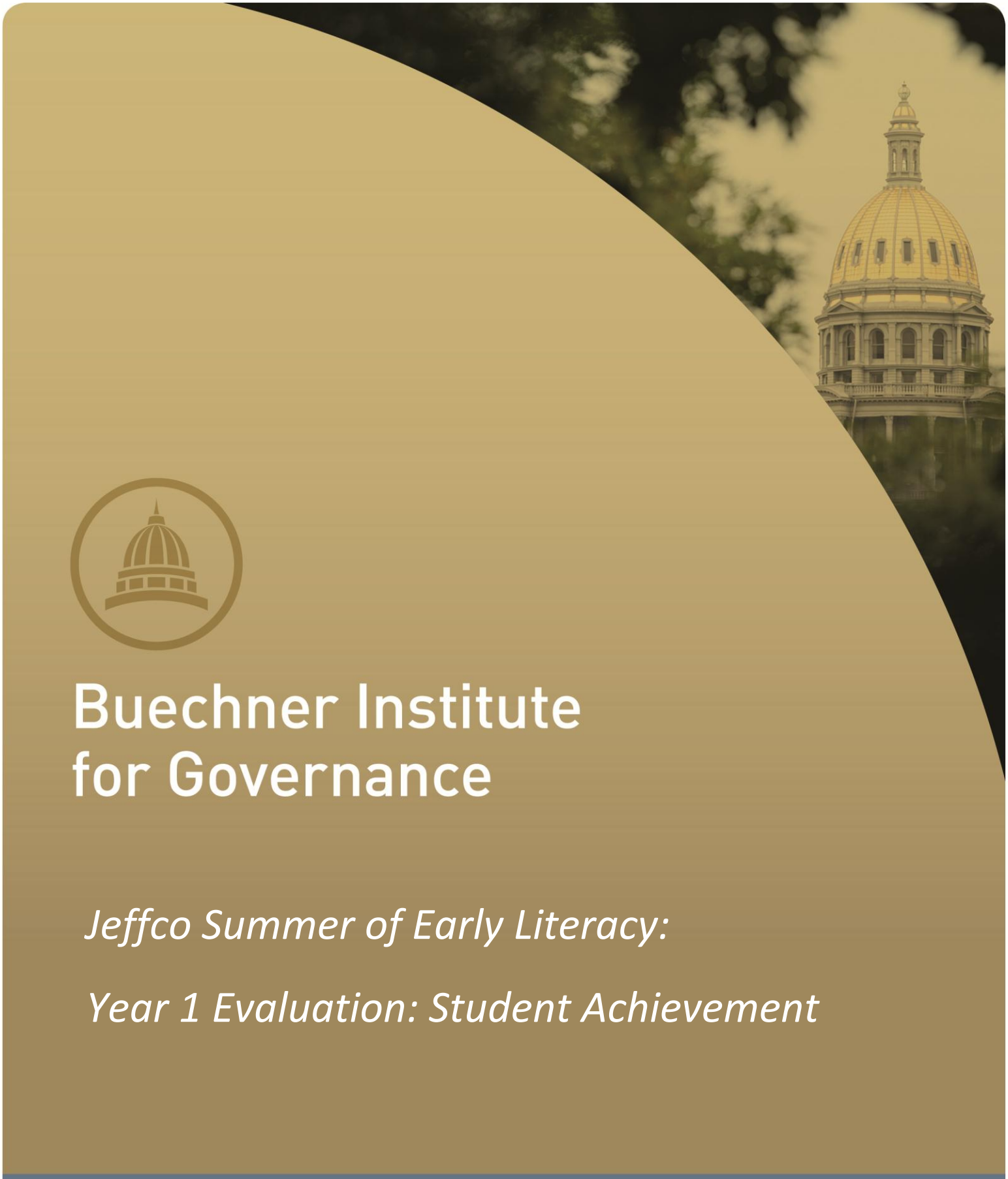


Buechner Institute for Governance

Jeffco Summer of Early Literacy:

Year 1 Evaluation: Student Achievement

School of Public Affairs
UNIVERSITY OF COLORADO DENVER

www.spa.ucdenver.edu/BIG

Abstract

This study evaluates the literacy achievement of students who participated in the inaugural year of the Jeffco Summer of Early Literacy, a six-week, foundation-funded summer program in which certified teachers provide three hours a day of intensive language arts instruction to students attending high-poverty public schools. Propensity-score matching created a control group by pairing 154 program participants entering grades 1 through 3 at two different schools with similar students at six local schools that did not offer the program. A Hierarchical Linear Model-based analysis found that, controlling for prior achievement at the student level and poverty rates at the school level, program participants scored one third of a standard deviation higher than the control group on a standardized measure of the DIBELS Next reading composite score. A binary logistic regression analysis indicated that, after controlling for prior achievement and special education status, program participants were nearly eight times more likely than matched pairs to attain proficiency on the third-grade Transitional Colorado Assessment Program reading exam. Findings should be interpreted with caution because test results were available for fewer than half the program participants and sample sizes were small. Future evaluations will track individual student achievement over time, as the program is scheduled to run from 2012 through 2015.

Summer literacy loss is a well-established problem, especially for low-income children, whose reading skills often stagnate or decline during long vacation breaks (Cooper, Nye, Charlton, Lindsay, & Greathouse, 1996; Entwisle, Alexander, & Olson, 1997). Such literacy losses help explain the achievement gap between low-income children and peers from more affluent families that are able to provide a more literacy-rich environment during the summer months (Reardon, 2011). Further, early literacy challenges can reverberate throughout a child's life. Students who are not reading proficiently by the end of third-grade are four times more likely to drop out of high school than are those reading proficiently by that time (Hernandez, 2011). The risk is even greater for non-proficient, low-income black and Hispanic students, who are eight times more likely to drop out than their peers who reached proficiency by grade 3 (Hernandez, 2011).

Summer literacy programs have had positive effects on nearly every aspect of reading achievement (Kim & Quinn, 2013). Effects have been particularly strong for a subgroup of students most at risk for falling behind, I. e. low-income children(Kim & Quinn, 2013).

Past research suggests that classroom-based summer literacy programs boosts student achievement by remediating past weaknesses and/or previewing skills and materials that will be taught during the following school year (Cooper, Charlton, Valentine, & Muhlenbruck, 2000; McCombs et al., 2011). This report examines achievement impacts of the first-year of one such reading program, the Jefferson Foundation's Jeffco Summer of Early Literacy (JSEL). As such, the goal is to explore the responses to the following research questions:

1. How does the achievement of JSEL program participants compare to the achievement of a similar group of students who did not participate?

2. How does the achievement of at-risk subgroups (low-income students, special education students and English learners) in the JSEL program compare to the achievement of JSEL participants who are not members of these sub-groups?

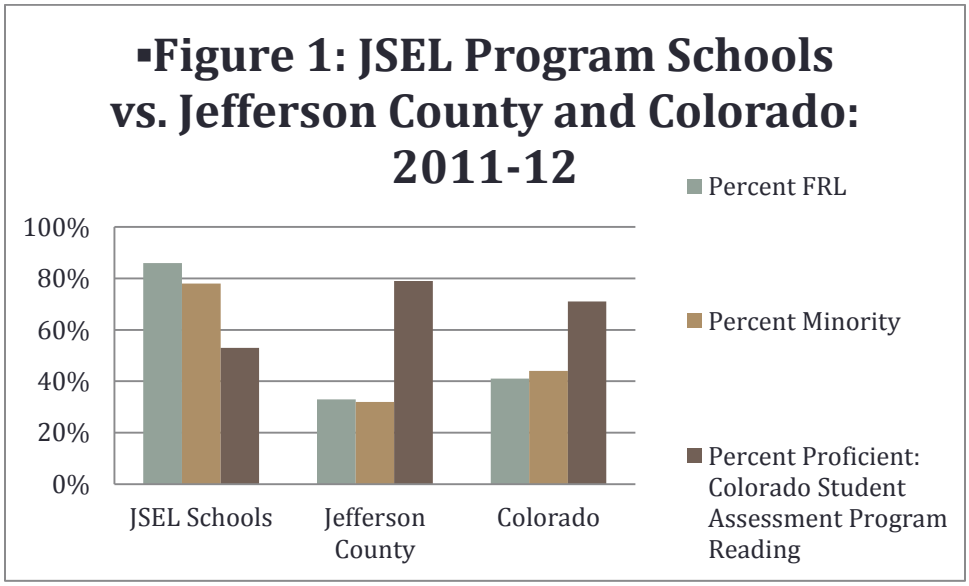
This evaluation begins by describing the setting, approach and demographics of the JSEL program and its participants. The report next explains the formation of the control group and the disposition of two outcome measures of interest: the DIBELS Next literacy test and the third-grade Transitional Colorado Assessment Program reading exam. A section on the analytic approach outlines the types of regression methods and descriptive statistics used to analyze the data. A results section describes the achievement outcomes identified by these methods. A final, concluding section places these results in the context of previous findings related to summer literacy programs.

Methodology

Setting

On average, Jefferson County, Colorado's 86,000 students have higher rates of reading achievement than their classmates in the rest of the state. They are also less likely to fall into subgroups with lower, average student achievement. However, these averages mask the diversity of the district's eastern region, an inner-ring of Denver suburbs that

house high proportions of lower-achieving low-income and minority students. The JSEL program serves early elementary students at four of these eastern ring schools: Edgewater, Lasley, Molholm and Swanson. Figure 1 contrasts the demographics of these four schools with those of public school students in Jefferson County and the rest of the state (Colorado Department of Education, 2011a, 2011b, 2012a, 2012b) Based on past research indicating that low-income students experience greater summer learning loss and, as a result, may benefit more than affluent students from summer literacy programs, the Jeffco Foundation and the Jefferson County School District made a strategic decision to focus their limited resources on offering the JSEL program at these four schools (Gershenson, 2013; Kim & Quinn, 2013).



Treatment

JSEL is a voluntary, intensive six-week literacy program offered during the summers of 2012 through 2015. Edgewater, Lasley, Molholm and Swanson students entering grades K-3 are eligible for the program. Children who enroll in the program are placed in small classes of no more than 15 students. They receive individual and small-

group instruction, coaching through the Oasis Tutoring program, and enrichment from the Spellbinders storytellers(Jefferson Foundation, 2012). They also attend field trips that were then used as the basis for further reading instruction. For students, the program lasts four hours per day. The schedule includes three hours of literacy instruction, plus a total of two hours for a breakfast and lunch that is served at no cost to families. The schedule mirrors the structure of the school day. It includes 100 to 125 minutes of reading.

JSEL also has a family outreach component. Parents are asked to sign a literacy promise. They are invited to breakfast and lunch during each day of the program so they can develop a relationship with their child's teacher. The program also hosts two parent nights.

Teachers, all of whom are employed at one of the four schools, also receive two full days of literacy-related professional development before the program began, plus weekly training sessions developed at their sites (Jefferson Foundation, 2012).

Participants

In Year 1 (Summer 2012), 326 K-3 children signed up for the JSEL program after being recruited by principals, teachers, take-home folders, parent nights and other regular communication channels between school and home. Preschoolers were invited to participate when they registered for kindergarten. Although enrollment was voluntary, participant demographics closely resembled demographics at the four schools.

This study does not focus on the full population of 326 JSEL participants. Rather, it analyzes achievement results from a subset of 154 students who enrolled in the programs

at Edgewater and Lasley and also had information available on their test scores, ethnicity and free and reduced-price meal status. This study focuses on Edgewater and Lasley students because both those schools piloted a new form of the Dynamic Indicators of Basic Early Literacy Skills (DIBELS Next) in 2011-12, one year before the entire district adopted this new version of an older assessment. As a result, these two schools had comparable pre and post-test results for the 2011-12 school year (before the program began) and the 2012-13 school year (after the program ended and the entire district adopted DIBELS Next). The two other JSEL schools (Swanson and Molholm) did not have comparable pre and post-test reading results for Year 1 of the program. Test scores from these two schools should be available for evaluations of years 2 through 4 of the program. Also excluded from this analysis are students who were Kindergartners in 2012-13. These students did not have pre-scores available either because they were not yet enrolled in public school or because their preschool programs did not provide DIBELS Next pre-test scores.

Propensity- Score Matching

This analysis employed Propensity -Score Matching (PSM) to pair each of the 154 Edgewater and Lasley JSEL participants with 154 similar students at six other Jefferson County elementary schools that also piloted DIBELS Next in 2011-12. PSM is used in situations in which it is not possible to control for differences in the treatment and control groups by randomly assigning participants to each group. PSM generally uses a logit or probit function to create a propensity score that represents the likelihood that a person would be assigned to the treatment group if this had, indeed, been a randomized, controlled experiment, which it was not (Heinrich, Maffioli, & Vazquez, 2010) . Observable characteristics such as demographics or test scores are used to predict membership in the

treatment group. This study employed a type of PSM called neighbor covariate matching. This method matches individuals from the "treatment" group (i.e. JSEL participants) with individuals from the "control" group who are most similar in terms of the observed characteristics used to generate the propensity scores (Heinrich et al., 2010).

The PSM process employed three variables to match JSEL participants with non-participants: These variables were minority status, whether or not the student was from a family with an income sufficiently low to qualify for the federal free or reduced-price federal meal program (FRL) and DIBELS Next composite scores based upon tests given prior to the treatment group's intervention (Table 1). Table 1 demonstrates that the PSM process resulted in the selection of a control group that did not differ significantly from the treatment group on any of those three measures.

Table 1
Descriptive Statistics for JSEL Participants (Treatment) and PSM Pairs (Control)

Variable	**Control**	**Treatment**	**Total**
Number per 2011-12 Grade Level			
K	47	47	94
1	54	54	108
2	51	51	102
3	2	2	4
Percent White	15	14	15
Percent FRL	79	80	79
Percent Special Education	8	9	9
Percent Gifted and Talented	2	4	3
Percent English Learners*	36	54	45
Percent days attended: JSEL program	N/A	14	N/A
DIBELS Next Z score			
Pretest (SD)	0.01 (.96)	-0.01 (1.04)	0 (1)
Post-test* (SD)	-0.09 (1.03)	0.09 (.96)	0 (1)
Grade 3 CSAP Reading: 2013			
Percent Proficient/Advanced*	51	76	67
Mean Scale Score (SD)*	520 (87)	534 (91)	527 (89)

Note.
* Differences between the treatment group and the control group were significant at F=p<.05 for continuous variables and z=p<.05 for categorical variables.

Although JSEL participants were similar to matched students prior to the intervention, their post-test results are significantly better: JSEL students earned higher scores on the DIBELS Next post-test (Cohen's d= .18). They are also significantly more likely to score proficient or advanced on the third-grade TCAP exam given the spring after the intervention. Scale scores for that same exam were also higher for JSEL participants, although the difference was not statistically significant.

DIBELS Next

The key outcome measure for this study was DIBELS Next, an assessment designed to measure whether children in grades Kindergarten through grade six have acquired early literacy skills. DIBELS Next consists of six possible subtests. (See Table 2.) All of the assessed skills are essential components of early literacy (National Reading Panel, 2000; National Research Council, 1998). As of the 2012-13 school year, DIBELS Next is administered by Jefferson County teachers at least twice per year, in grades K through 3.

This study uses the DIBELS Next Composite Score, a combination of multiple subtest scores that is described by the test authors as "the best overall estimate of the student's early literacy skills and/or reading proficiency" (Good et al., 2013, p. 9). Test authors state that appropriate uses of DIBELS Next include assessing individual student progress over time and examining "the effectiveness of a school's system of instructional supports" (Good et al., 2013, p. 6). This study makes use of this test in these ways by comparing individual student progress toward the acquisition of key literacy skills and assessing the effectiveness of the JSEL instructional support.

For the purposes of the Year 1 analysis, we did not have a big enough sample size to run separate models for each grade level. This posed an obstacle because DIBELS Next composite scores at different grade levels are calculated using different subtests.[1] (See Table 2.)

[1] The meaning of the score also changes with time of year: For instance, a score of 26 that is considered proficient at the beginning of kindergarten is a cause for concern by the end of the year. Treatment and control group participants took DIBELS tests at the same time of year, so the time-of-year issue was not a concern for this study.

Table 2
Subtests used to Calculate DIBELS Next Composite Scores for Each Grade Level

Subtest	Grade Level				
	K	1	2	3	4[a]
First Sound Fluency	X				
Phoneme Segmentation Fluency	X	X			
Nonsense Word Fluency: Correct Letter Sounds	X	X	X		
Nonsense Word Fluency: Whole Words Read		X	X		
DIBELS Oral Reading Fluency: Words Correct		X	X	X	X
DIBELS Oral Reading Fluency: Accuracy		X	X	X	X
Retell		X	X	X	X
Retell Quality of Response			X	X	X
Daze Adjusted Score (Reading Comprehension)				X	X

Note.
a Although the JSEL program targets students entering grades K through 3, a small number of participants were entering fourth-grade, including two treatment group and two control group students.

The study addressed the issue by conducting separate rounds of propensity score matching for each grade level. As previously explained, the matching process used DIBELS Next composite scores compiled prior to the intervention as one of three variables to pair treatment group participants with members of a control group. Once the matching process was complete, the pre-test composite scores of treatment and control group participants alike were combined. The entire group (of JSEL participants and their matched pairs) was then normed at each grade level. This process created, from the pretest composite scores, four sets of Z scores, one for each grade level, from grades kindergarten through 3. Z scores are standardized scores with a mean of 0 and a standard deviation of 1. The Z scores were then combined for all grade levels and re-normed for the entire group of 308 students, in grades K through 3. The re-norming process produced a new set of Z scores that was

appropriate to use to compare children in different grade levels because it accounted for their relative positions within their own grades. The same process was then repeated for treatment and control groups, for the DIBELS Next post-test. The end result was two sets of Z scores. The pre-intervention scores used to control for the students different starting points. The post-intervention scores became the outcome variable. Table 1 summarizes these scores for treatment and control groups.

Transitional Colorado Assessment Program

An additional outcome measure was the Transitional Colorado Assessment Program (TCAP) reading score for 99 treatment and control group students who were third-graders in the spring of 2013. TCAP is Colorado's official state assessment. The test, which included both multiple choice and constructed response items, is created by CTB McGraw Hill. Although the 2013 technical report was not yet available as of the time of this report, the 2012 report indicates that third-grade reading scale scores ranged from 150 to795, with a median of 569, a mean of 563 and a standard deviation of 75 (CTB McGraw-Hill, 2012). The scale scores are used to derive four proficiency levels: unsatisfactory, partially proficient, proficient and advanced. In 2013, 73 percent of Colorado third-graders and 80 percent of Jefferson County third-graders were proficient or advanced. Table 1 summarizes CSAP results for the study participants who took the exam.

Analytic Approach: HLM

This study employed Hierarchical Linear Modeling (HLM). This form of regression analysis is appropriate for this evaluation because the data is structured such that students are nested in schools, meaning it is reasonable to assume that students enrolled in the

same school may share certain commonalities(Raudenbush & Bryk, 2001). While other approaches such as OLS regression require observations (e.g. student test scores) to be independent, HLM does not make this assumption. Rather, in HLM, dependence between observations is modeled as part of the analysis. This permits researchers to decompose the variation attributable to differences between schools versus the variation attributable to differences between individual students within a school. A second reason for using HLM for this evaluation is that analyses of future years will examine growth over time as JSEL students' progress through the elementary grades. HLM can also accommodate such longitudinal growth modeling, with multiple observations spanning multiple time periods nested within each individual student subject, who is in turn nested within her individual school (Raudenbush & Bryk, 2001).

The dependent variable of interest was the DIBELS Next composite score, standardized as described above. The analytic approach for both dependent variables began with an unconditional means model. The unconditional means model contains no predictor variables. It is designed to parse to Levels 1 and 2 the base levels of variance found in the dependent variable. For both dependent variables, both of which measure student achievement, the unconditional means model for student i in school j is

$$DIBELS\ Next_{ij} = \gamma_{00} + u_{0j} + r_{ij}$$

in which γ_{00} is a fixed effect representing mean, school level achievement, u_{0j} is a school level variance component and r_{ij} is a student-level variance component.

From there, the key variable of interest (whether or not the student was in the treatment group or control group) was added at Level 2. Treatment status was treated as a school-level variable because all the students drawn from a particular school either

attended JSEL (in the case of Edgewater and Lasley) or did not (in the case of the six other Jefferson County elementary schools that piloted DIBELS Next in 2011-12). Next, the remaining independent variables were added, one by one. If non-significant, they were then excluded from the model. This was because the small sample size

At Level 1, the independent variables tested for significance were standardized DIBELS Next pretest scores, special education status, gifted and talented status, English language learner status, attendance rate (with non-participants coded as zeroes), FRL status, and race. Table I provides descriptive statistics on all these Level 1 variables as well as on the dependent variable, the standardized DIBELS Next post-test.

Table 3 provides descriptive statistics for the available Level 2 variables. In addition to treatment status of the school, these variables were: the percentage of students enrolled in special education; the percent of students proficient or advanced in TCAP reading; the total number of students enrolled in the school (school size); the percent of points earned from the state accountability system; and the percentage of students who were white, low-SES, or English learners.

Table 3
Descriptive Statistics for School-level Indicators (Level 2)

School	FRL	N STUDY	N TOTAL	TCAP[a]	PTS[b]	WHITE	ELL	SPED
Bergen Meadow	9%	4	336	91%	80%	88%	3%	7%
Deane	85%	39	501	66%	63%	19%	42%	4%
Edgewater[c]	93%	108	464	65%	78%	53%	47%	13%
Kendallvue	37%	3	541	77%	90%	72%	2%	9%
Kendrick Lakes	33%	15	447	82%	74%	71%	7%	8%
Lasley[c]	82%	100	554	59%	62%	20%	35%	5%
Ryan	24%	18	524	87%	78%	66%	11%	7%
Weber	37%	21	507	75%	75%	70%	8%	9%
Total (Mean)	50%	39	484	75%	75%	57%	19%	8%

Note.
a. Schoolwide percent proficient and advanced in TCAP reading, 2013
b. Percent of points earned from statewide accountability system based on achievement and growth over time.
c. Edgewater and Lasley are treatment schools. The remaining schools are control schools

It is important to note that the sample sizes for these analyses are extremely small, especially when it comes to Level 2 of the HLM Model. For example, Table 3 shows that the PSM process selected just four students from Bergen Meadow Elementary and three from Kendallvue. The problem is even more severe when it comes to measuring student achievement using grade three TCAP reading results. Just 99 total students in the sample were the right age to take that exam in 2013. No Bergen Meadow or Kendallvue students in the sample were the right age to take that exam. Given these limitations, HLM was simply not an appropriate method for analyzing Year 1 TCAP outcomes. Instead, TCAP results are modeled using non-hierarchical regression methods. One analysis used OLS regression to model TCAP scale score achievement. Another used binary, regression to predict whether or not a student would be proficient (0= unsatisfactory or partially

proficient, 1= proficient or advanced). All the variables previously described as Level 1 indicators were considered for inclusion in these models. Each was added one at a time then rejected if not significant.

A final caveat is that none of the regression models described above contained interactions because the small sample size rendered them statistically and/or substantively insignificant. As a result, the HLM model, for example, controlled for FRL status and pre-program achievement levels but did not provide information on whether JSEL program differentially made a bigger (or smaller) difference for FRL participants or for those achieving at higher or lower levels prior to the intervention. In order to get a general sense of whether JSEL impacted some groups differently than others, the analysis examined descriptive statistics for subgroups within the treatment group. (For example, how did the achievement of English learners who participated in JSEL compare to the achievement of non-English learners who participated in JSEL?) In order to facilitate this comparison, the DIBELS Next pre-test and post-test scores were re-normed by grade level, and then as a pool, (as described above) against other members of the treatment group, i.e. control group members were excluded from the norming process. Pre-test Z scores were subtracted from post-test Z scores to create a rough measure of growth.

Results

The final HLM model selected to predict the DIBELS Next achievement of JSEL and non-JSEL participants is represented at Level 1 as:

$$ZALLPOST_{ij} = \beta_{0j} + \beta_{1j}*(ZALLPRE_{ij}) + r_{ij}$$

This model represented the DIBELS Next achievement of student i in school j, controlling for that student's score on the DIBELS Next pre-test (β_{1j}) . r_{ij} represents the variance remaining at Level 1. At Level 2, the model took the form of

$$\beta_{0j} = \gamma_{00} + \gamma_{01}*(TREAT_j) + \gamma_{02}*(SCHLFRL_j) + u_{0j}$$

$$\beta_{1j} = \gamma_{10} + u_{1j}$$

$$\beta_{2j} = \gamma_{20} + u_{2j}$$

in which γ_{01} represented whether the school was a treatment (JSEL) or control school and γ_{02} represented the percentage of FRL students at each school. Variance components were estimated at the school level for the intercept, and mean FRL and DIBELS pretest achievement. None of the independent variables excluded from the model were statistically significant.

The model predicts that JSEL participants performed significantly better (.31 of a standard deviation) than non-participants after controlling for school-level FRL and pre-intervention achievement. Table 4 summarizes model results.

Table 4
Fixed Effects (Top) and Variance Components (Bottom) for Unconditional Means Model and Final Model Predicting DIBELS Next Post-test Achievement

Parameter	UMM	Final Model
Intercept	0.06 (.11)	.28 (.11)
Level 1		
DIBELS Prescore		.91*** (.06)
Level 2		
Treatment		.33** (.09)
FRL		.91*** (.06)
	Variance Components	
Level 2		
Intercept u0 (ICC)	0.06 (6%)	0.0004
Pseduo R^2		0.99
Level 1		
Intercept *r (*ICC*)*	0.96 (94%)	0.22
DIBELS Prescore, u1		0.01
Pseudo R^2		0.74

Note.
**= p< .05. *** p< .001
ICC= Intra-class correlation

The intra-class correlation attached to the unconditional means model, also summarized in Table 4, indicated that the vast majority of the variability (94 percent) was

at the first or individual student level. The final model explained most of the variability at both levels, including 99 percent at Level 2 and 74 percent at Level 1.

On average, JSEL participants earned higher TCAP reading scale scores than did non-participants. (See Table 1.) This difference was not statistically significant. OLS regression models that controlled for prior achievement and other factors also failed to identify any statistically significant differences between the scale scores of treatment and control groups.

The descriptive statistics provided in Table I did suggest that treatment students were significantly more likely to be proficient and advanced than were students in the control group. A binary logistic regression model examined these results in further depth, controlling for pre-intervention achievement and special education status. That model took the following form:

$$\eta i = 1n\left(\frac{\Pr(Proficient\ or\ Advanced)}{1-\Pr(Proficient\ or\ Advanced)}\right) = \beta\ 0 + B1(DIBELS\ Prescore)i + \beta 2\ (SPED) + \beta 3\ (TREATMENT) + \epsilon i$$

This model predicted that, controlling for pre-intervention achievement and special education status, the 50 Year 1 JSEL students who took the third-grade TCAP in 2013 were nearly eight times more likely to reach proficiency than were the 49 non-participants. (Table 5).

Table 5

Predictors of Proficiency on 2013 Grade 3 TCAP Reading

Parameter	B	S.E.	Exp (B)
Intercept	0.16	0.38	1.17
DIBELS Prescore	1.89***	0.39	6.61
Treatment	2.05***	0.66	7.75
Special Education	-3.46**	1.58	0.03
Pseudo R2	.59		

Note.
**<.05. ***p<.001

As previously explained, sample sizes were too small to include interactions in the two regression models described above. However, descriptive statistics suggest that one subgroup of students made significantly less growth than their peers who attended the JSEL program: Special needs students made less growth than non-special needs JSEL participants when the DIBELS Next pre-test and post-test scores were compared. This result should be interpreted with caution as data was only available for nine special education students. There were no significant differences between the growth rates of JSEL participants in other subgroups including FRL, gifted and talented, English learners, white students, students with below-average attendance or students in different grade levels. This result should also be interpreted with caution since these groups were also small.

Conclusion and Discussion

The Year 1 results of the Jeffco Summer of Early Literacy suggest that the program is a promising intervention: Controlling for prior achievement levels and a measure of school-

level poverty, the average JSEL participant started the 2012-13 school year with reading achievement results that were nearly a third of a standard deviation higher than the results from similar students who did not attend the program. Despite the small sample sizes, this is a meaningful difference, especially given that the program lasted just six weeks. Also encouraging is that, in their recent meta-analysis of 26 classroom-based summer reading interventions, Kim and Quinn (2013) found a mean effect size of Cohen's d= .09 (for overall reading achievement). For JSEL, the corresponding effect size on DIBELS Next composite reading achievement was .18, or twice that amount.

Also encouraging is the fact that, controlling for prior achievement and special education status, the average JSEL participant who took the third-grade reading TCAP exam in the spring of 2013 was nearly eight times more likely to earn a score of proficient or advanced than was a similar pool of non-program participants. This suggests that the impact of the program may endure since the TCAP was taken more than six months after the intervention ended.

Unlike past research studies, this evaluation did not find that JSEL had a different and/or larger impact on students from low-income families (Kim & Quinn, 2013): JSEL participants who qualified for free or reduced-price meals made comparable achievement gains to those who did not. However, this finding should be interpreted with caution because of the small sample sizes. In addition, virtually all of the participants (and control group members) qualified for free or reduced-price meals. Since their schools drew from the same neighborhood zones, the non-FRL JSEL participants, who comprised just 20 percent of the treatment group, were likely from similar socio-economic backgrounds as their FRL peers even if they were not currently eligible for free or reduced-price meals.

This demographic uniformity is based on a strategic and research-based decision to focus the JSEL intervention on schools that serve large numbers of students from low-income families. However, an unintended consequence of this decision is that it may limit this study's ability to tease out whether or not the program is differentially impacting low-income students.

As previously explained, findings about special education program participants (whose achievement lagged behind non-special education participants on the DIBELS Next) should also be interpreted with caution since so few (n=9) treatment group students were classified as special needs. Future analyses involving larger numbers of students may be able to better identify differential impacts for special needs students and also for English learners, whose achievement gains in Year I did not differ significantly from the gains of non-English learners.

Yet the achievement of treatment group members overall did differ from the achievement of the control group members. Despite the small sample size, this result is clearly visible not just from the regression equations described in the results sections but from the simple, descriptive statistics presented in Table 1. As JSEL continues down the path of summer literacy instruction, all signposts point to its ability to help boost the reading achievement of the students it serves.

References

Colorado Department of Education. (2011a). *Fall 2011 pupil membership by school, ethnicity, gender and grade* Retrieved from: http://www.cde.state.co.us/cdereval/rv2011pmlinks.htm

Colorado Department of Education. (2011b). *K-12 free and reduced lunch eligibility by district, and school* Retrieved from: http://www.cde.state.co.us/cdereval/rv2011pmlinks.htm

Colorado Department of Education. (2012a). *2012 TCAP state summary results: Reading Grades 3-10*. Retrieved from: http://www.cde.state.co.us/assessment/CoAssess-DataAndResults#summarydata

Colorado Department of Education. (2012b). Schoolview Data Center. from https://edx.cde.state.co.us/SchoolView/DataCenter/reports.jspx?_afrLoop=16111942205420299&_afrWindowMode=0&_adf.ctrl-state=jxn16fi9f_4

Cooper, H., Charlton, K., Valentine, J. C., & Muhlenbruck, L. (2000). Making the most of summer school: A meta-analytic and narrative review. *Monographs of the Society for Research in Child Development, 65*(260), 1-118.

Cooper, H., Nye, B., Charlton, K., Lindsay, J., & Greathouse, S. (1996). The effects of summer vacation on achievement scores: A narrative and meta-analytic review. *Review of Educational Research, 66*, 227-268. doi: 10.3102/00346543066003227

CTB McGraw-Hill. (2012). Transitional Colorado Assessment Program: Technical report 2012. Denver, CO: Colorado Department of Education.

Entwisle, D. R., Alexander, K. L., & Olson, L. S. (1997). *Children, schools and inequality*. Boulder, CO: Westview Press.

Gershenson, S. (2013). Do Summer Time-Use Gaps Vary by Socioeconomic Status? *American Educational Research Journal*. doi: 10.3102/0002831213502516

Good, R., Kaminski, R., Dewey, E., Wallin, J., Powell-Smith, K., & Latimer, R. (2013). DIBELS Next technical manual. Eugene, OR: Dynamic Measurement Group.

Heinrich, C., Maffioli, A., & Vazquez, G. (2010). A primer for applying Propensity-Score Matching: Impact-evaluation guidelines. Washington, DC: Inter-American Development Bank.

Hernandez, D. J. (2011). Double jeopardy: How reading skills and poverty influence high school graduation. . Baltimore, MD: Annie E. Casey Foundation.

Jefferson Foundation. (2012). Jeffco summer of early literacy: A formative evaluation. Lakewood, CO: Jefferson Foundation.

Kim, J., & Quinn, D. (2013). The effects of summer reading on low-income children's literacy achievement from Kindergarten to grade 8: A Meta-analysis of classroom and home interventions. *Review of Educational Research, 83*(3), 386-431. doi: 10.3102/0034654313483906

McCombs, J. S., Augustine, C. H., Schwartz, H. L., Bodilly, S. J., McInnis, B., Lichter, D. S., & Cross, A. B. (2011). Making summer count: How summer programs can boost children's learning. Santa Monica, CA: RAND Corporation.

National Reading Panel. (2000). Teaching children to read: An evidence-based assessment of the scientific

research literature on reading and its implications for reading instruction. Washington, DC: National Institute of Child Health and Human Development.

National Research Council. (1998). *Preventing reading difficulties in young children*. Washington, DC: National Academy Press.

Raudenbush, S., & Bryk, A. (2001). *Hierarchical linear models: Applications and data analysis methods: 2nd edition*. Thousand Oaks, CA: SAGE Publicarions.

Reardon, S. (2011). The widening academic achievement gap between the rich and the poor: New evidence and possible explanations. In G. Duncan & R. Murname (Eds.), *Whither opportunity?* (pp. 91-116). New York, NY: Russell Sage.

For more information on the Buechner Institute for Governance, please visit out website and learn more about:

- Current research projects
- Areas of expertise
- Leadership and professional development programs
- Our public policy forum series and other hosted events
- Solutions—Health policy journalism, ideas and analysis
- The Buechner Institute for Governance Advisory Board

www.spa.ucdenver.edu/BIG

303-315-2490

big@ucdenver.edu

Buechner Institute for Governance

Lead. Solve. Change.

School of Public Affairs

UNIVERSITY OF COLORADO **DENVER**