groups	27
Did Black students benefit from FF because they were mostly in Alabama?	29
Winter Star Reading score growth	29
Spring Star Reading score growth	30
Full-year Star Reading growth	30
FF promoting an interest in reading books in Alabama students	31
Understanding the null impact of fall FF in Wisconsin	32
Summary	33
Appendix	35
References	36



Table of Tables

Table 1: Study schools 3
Table 2: Characteristics of study participants 7
Table 3: Baseline attendance, SEL, and reading achievement of study participants 8
Table 4: Future Forward tutor characteristics 12
Table 5: Tutoring implementation
Table 6: FE implementation
Table 7: Characteristics of students included in the winter and spring analytic sample
Table 8: Baseline (fall) and winter attendance and assessment results for students in the winter analytic
sample
Table 9: Fall, winter, and spring attendance and assessment descriptive results for students in the spring or
full-year analytic sample
Table 10: Winter, spring, and full year assessment and attendance impacts on students receiving FF in the
fall compared to students receiving it in the spring
Table 11: Winter differential impact estimates of FF group for
Table 12: Spring differential impact estimates of FF group for
Table 13: Full-year differential impact estimates of FF group for
Table 14: Baseline (fall) and post (winter) attendance and assessment results for Alabama students in the
analytic sample



Table of Figures

Figure 1: Switched-replication assignment conditions	6
Figure 2: Pattern of results that suggests FF had a sustained impact on students after the fall semester	11
Figure 3: Number of students <i>frequently</i> showing an interest in reading books	21
Figure 4: Number of students <i>frequently</i> showing an interest in reading books by state	27
Figure 5: Number of Black and White students <i>frequently</i> showing an interest in reading books	28
Figure 6: Fall to winter Star Reading scale score growth in Alabama by race	29
Figure 7: Winter to spring Star Reading scale score growth in Alabama by race	30
Figure 8: The full-year Star Reading growth in Alabama by race	31
Figure 9: Number of Black and White Alabama students <i>frequently</i> showing an interest in reading	32



2021-22 Future Forward Switched Replication Study: Final Report

Future Forward (FF) is an early elementary literacy program that 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, FF received an Education Innovation and Research (EIR) Mid-Phase grant to expand 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, Price, and Pyatigorsky, 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, typically a certified teacher, manages personalized one-on-one tutoring provided by paraprofessionals. The site manager works with teachers 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 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 inperson meetings, phone call conversations, home visits, or email or text conversations.

FF also impacts social-emotional learning (SEL) directly through the strong relationship developed between the tutor and the students. Student school attendance is impacted indirectly through its engagement and support of families. A fuller description of FF has been published elsewhere (Jones & Christian, 2021). Historically, students have received FF for the entire school year. However, partially due to the increasing need for support in schools caused by the COVID-19 disruptions to school, the program length was changed from one year to one semester during the 2021-22 program year.

Current Study of Future Forward

The 2021-22 FF program was limited to three of the 14 schools originally included in the EIRfunded 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 FF than was planned or budgeted for. In 2021-22, these leftover funds allowed FF to be implemented in one Wisconsin and two Alabama schools (Table 1). EA partnered with two local Boys & Girls Clubs to implement FF in these three schools.

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. Because of this. they do not provide students Tier II intervention services. Both schools mostly served White students and families before a 2015 consent order 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 and believe that FF provides a critical resource for them.

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



	Community	Reading	Percent	Percent	Grades of
	Type	Proficiency	White	Low-income	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

Table 1: Study schools

The evaluation 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. Further, our use of a switched-replication study allows us to examine the sustained impact of FF one semester after students received it in the fall semester. Examining the sustained impact of FF defines program success by what benefits students keep after the program ends. Typically, literacy intervention impacts dissipate quickly after program end (D'Agostino, Lose, & Kelly, 2017; Jesson & Limbrick, 2014; Hurry & Sylva, 2007), suggesting that students are not any better off for having participated in the program.

Our study reports on the implementation and impact of FF in the 2021-22 school year. Impacts include reading achievement, school attendance, and social-emotional development (SEL). We report on the impact of FF in the fall and spring semesters separately. We also compare the full year's academic and SEL growth of students receiving FF in the fall to students receiving it in the spring. The full year comparison's purpose is to examine the sustained impact of FF one semester after participation ended.

Research Questions

How was Future Forward implemented in 2021-22 school year?

What was the impact of one semester of the Future Forward on the regular-school-day attendance, social-emotional learning, and reading achievement of K-3rd students as compared to BAU reading instruction?



What was the sustained impact of Future Forward one semester after participation ended?

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. Evidence of a sustained impact would require two conditions to be true. First, a significant impact is observed on students who received FF in the fall semester, Second, end-of-year outcomes of students who received FF in the fall and spring semester are equivalence, i.e. students who receive it in the spring semester are not outperforming students who received it in the fall semester.

Outcomes

School attendance was measured covering three time periods, in the weeks prior to the start of the Fall FF program, during the fall FF program, and during the spring FF 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 of the year, at the end of the fall FF program in January, and at the end of the spring FF program in May. 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). An additional question was added to the SEL assessment



asking about students' interest in reading. Student interest in reading is viewed as a mediator for the impact of FF on reading achievement.

The Star Reading assessment was administered to all students at the beginning of the year, at the end of the fall FF program in January, and at the end of the spring FF program in May. 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, interrater 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

Eligible study participants included Kindergarten, first, second, or third grade students without an individualized education plan (IEP) and who were not English language learners. One hundred twenty-seven student families consented to be in the study. Random 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 fall (Figure 1). In the spring, the BAU students receive FF and the FF students receive 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 block¹. Star Reading was used to calibrate assignment. Rerandomization was done until the two groups were balanced.

Figure 1: Switched-replication assignment conditions



¹ Our power analyses leveraged data from previous years to estimate the amount of variance accounted for by level 1 covariates and block. Assuming a fixed program effect and 65% of the variance in reading explained by covariates, 25% of the variance in attendance explained by covariates, and 70% of the variance in SEL explained by covariates, the current study, prior to attrition, had an 80% likelihood of detecting an impact of 0.313 standardized units on reading achievement, an impact of 0.458 standardized units on student attendance, and an impact of 0.289 standardized units on SEL.



Study Participants

Table 2 includes the characteristics of study participants. Most of the 127 participants were eligible for free or reduced lunch (88%), while more than one-third were Black (38%) and half were White (53%).

Table 2: Characteristics of study participants

		Spring FF	Fall FF	Total
		Group	Group	
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 represents baseline attendance, SEL, and reading achievement. Students who received FF in the fall or spring semester started the study with similar school attendance, SEL, and Star Reading. However, students assigned to FF in the fall semester had higher baseline local reading assessment results than students assigned to receive FF in the spring. Local assessment results were not available to the evaluation team at the time of random assignment.



		Spring FF	Fall FF	Total
		Group	Group	
Attendance Rate	Mean	91.64	92.84	92.25
	SD	9.92	8.95	9.42
	n	61	64	125
	SEL: DI	ESSA	i	i
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

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

Attrition and Students Included in the Final Analysis

The final analytic sample for winter data analyses included between 121 and122 students, depending on outcome. Only five to six students attrited (3.9%to 4.7%), three spring FF (4.8%) and two to three fall FF (3.1% to 4.6%). Five students transferred out of their school. One student left and then later returned to their school at the end of the semester. Four to five of 83 Alabama students (4.8% to 6.0%) and one of 44 Wisconsin students (2.3%) did not have winter assessment results. Differential attrition in Alabama (4.7% to 2.2%) and Wisconsin (4.2%) was small. The overall and differential attrition are within the conservative levels of acceptability as established by the What Works Clearinghouse (2020).



After the switching groups in the spring, one FF student was referred for special education services and removed from the study. No other students were attrited in the spring. The final end-of-year sample included 58 spring FF and 63 fall FF students.

Analysis Strategy

We used general linear models (GLM) with fixed block effects to estimate the impact of FF on winter and spring outcomes. Across all models, 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.

For winter outcomes, we use equation 1:

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

 Y_{ij} is the *fall 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 the FF group (fall or spring); 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 baseline local reading assessment; $\beta_{7.m}$ is the effects of *m* student demographics; and $\beta_{8.j}$ is the effect of block.

For spring outcomes, we use equation 2:

(2)
$$Y_{ij} = \beta_0 + \beta_1 (FFgroup_{ij}) + \beta_2 (OT_{ij1}) + \beta_3 (OT_{ij2}) + \beta_4 (GD_{ij1}) + \beta_5 (GD_{ij2}) + \beta_6 (A_{ij1}) + \beta_7 (A_{ij2}) + \beta_8 (Star_{ij1}) + \beta_9 (Star_{ij2}) + \beta_{10a} (PALS_{ij}) + \beta_{10b} (DIBELS_{ij}) + \sum_{m=1}^{M} \beta_{11.m} X_{mij} + \sum_{j=1}^{J-1} \beta_{12.j} Block_j + \varepsilon_i$$

 Y_{ij} is the *spring regular-school-day attendance, spring SEL, or spring Star Reading* for the *i*th student in the *j*th block; β_1 is the impact of FF group (fall or spring); β_2 and β_3 are the effects of baseline SEL scores; β_4 and β_5 are the effects of winter SEL scores; β_6 is the effect of



attendance in the months before the study started; β_7 is the effect of attendance in the fall semester after the program started; β_8 is the effect of baseline Star Reading; β_9 is the effect of winter Star Reading; β_{10} is the effect of baseline local reading assessment, $\beta_{11.m}$ is the effects of *m* student demographics, and $\beta_{12.j}$ is the effect of block.

For the sustained impact analysis, we use equation 3:

(3)
$$Y_{ij} = \beta_0 + \beta_1 (FFgroup_{ij}) + \beta_2 (OT_{ij}) + \beta_3 (GD_{ij}) + \beta_4 (A_{ij}) + \beta_5 (Star_{ij}) + \beta_{6a} (PALS_{ij}) + \beta_{6b} (DIBELS_{ij}) + \sum_{m=1}^{M} \beta_{7,m} X_{mij} + \sum_{j=1}^{J-1} \beta_{8,j} Block_j + \varepsilon_i$$

 Y_{ij} is the spring regular-school-day attendance, spring SEL, or spring Star Reading for the *i*th student in the *j*th block; β_1 is the impact of FF group (fall or spring); 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 baseline local reading assessment, $\beta_{7.m}$ is the effects of *m* student demographics, and $\beta_{8.j}$ is the effect of block.

Evidence for a sustained impact requires two conditions. First, the impact of FF group (β_1) measured in equation 1, would need to be statistically significant. Second, impact of FF group (β_1) measured in equation 3, would need to NOT be statistically significant. Further, with the small sample, even if not statistically significant, it will be important that the differences in outcomes between FF groups be small; students receiving FF in the fall semester would need to have comparable outcomes at the end of the year to students who received FF in the spring.

Figure 2 below presents the pattern of results that would suggest FF had a sustained impact on students. There are three outcomes that would suggest a sustained impact. One, both groups have equivalent reading scores at the beginning of the study. Two, the winter results of students assigned to receive FF in the fall semester are higher than students assigned to receive FF in the spring semester. Three, that both groups have equivalent reading scores at the end of the year.



Figure 2: Pattern of results that suggests FF had a sustained impact on students after the fall

semester





Future Forward Implementation

Tutoring

All tutoring was scheduled in person during the 2021-22 school year. The only exception is that one student who was quarantined due to COVID-19 was set up for Zoom lessons for ten days. There was a team of sixteen tutors supporting FF students (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

	Tutors	White	Other	Female	College	Have Teaching	Students
			Race		Graduate	Experience	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

Tutoring sessions were scheduled for 30 minutes three times per week (90 possible 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. FF tutoring was successfully implemented in the 2021-22 school year (Table 5). Throughout the year, all but three students received the targeted amount of tutoring.

FE also 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. FF was more effective engaging more families in the spring than in the fall (63.5% versus 72.4% engaged at least once per month) (Table 6). However, the typical family was engaged about the same number of times each month in the fall and the spring (~1.6 times).



Table 5: Tutoring implementation

Fall FF (Nov to Jan) Average Average Students total sessions receiving sessions p/w 2+ sessions (SD) (SD) p/w (%) AL 20.3 2.2 18 SCH 1 (5.1) (0.6) (90.0%) AL 22.9 2.5 20 SCH 2 (1.5) (0.2) (100%) WI 32.3 3.0 23 SCH 1 (3.6) (0.3) (100%)				Spring FF (Feb to May)				Full-Year FF (Nov to May)				
	Average	Average	Students		Average	Average	Students		Average	Average	Students	
	total	sessions	receiving	n	total	sessions	receiving	n	total	sessions	receiving	n
	sessions	p/w	2+ sessions		sessions	p/w	2+ sessions		sessions	p/w	2+ sessions	
	(SD)	(SD)	p/w (%)		(SD)	(SD)	p/w (%)		(SD)	(SD)	p/w (%)	
AL	20.3	2.2	18	20	35.7	2.7	19 (100%)	19	27.8	2.4	37	39
SCH 1	(5.1)	(0.6)	(90.0%)		(1.9)	(0.1)			(8.7)	(0.5)	(94.9%)	
AL	22.9	2.5	20	20	36.9	2.8	18	19	29.7	2.6	38	39
SCH 2	(1.5)	(0.2)	(100%)		(5.1)	(0.4)	(94.7%)		(8.0)	(0.3)	(97.4%)	
WI	32.3	3.0	23	23	39.5	2.8	20 (100%)	20	35.6	2.9	43 (100%)	43
SCH 1	(3.6)	(0.3)	(100%)		(2.6)	(0.2)			(4.8)	(0.3)		
Overall	25.5	2.6	61	63	37.4	2.7	57	58	31.2	2.7	118	121
	(6.4)	(0.5)	(96.8%)		(3.7)	(0.3)	(98.3%)		(8.0)	(0.4)	(97.5%)	



Table 6: FE implementation

		Fall FF	(Nov to Jai	n)		Spring FF (Feb to May)				Full-Year FF (Nov to May)					
	Avg.	Avg.	2+	1+		Avg.	Avg.	2+	1+		Avg.	Avg.	2+	1+	
	total	contact	contacts	contact	n	total	contact	contacts	contact	n	total	contact	contacts	contact	n
	(SD)	p/m	p/m	p/m		(SD)	p/m	p/m	p/m		(SD)	p/m	p/m	p/m (%)	
		(SD)	(%)	(%)			(SD)	(%)	(%)			(SD)	(%)		
AL	5.1	1.7	7	13	20	7.4	1.8	7	16	19	6.2	1.8	14	29	39
SCH 1	(4.8)	(1.6)	(35.0%)	(65.0%)		(4.9)	(1.2)	(36.8%)	(84.2%)		(4.9)	(1.4)	(35.9%)	(74.4%)	
AL	5.5	1.9	9	15	20	7.9	2.0	7	15	19	6.7	1.9	16	30	39
SCH 2	(3.5)	(1.2)	(45.0%)	(75.0%)		(6.9)	(1.7)	(36.8%)	(78.9%)		(5.5)	(1.4)	(41.0%)	(76.9%)	
WI	3.3	1.1	4	12	23	4.6	1.1	3	11	20	3.9	1.1	7	23	43
SCH 1	(2.7)	(0.9)	(17.4%)	(52.2%)		(3.1)	(0.8)	(15.0%)	(55.0%)		(2.9)	(0.8)	(16.3%)	(53.5%)	
Overall	4.6	1.5	20	40	63	6.6	1.6	17	42	58	5.5	1.6	37	82	121
	(3.8)	(1.3)	(31.7%)	(63.5%)		(5.3)	(1.3)	(29.3%)	(72.4%)		(4.7)	(1.3)	(30.1%)	(67.8%)	





Future Forward Impact

Overall Results

Descriptive results

Table 7 presents characteristics of students included in the winter and spring analytic samples. Table 8 reports baseline (fall) and winter attendance and assessment results for students in the winter analytic sample. Table 9 reports baseline (fall), winter, and spring attendance and assessment results for students in the spring or full-year analytic sample.

After attrition, fall and spring FF participants continued to have comparable baseline school attendance, SEL, and reading assessment results (Table 8). Students receiving FF in the fall demonstrated slightly higher baseline local reading assessment results than students receiving FF in the spring.

In the winter, after the fall FF program, the attendance and DESSA results of students assigned to received FF in the fall remained similar to those of students assigned to receive it in the spring (Table 8). The biggest change from baseline to winter scores was observed in Star Reading. Students who had participated in the fall FF program scored higher than students who were assigned to received FF in the spring (0.29 standard deviations). This was in comparison to a 0.09 standard deviation difference observed at baseline. Students who received FF in the fall also demonstrated somewhat higher SEL and attendance results. Winter local reading assessment results are not presented because only Alabama schools administered a winter reading assessment.

At the end of the year, students who participated in the spring FF program mostly caught up to students who received it in the fall (Table 9). There were almost no differences in SEL and Star Reading results between the two groups. Although students who received FF in the fall continued to have slightly better school attendance and local assessment results than students who received FF in the spring, the differences between the groups on these outcomes were smaller than they were at the beginning of the year.



Table 7: Characteristics of students included in the winter and spring analytic samples

			Winter			Spring	
		Spring FF	Fall FF	Total	Spring FF	Fall FF	Total
		Group	Group		Group	Group	
Grade Level	KG	20 (33.9%)	21 (33.3%)	41 (33.6%)	20 (34.5%)	21 (33.3%)	41 (33.9%)
	1st	15 (25.4%)	16 (25.4%)	31 (25.4%)	15 (25.9%)	16 (25.4%)	31 (25.6%)
	2nd	17 (28.8%)	21 (33.3%)	38(31.1%)	16 (27.6%)	21 (33.3%)	37 (30.6%)
	3rd	7 (11.9%)	5 (7.9%)	12 (9.8%)	7 (12.1%)	5 (7.9%)	12 (9.9%)
School	AL school 1	19 (32.2%)	20 (31.7%)	39 (32.0%)	19 (32.8%)	20 (31.7%)	39 (32.2%)
	AL school 2	20 (33.9%)	20 (31.7%)	40 (32.8%)	19 (32.8%)	20 (31.7%)	39 (32.2%)
	WI school 1	20 (33.9%)	23 (36.5%)	43 (35.2%)	20 (34.5%)	23 (36.5%)	43 (35.5%)
Race/ Ethnicity	Black	20 (33.9%)	25 (39.7%)	45 (36.9%)	20 (34.5%)	25 (39.7%)	45 (37.2%)
	White	36 (61.0%)	30 (47.6%)	66 (54.1%)	35 (60.3%)	30 (47.6%)	65 (53.7%)
	Other	3 (5.1%)	8 (12.7%)	11 (9.0%)	3 (5.2%)	8 (12.7%)	11 (9.1%)
Gender	Female	31 (52.5%)	35 (55.6%)	66 (54.1%)	31 (53.4%)	35 (55.6%)	66 (54.5%)
	Male	28 (47.5%)	28 (44.4%)	56 (45.9%)	27 (46.6%)	28 (44.4%)	55 (45.5%)
F/R Lunch	No	8 (13.6%)	7 (11.1%)	15 (12.3%)	8 (13.8%)	7 (11.1%)	15 (12.4%)
	Yes	51 (86.4%)	56 (88.9%)	107 (87.7%)	50 (86.2%)	56 (88.9%)	106 (87.6%)
Total		59 (100.0%)	63 (100.0%)	122 (100.0%)	58 (100.0%)	63 (100.0%)	121 (100.0%)

Note: The winter analytic sample for all measures except Star Reading includes 63 students who received FF in the fall.





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

Winter Analytic S	ample	Ba	aseline (fa	ull)	Winter			
		Spring	Fall	Total	Spring	Fall	Total	
		FF	FF		FF	FF		
		Group	Group		Group	Group		
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	Mean	-0.16	0.21	0.03				
Assessment*	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	62	121	59	62	121**	

* The Winter local reading assessment was only available as an outcome in the 2 Alabama schools. **One Fall FF student in Alabama moved to another school after the FF tutoring started

and moved back before FF tutoring ended. This student was not assessed with Star Reading in the winter.



		Ba	aseline (fa	ll)		Winter			Spring	
		Spring	Fall	Total	Spring	Fall	Total	Spring	Fall	Total
		FF	FF		FF	FF		FF	FF	
		Group	Group		Group	Group		Group	Group	
Attendance Rate	Mean	91.52	92.97	92.28	91.77	92.75	92.28	92.36	93.38	92.89
	SD	10.04	8.96	9.48	6.70	7.96	7.37	7.68	6.13	6.91
	n	58	63	121	58	63	121	58	63	121
OT DESSA	Mean	-0.01	0.09	0.04	-0.16	0.09	-0.02	-0.03	0.01	-0.01
	SD	1.00	1.02	1.01	1.09	0.92	1.00	1.05	0.97	1.00
	n	54	60	114	54	60	114	54	60	114
GD DESSA	Mean	-0.09	0.07	-0.01	-0.12	0.13	0.02	-0.01	-0.00	-0.01
	SD	1.01	1.01	1.01	1.04	0.99	1.02	1.04	0.98	1.00
	n	54	60	114	54	60	114	54	60	114
Local Reading	Mean	-0.17	0.19	0.02				-0.06	0.10	0.02
Assessment*	SD	0.83	1.07	0.97				0.94	0.99	0.96
	n	58	63	121				58	63	121
Star Reading	Mean	-0.02	0.08	0.03	-0.17	0.14	-0.01	-0.03	0.01	-0.01
	SD	0.96	1.03	0.99	0.95	1.02	1.00	1.03	0.98	1.00
	n	58	63	121	58	62	120**	58	63	121

Table 9: Fall, v	vinter, and s	spring attendanc	e and assessment	descriptive re	esults for stu	dents in the s	pring or full-	year analy	tic samp	ole
	,			1					/ I	

* The Winter local reading assessment was only available as an outcome in the 2 Alabama schools.
**One Fall FF student in Alabama moved to another school in Oct 2021 after the FF tutoring started and moved back in January 2022 before FF tutoring ended. This student was not assessed with Star Reading in the winter.





Modeling results

Winter statistical modeling results. After adjusting for baseline (fall) attendance and achievement, block effects, and student demographics (equation 1), FF in the fall semester was not found to have statistically significant impacts on school attendance or SEL (Table 10). FF did have a statistically significant impact on Star Reading (B = 0.289, p = .005). Even though they were only available for students in the two Alabama schools, we also examined the impact of FF on local reading assessments (DIBELS). Consistent with Star Reading results, FF had a statistically significant impact on DIBELS (0.31 standard deviations, t = 2.61, p = .012).

Spring statistical modeling results. Models used to measure the impact of FF on spring outcomes controlled for both baseline (fall) and winter assessments (equation 2). Since group assignment was made in the fall, before winter assessments, these analyses are considered exploratory. Generally, you would not control for data that was collected after assignment. It was done in this case in an effort to equalize the two FF groups at the start of groups switching, i.e. before the spring FF group started their participation. Thus, these analyses test the impact of switching from BAU to FF, rather than the impact of being assigned to FF.

Spring FF was not found to have statistically significant impacts on spring school attendance, Star reading or local reading assessments (Table 10). Regarding SEL, the results do suggest that students who received FF in the spring demonstrated greater SEL growth from winter to spring, but these results were not statistically significant.

Full-year statistical modeling results. Finally, we examined the full-year estimate by comparing spring results for the fall and spring FF groups, adjusting only for baseline (fall) attendance and achievement, block effects, and student demographics (equation 3). These analyses examined if end-of-year (spring) assessments differed depending on when students received FF. If the outcomes of the two groups were equivalent at the end of the year, then it likely did not matter when students participated in FF, i.e. students who participated in FF in the fall semester continued to benefit from their participation one semester after it ended (Figure 2). That is exactly what these results suggest (Table 10). Across all five measured outcomes, the fall and spring FF groups were doing equally well at the end of the year. It did not matter when students received FF. All three outcome conditions outlined in Figure 2 were satisfied, suggesting the benefit of the fall FF program was sustained through the end of the school year.



Table 10: Winter, spring, and full year assessment and attendance impacts on students receiving FF in the fall compared to students receiving it in the spring

Outcome		Winter results	Spring results	Full-Year results
		(Fall to Winter)	(Winter to Spring)	(Fall to Spring)
School Attendance	В	0.03%	0.26%	0.05%
	SE	1.232	1.285	1.253
	t	0.026	0.199	0.037
	р	0.979	0.843	0.971
DESSA – Goal	В	-0.034	-0.209	-0.109
Directed Behaviors	SE	0.147	0.140	0.155
(02)	t	0.233	-1.498	-0.701
	р	0.816	0.139	0.485
DESSA - Optimistic Thinking (OT)	В	0.082	-0.174	-0.049
5 ()	SE	0.116	0.124	0.138
	t	0.710	-1.402	-0.356
	р	0.480	0.166	0.723
Star Reading	В	0.289	-0.052	0.023
	SE	0.100	0.125	0.117
	t	2.899	-0.411	0.195
	р	0.005*	0.682	0.846
Local Reading	В	0.305	0.028	-0.007
	SE	0.117	0.135	0.125
	t	2.610	0.211	-0.054
	р	0.012*	0.834	0.957

Notes: The Winter local reading assessment was only available as an outcome in the two Alabama schools. The reference is the Spring FF Group.

* p < .05





FF promoting an interest in reading books

One of FF's goals is making reading enjoyable to students. A student's attitude toward reading should change as they gain confidence as a reader. Further, increasing student interest in reading is considered by the program as a mediator for impacting student reading achievement, i.e. as students develop an interest in reading, they improve their reading abilities.

Classroom teachers were asked to report in the fall, winter, and spring how often each student showed an interest in reading books, *frequently*, *occasionally*, or *never/rarely*. Figure 3 presents the number of students who *frequently* showed an interest in reading books from fall to winter to spring. In the fall, before FF was implemented, roughly the same number of students assigned to receive FF in the fall (32) and spring (31) were reported as frequently showing an interest in reading books. In the winter, after the first semester, more students who had received FF (33) were reported as frequently showing an interest in reading books than students who were to receive it in the spring (22). This suggests that many students who did not participate in the fall FF program became less interested in reading by the end of the semester. However, by the end of the school year, after all students had the opportunity to participate in FF, both fall and spring FF students were again equally likely to frequently show a love of reading. One additional implication of these findings is that the students who participated in the fall FF maintained their love of reading in the semester after their participation had ended.



Figure 3: Number of students frequently showing an interest in reading books



Differential Impact Results

We explored two possibilities regarding student groups who might have benefited from their participation in FF. 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 implementation systems, so it made sense to combine them in a comparison with students in the Wisconsin school. This line of analysis was designed to explore the possibility that differences in program implementation could explain the overall impact of FF. 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).

Winter differential impact results

For winter results (Fall to Winter), we conducted separate statistical models (using equation 1) for Black students, White students, Wisconsin students, and Alabama students for each of the four outcomes (Table 11). The results of these models show the fall FF program had a much larger impact on Star Reading scores of Black students (B = 0.671, p = .005). FF also had a much larger impact on the Star Reading scores of Alabama students (B = 0.473, p = .001). Conversely, the results show negative or null effects on White students and students in Wisconsin.

Spring differential impact results

For spring results, we again performed four separate statistical models (using equation 2) for Black students, White students, Wisconsin students, and Alabama students for each of the four outcomes (Table 12). These models controlled for both baseline data and winter data, after the fall FF group had completed their participation. These analyses are considered exploratory since we controlled for data that was collected after assignment was done in the fall. The purpose for doing this was to try to make the two groups as equivalent as possible to more precisely measure the impact of the spring FF program.

The results of these models again suggest the spring FF program likely had a differential impact on Black student Star Reading scores (B = -0.576, p = .051). The results also suggest the spring FF program may have impacted Wisconsin student OT DESSA (B = -0.305, p = .118) and GD DESSA (B = -0.489, p = .062) scores.





Full-year differential impact results

For the full-year results, we conducted four separate statistical models using equation 3, separately for Black students, White students, Wisconsin students, and Alabama students on each of the five spring outcomes (Table 13). These models examined if end of year results were affected by when students received FF, the fall or the spring. The lack of statistical significance and near-zero coefficients across these outcomes suggest the outcomes for students who received FF in the fall and the spring were about the same. This finding suggests that the impact of FF in the fall semester was likely sustained through the end of the year.



	В	Standard error	t	р
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	0.530
White students	0.85%	2.162	0.393	0.696
WI students	-1.34%	2.517	-0.530	0.601
AL Students	1.59%	1.260	1.260	0.214
OT DESSA				
Black students	0.050	0.250	0.199	0.845
White students	-0.151	0.190	-0.799	0.431
WI students	-0.107	0.160	-0.667	0.511
AL Students	0.155	0.164	0.943	0.351
GD DESSA				
Black students	0.249	0.308	0.807	0.433
White students	-0.267	0.253	1.056	0.299
WI students	-0.139	0.222	-0.625	0.538
AL Students	0.061	0.202	0.303	0.764

Table 11: Winter differential impact estimates of FF group for...

Note: The reference group is spring FF. The results indicate the impact of the fall FF program. * *P* < .05





	В	Standard error	t	р			
Star Reading							
Black students	-0.576	0.256	-2.246	0.051			
White students	-0.068	0.224	-0.302	0.765			
WI students	0.161	0.250	0.643	0.528			
AL Students	-0.119	0.173	-0.687	0.496			
Attendance							
Black students	-0.654%	2.956	-0.221	0.830			
White students	2.208%	2.143	1.030	0.312			
WI students	0.902%	1.681	0.537	0.598			
AL Students	0.629%	1.807	0.348	0.730			
OT DESSA							
Black students	-0.197	0.384	-0.514	0.623			
White students	-0.087	0.131	-0.662	0.514			
WI students	-0.305	0.185	-1.654	0.118			
AL Students	-0.162	0.193	-0.841	0.406			
GD DESSA							
Black students	-0.396	0.491	-0.805	0.447			
White students	-0.164	0.219	-0.746	0.463			
WI students	-0.489	0.244	-2.007	0.062			
AL Students	-0.236	0.187	-1.264	0.214			
Local Assessments							
Black students	0.416	0.357	1.165	0.274			
White students	-0.027	0.231	-0.117	0.908			
WI students	-0.229	0.300	-0.764	0.454			
AL Students	0.030	0.160	0.187	0.853			

Table 12: Spring differential impact estimates of FF group for...

Note: The reference group is spring FF. The results indicate the impact of the fall FF program.



Fall to Spring	В	Standard error	t	р			
Star Reading							
Black students	0.006	0.216	0.026	0.980			
White students	-0.093	0.210	0.445	0.659			
WI students	0.036	0.228	0.156	0.877			
AL Students	0.070	0.145	0.484	0.631			
Attendance							
Black students	-1.57%	2.138	0.735	0.474			
White students	1.93%	2.087	0.924	0.363			
WI students	0.88%	1.451	0.604	0.552			
AL Students	0.45%	1.695	0.267	0.791			
OT DESSA							
Black students	0.221	0.293	0.753	0.466			
White students	-0.236	0.197	-1.196	0.242			
WI students	-0.288	0.192	-1.502	0.149			
AL Students	0.098	0.195	0.505	0.617			
GD DESSA							
Black students	0.054	0.348	0.156	0.879			
White students	-0.328	0.252	-1.303	0.204			
WI students	-0.445	0.266	-1.672	0.110			
AL Students	-0.056	0.200	-0.280	0.781			
Local Assessments							
Black students	0.125	0.206	0.607	0.554			
White students	-0.066	0.223	-0.294	0.771			
WI students	-0.092	0.263	-0.350	0.730			
AL Students	0.043	0.139	0.306	0.761			

Table 13: Full-year differential impact estimates of FF group for...

Note: The reference group is students receiving FF in the spring semester.





FF promoting an interest in reading books in different student groups

As with other outcomes, we examined the differential impact of FF on student motivation for reading for Black, White, Wisconsin, and Alabama students. Figure 4 presents the number of Alabama and Wisconsin students who *frequently* showed an interest in reading books from fall to winter to spring. These data suggest the overall trend presented in Figure 3 was largely driven by the impact of FF in Alabama. Among Alabama students who received FF in the fall, 16 initially frequently showed an interest in reading, which increased to 18 in the winter and 20 in the spring. Among students assigned to receive FF in the spring, 20 initially frequently showed an interest to 11 in the winter. However, after participating in FF in the spring, 20 students again frequently demonstrated an interest in reading. Participation in the Wisconsin FF program was not associated with student interest in reading.

Figure 4: Number of Alabama and Wisconsin students *frequently* showing an interest in reading books by state



Figure 5 presents the number of Black and White students who frequently showed an interest in reading books from fall to winter to spring. These data suggest the overall trend presented in Figure 3 was partially driven by impacts on both Black and White students. Both Black and



White students who received FF in the spring increased their interest in reading books after they participated in FF. With the combination of differential impacts by state and race, these results suggest the overall impact of FF on student interest in reading was likely driven by an Alabama program effect.

Figure 5: Number of Black and White students *frequently* showing an interest in reading books







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

Forty-one of the 45 Black students enrolled in the study attended the two Alabama schools. These 41 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 in the fall and Black students in the spring. To unpack these findings, *within* Alabama we compared the Star Reading scale score growth of Black and White FF students.

Winter Star Reading score growth

Both Black and White Alabama students assigned to receive FF in the fall demonstrated nearly twice as much growth during the fall semester compared to students assigned to receive FF in the spring (Figure 6). This suggests both Black and White Alabama students benefitted from their participation in the fall FF program. Thus, the overall effect of the fall FF program was driven by the differential effect of the Alabama FF program, not the differential effect of FF on Black students.

Figure 6: Fall to winter Star Reading scale score growth in Alabama by race





Spring Star Reading score growth

The spring results suggest a similar pattern. Both Black and White Alabama students who participated in the spring FF program gained more than students who had participated in the fall program (Figure 7). Growth was reduced for every group from the winter to spring than it was from the fall to winter.

Figure 7: Winter to spring Star Reading scale score growth in Alabama by race



Full-year Star Reading growth

The full-year Star Reading score growth of Alabama students suggests Black and White students who participated in the fall or spring FF program grew about the same amount during the year (Figure 8).









FF promoting an interest in reading books in Alabama students

Figure 9 presents the number of Black and White Alabama students who frequently showed an interest in reading books in the fall, winter, and spring. These results show that both White and Black Alabama students who participated in the fall FF program maintained their interest in reading throughout the year. Both White and Black students who participated in the spring FF program decreased their interest in reading from the fall to the winter. However, after they participated in FF in the spring, both Black and White students increased their interest in reading. These results further clarify the overall differential impact of FF on Black students as likely being a function of the differential impact of the Alabama FF program, which included nearly all of the Black students who received FF in the current study.







Understanding the null impact of fall FF in Wisconsin

Although not statistically significant, the finding that Wisconsin students assigned to receive FF in the fall performed worse on the winter Star Reading assessment (B = -0.192) than students assigned to receive FF in the spring needs some clarification (Table 11). There were two circumstances that we believe help explain the negative measured effect of the fall FF program in Wisconsin. First, in Wisconsin the Winter Star assessment for fall FF students was completed by early February, while it was not completed for spring FF students until mid-March.² Thus, the Winter assessment for spring FF students accounted for over a month of additional reading development. Further, switching from the Fall FF to Spring FF occurred at the beginning of February. So, the additional time between assessments also reflected a partial impact of Spring FF on Spring FF students. A second possible explanation for the null finding of the Wisconsin fall FF program was that Wisconsin program assignment was somewhat contaminated by local implementation of Tier II interventions. In the fall, three Spring FF students received tier II

² Fall and spring FF Alabama students were all assessed within the same two weeks.





intervention services from their school. No students assigned to Fall FF did though. 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 business-as-usual instruction, before providing it. Conversations with the site reveal a complication with using the switched-replication design. The site misunderstood the assignment conditions, believing it was acceptable to change regular reading instructional supports to accommodate FF supports. 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 The current study, as implemented in the Wisconsin school, rather than preventing study contamination, the switched-replication design

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 2.7 tutoring sessions each week and all but three students (98%) received the targeted amount of tutoring (at least two sessions each week). Family engagement was implemented with a combination of in-person and virtual events. The average student's family was engaged successfully 1.6 times per month and most families (about 68%) were contacted at least once per month.

Due to our small sample size, our ability to detect program impacts was greatly diminished, only likely to measure large program effects. We did find that the fall FF program had a large, statistically significant, positive impact on winter reading achievement (0.29 standard

³ This was not a problem in Alabama. Alabama sites did not provide Tier II interventions.



deviations). This impact was driven by a large impact on students in the two Alabama schools. We also found that FF had a statistically significant, positive impact on winter local reading assessment results in Alabama (0.31 standard deviations). However, we did not measure statistically significant impact of FF in the spring semester.

There is evidence that FF impacted student attitudes toward reading. In the fall, classroom teachers reported the same number of Fall FF and Spring FF students frequently showing an interest in reading books. In the winter more students who had received FF in the fall than students scheduled to receive FF in the spring frequently showed an interest in reading books. By the end of the year, students who participated in the spring FF program were, again, as likely as students who received FF in the fall to frequently show an interest in reading books. As both groups of students participated in FF, their interest in reading books increased.

Finally, the combination of findings presented in this report provide evidence of a sustained impact for FF. First, we found that the fall FF program had a significant impact on winter reading assessment results. Next, we found that both students who received FF in the fall and spring ended the year with similar outcomes. So while students who received FF in the spring caught up to the students who received it in the fall, they did not do better on end-of-year assessments than students who received FF in the fall. This suggests a likely sustained impact of FF for at least one semester. Ultimately, it did not matter when students received FF. The end of year outcomes of both groups of students were the same. These results, along with our previous sustained impact study of FF (Jones, Reeves, Li, & Gilman, 2021), suggest that students in FF continue to benefit from their participation after they leave FF. This contrasts with the literature of literacy interventions that suggest program impacts typically dissipate after students complete their participation (Hurry & Sylva, 2007). FF is different than other literacy programs in that it is intentionally designed to have a sustained impact. Participant families do not completely leave the program after tutoring ends. Families are more engaged with their student's school and community resources like FF. It is this type of contextual approach to develop student literacy that makes a difference in the lives of students.





Appendix

Table 14: Baseline (fall) and post (winter) attendance and assessment results for Alabama students in the analytic sample

		Baseline (Fall)			Winter		
		Spring	Fall FF	Total	Spring	Fall FF	Total
		FF			FF		
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
	п	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
	п	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
	п	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
	п	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
	п	39	39	78	39	39	78

Note: This table is provided to help assess the rigor of the analysis specific to the fall Alabama FF program.



References

- 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.
- DIBELS. (2021). What are DIBELS? <u>https://dibels.uoregon.edu/assessment/dibels#:~:text=The%20Dynamic%20Indicators%2</u> <u>0of%20</u>
- Epstein, J.L. (2001). School, Family, and Community Partnerships: Preparing Educators and Improving Schools. Boulder, CO: Westview Press.
- Hurry, J. & Sylva, K. (2007). Long-term outcomes of early reading intervention. *Journal of Research in Reading*, 30, 227-248.
- 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.
- 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). 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. <u>https://uwm.edu/sreed/wp-content/uploads/sites/502/2020/10/Regression-Discontinuity-Study-of-Future-Forward.pdf</u>





- Jones, C. J., & Li, D. (2021, March). *Testing the Impact and Scalability of the Future Forward Literacy Program.* <u>https://uwm.edu/sreed/wp-content/uploads/sites/502/2021/03/FF-</u> <u>attendance-impact-evaluation-19-20.pdf</u>
- 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? <u>https://uwm.edu/sreed/wp-</u> 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*. <u>https://uwm.edu/sreed/wp-content/uploads/sites/502/2022/02/FF-2020-21-report.pdf</u>
- 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. <u>https://dibels.uoregon.edu/sites/dibels1.uoregon.edu/files/DIBELS8thEdition_TechRpt18</u> 01_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. <u>https://dibels.uoregon.edu/docs/UnderstandingDIBELS8CompositeScores.pdf</u>

What Works Clearinghouse. (2020). Standards Handbook, Version 4.1.

https://ies.ed.gov/ncee/wwc/Docs/referenceresources/WWC-Standards-Handbook-v4-1-508.pdf