

Registry of Efficacy and Effectiveness Studies

Study Title:

473 Efficacy of the Burst early reading intervention

Registry ID: 473.1v3

Version History

Changes were published on November 14, 2018 5:40:04 PM EST

Currently viewing this version.

Description of changes published:

Prior to initial posting of the analysis plan, as well as to the posting of the first revision, preliminary statistical analyses were conducted without use of outcomes within the treatment group. This round of revisions to the analysis plan were made following our first pass at estimating a "main effect" of the treatment, and address several issues we became aware of during this process. A detailed list of the current changes is included with comments to Section VIII, at the end of this registry entry.

Changes were published on October 22, 2018 12:30:06 PM EDT

[Review this version.](#)

Description of changes published:

Made modest changes to analytic plan prior to analysis.

The first version of this entry was published on October 19, 2018 9:11:21 AM EDT

[Review this version.](#)

Section I: General Study Information

PI name: Brian Rowan

PI affiliation: University of Michigan

Co-PI name: Ben Hansen

Co-PI affiliation: University of Michigan

Primary Funding Source(s):

Institute of Education Sciences

National Center for Education Research

Award Number(s):

R305A120811

IRB Name:

University of Michigan Health Sciences and Behavioral Sciences Institutional Review Board

IRB Approval Date:

2014-03-27

IRB Approval Number:

IRB00000246

Other Registration Name:

-

Other Registration Date:

-

Other Registration Number:

-

Study Start Date:

2012-09-01

Study End Date:

2018-10-31

Intervention Start Date:

-

Timing of entry:

Prior to analysis of outcome data

Brief Abstract:

Researchers will test the efficacy of the intervention by conducting a four-year longitudinal, randomized control trial. The intervention will be delivered to students in Grades K-3 across four years, thereby allowing some students to receive the intervention for more than one year (e.g., those in kindergarten during Years 1 and 3 will receive the intervention for four and two years, respectively). Implementing teachers will receive training in use of the software, which is intended to increase their ability to make data-driven instructional decisions for individual students with confidence. The researchers will also collect information about instructional fidelity and classroom activities.

Keywords:

beginning reading; individualized instruction; computer-based intervention

Comments:

Elaboration of response to timing of registry entry prompt: This registry entry is posted prior to comparative analysis of treatment and control groups' SEL scores, the outcome for each of the confirmatory and exploratory research questions listed here. However, study team members have examined the SEL outcome for control group members (only), and other study group members have compared treatment to control in terms of outcomes not discussed in this registry entry (e.g. teacher survey responses).

Section II starts on the next page.

Section II: Description of Study

Type of Intervention:

Curriculum/Product, Professional Development

Topic Area of Intervention:

Education Technology, Reading and Writing

Number of intervention arms:

1

Target school level:

1, 2, 3, Kindergarten

Target school type:

Rural, Suburban, Urban

Location of Implementation:

United States: United States : Northeast, United States : South, United States : West

Further description of location:

Fifty-two K-3 serving schools spread across 11 districts in the states of AR, CA, CO, CT, GA, KY, LA, MD, TN, and WV.

Brief Description of Intervention Condition:

Burst:Reading is a fully developed software intervention that both assesses reading and related skills (e.g., phonological awareness) of students in grades K-3 and provides aligned instructional content. Teachers use handheld and desktop devices to administer reading assessments individually to students that cover multiple skill categories. This allows the research-based and technology-driven algorithms to create profiles of each student's skills and instructional needs. Additionally, the algorithms consider classroom characteristics (e.g., staffing capacities, time constraints, etc.) and generate small-group suggestions based on students' skill profiles.

The system offers lessons with appropriate material for the various groups, with instructional cycles lasting about 9-18 days, followed by progress monitoring assessments, resulting in additional material covering the same skills or new material (and possibly grouping) covering new skills.

In addition to modules related to instruction and assessment, the intervention also includes a coherent set of professional development modules covering content knowledge and instruction on using the intervention.

Brief Description of Comparison Condition:

Schools assigned to the control condition will not receive the Burst intervention, although they will be given free use of the DIBELS:Next assessment, which they may or may not otherwise have used to diagnose progress among early readers, and they agree to administer it to K-3 students. In other respects they will operate under typical conditions, which could include the use of technology-based reading assessments, and/or technology-delivered instructional content.

Comparison condition:

Business-as-usual

Comments:

-

Section III: Research Questions

Confirmatory research questions:

Question 1:

Among experimental schools assigned to treatment, how did Burst:Reading (as compared to usual practice) affect reading achievement as averaged over students' progression through grades K-3?

Question 2:

How are the effects of Burst:Reading moderated by whether a student tests at, below, or far below benchmark levels at the first time she's tested during the study period?

Question 3:

What are Burst:Reading's effects on the subgroup of students for whom it is most clearly intended, namely K-3 students who (i) test below benchmark levels (calibrated to the student's grade and the time of year) on acquisition of early literacy skills at some point during the study period, (ii) at the first such occasion score "below benchmark", rather than "well below benchmark"?

Question 4:

Does a school's adoption of Burst:Reading have effects on the reading achievement of students other than those for whom it is intended, but at similar grade levels? Specifically, on the subgroup of K-3 students testing consistently at or above benchmark on early literacy skills, having been observed at grades K-3 during at least 3 of the 4 years of the study?

Question 5:

For the targeted subgroup of K-3 students falling behind in acquisition of literacy skills, are Burst:Reading's effects on reading further moderated by student stability as a matriculant at the same school? In particular, what are the program's effects on the subgroup of students who at some point test below benchmark levels on acquisition of early literacy skills, and who are also observed to attend the same school for 3 or 4 of the four years of the study?

Question 6:

Based on separate data shared by the intervention provider, we know that not all schools subscribing to the intervention actually use it. Analyzing these data, one can model use of the program in the intended manner as a function of baseline characteristics of schools and districts. Among schools that would be predicted to provide the intervention to at least some students with the frequency recommended by the intervention provider, how did Burst:Reading affect reading achievement?

Exploratory research questions:

Question 1:

A related investigation by the study team examined characteristics and program usage habits of schools subscribing to the Burst:Reading service independently of the RCT, in particular using CCD and SEA-reported school achievement variables to estimate a model of schools' propensity to subscribe to Burst. In the RCT sample, are the Burst:Reading intervention's effects on student achievement moderated by schools' propensity to subscribe to Burst?

Question 2:

A related investigation by the study team examined characteristics and program usage habits of schools subscribing

to the Burst:Reading service independently of the RCT, in particular combining CCD and SEA-reported school achievement variables with school-level program implementation data to estimate a model of the a priori likelihood that a subscribing school will implement with fidelity. In the RCT sample, are effects of Burst:Reading moderated by this implementation prognosis variable?

Question 3:

A related investigation by the study team examined characteristics and program usage habits of schools subscribing to the Burst:Reading service independently of the RCT, in particular combining CCD and SEA-reported school achievement variables with school-level program implementation data to estimate the regression of fall-to-spring gains in early literacy skills (as measured by Dibels:Next) on school characteristics. In the RCT sample, are effects of Burst:Reading moderated by this literacy skill-building prognostic score?

Question 4:

For the RCT sample we constructed a model of student-level program implementation as a combination of baseline and time-varying combinations of student and school characteristics, along with random effects. In combination with school-level baseline characteristics (including aggregated student variables), this model's fixed effects corresponding to pre-treatment variables give rise to a one-dimensional constructed variable, interpretable as the study population regression of implementation on baseline characteristics. In the RCT sample, are effects of Burst:Reading moderated by the regression of implementation on baseline characteristics of schools? (The proposed moderator is the full-study-population regression, which we would have been able to determine directly had, counter to fact, the BURST:Reading intervention been offered to the entire RCT sample. That intervention having only been offered to the treatment group, in practice we can estimate but not obtain this full-sample regression; nonetheless this research question pertains to the hypothetical full-sample regression.)

Question 5:

The estimation strategy for the study's primary research question calls for Peters-Belson style covariate adjustment, in which covariance coefficients are estimated separately and prior to the comparison of treatment and control subjects, using control group observations only. This preliminary modeling estimates the regression on baseline characteristics of students' control-condition early reading achievement, which is observed in the case of the control group but which for treatment group members is an unobserved counterfactual variable. In the RCT sample, are effects of Burst:Reading moderated by the regression of control-condition early reading achievement on baseline characteristics of students and schools? (To be clear, the proposed moderator is the full-study-population regression, which we would have been able to determine directly had, counter to fact, the BURST:Reading intervention been withheld from all schools in the RCT sample.)

Question 6:

Are the effects on K-3 student reading achievement of a school's adopting the Burst:Reading program mediated by fidelity of Burst implementation at the school -- more specifically, by the extent to which the school tends to administer the recommended "dose" of the program to children whose test scores indicate a need for it, net of expected implementation among schools with similar baseline characteristics? (We measure this extent as a school-level random effect in a mixed model predicting student dosage from baseline and time-varying variables. Note that a separate exploratory question asks whether effects of the treatment are moderated by the regression of implementation on baseline variables only.)

Question 7:

At the student level, are effect of the program on student achievement mediated by "dose" of the treatment that the student received, as compared to doses administered to similar students at baseline-similar schools? Specifically, do our estimates of program benefits to students, the differences of treatment group students' mean reading achievement scores versus means of regression predictions of these outcomes from pre-treatment student

characteristics, associate positively with sums of school and student random effects from the implementation model? (These sums being interpretable as the number of cycles the student received minus the number of cycles that a student with his baseline characteristics would be expected to receive.)

Comments:

-

Section IV-A: Study Design (Selection)

Study Design:

Randomized Trial (RT)

Comments:

-

Section IV-B: Study Design (Input)

Study Design: Input

Unit of random assignment of intervention:

School

Assignment within sites or blocks:

Yes

Define the sites or blocks:

Matched pairs and triples engineered to decrease within-stratum variability on school demographics a

Probability of assignment to treatment the same across sites or blocks:

No

Probability of assignment to treatment:

In each of 24 matched pairs of school, the probability of assignment to treatment was 1/2. There were also two matched triples of schools, one of which assigned 2 schools to treatment and 1 to control and the other assigning 1 to treatment and 2 to control. For schools placed in these triples, the probabilities of assignment to treatment were 2/3 and 1/3, respectively.

Unit outcome data measured:

Student

Intermediate clusters between unit of random assignment and unit of measurement:

No

Comments:

Supplemental exploratory analyses, not enumerated in this protocol, considered outcomes at the teacher level, specifically teacher responses to questionnaires about time use and teacher practices.

Design Classification

Based on the responses above, this study has been classified as:

RT: Multisite (Blocked) Cluster Randomized Trial

Section V: Sample Characteristics

Approximate number of students per school: 200

Approximate number of schools in the comparison condition within each block (Matched pairs and triples engineered to decrease within-stratum variability on school demographics a): 1

Approximate number of schools in the intervention condition within each block (Matched pairs and triples engineered to decrease within-stratum variability on school demographics a): 1

Number of blocks (Matched pairs and triples engineered to decrease within-stratum variability on school demographics a): 26

Yes - K-3 students testing below grade- and time of year-specific benchmarks for acquisition of early reading skills.

No Yes - The study team concentrated school recruitment efforts on high-poverty schools. No No No

Comments:

Among schools that had initially agreed to participate in the study, some withdrew participation after finding that they had been assigned to the control condition. Our policy, communicated prior to random assignment in most cases, was to drop all schools from a given district in the event that any of them dropped after initially agreeing to participate. As a side-effect of this policy, schools belonging to districts in which another school withdrew following random assignment might be considered as having been excluded from this study.

Section VI: Outcomes (Input)

Confirmatory question 1: Outcome Measure 1

Outcome domain: Student Achievement - Literacy

Minimum detectable effect size: Approx. 0.2 (SDs), with 80% power

Outcome measure: Star Early Literacy Enterprise assessment

Scale of outcome measure: Continuous

Normed or state test: Yes

Same outcome measure in treatment and comparison groups: Yes

Confirmatory question 1: Outcome Measure 2

Outcome domain: Student Achievement - Literacy

Minimum detectable effect size: Approx. 0.2 (SDs), with 80% power

Outcome measure: Star Reading Enterprise assessment

Scale of outcome measure: Continuous

Normed or state test: Yes

Same outcome measure in treatment and comparison groups: Yes

Confirmatory question 2: Outcome Measure 1

Outcome domain: Student Achievement - Literacy

Minimum detectable effect size: (not calculated)

Outcome measure: Star Early Literacy Enterprise assessment

Scale of outcome measure: Continuous

Normed or state test: Yes

Same outcome measure in treatment and comparison groups: Yes

Confirmatory question 2: Outcome Measure 2

Outcome domain: Student Achievement - Literacy

Minimum detectable effect size: (not calculated)

Outcome measure: Star Reading Enterprise assessment

Scale of outcome measure: Continuous

Normed or state test: Yes

Same outcome measure in treatment and comparison groups: Yes

Confirmatory question 3: Outcome Measure 1

Outcome domain: Student Achievement - Literacy

Minimum detectable effect size: (not calculated)

Outcome measure: Star Early Literacy Enterprise assessment

Scale of outcome measure: Continuous

Normed or state test: Yes

Same outcome measure in treatment and comparison groups: Yes

Confirmatory question 3: Outcome Measure 2

Outcome domain: Student Achievement - Literacy

Minimum detectable effect size: (not calculated)

Outcome measure: Star Reading Enterprise assessment

Scale of outcome measure: Continuous

Normed or state test: Yes

Same outcome measure in treatment and comparison groups: Yes

Confirmatory question 4: Outcome Measure 1

Outcome domain: Student Achievement - Literacy

Minimum detectable effect size: (not calculated)

Outcome measure: Star Early Literacy Enterprise assessment

Scale of outcome measure:

Normed or state test:

Same outcome measure in treatment and comparison groups:

Confirmatory question 4: Outcome Measure 2

Outcome domain: Student Achievement - Literacy

Minimum detectable effect size: (not calculated)

Outcome measure: Star Reading Enterprise assessment

Scale of outcome measure: Continuous

Normed or state test: Yes

Same outcome measure in treatment and comparison groups: Yes

Confirmatory question 5: Outcome Measure 1

Outcome domain: Student Achievement - Literacy

Minimum detectable effect size: not calculated

Outcome measure: Star Early Literacy Enterprise assessment

Scale of outcome measure: Continuous

Normed or state test: Yes

Same outcome measure in treatment and comparison groups: Yes

Confirmatory question 5: Outcome Measure 2

Outcome domain: Student Achievement - Literacy

Minimum detectable effect size: not calculated

Outcome measure: Star Reading Enterprise assessment

Scale of outcome measure: Continuous

Normed or state test: Yes

Same outcome measure in treatment and comparison groups: Yes

Confirmatory question 6: Outcome Measure 1

Outcome domain: Student Achievement - Literacy

Minimum detectable effect size: not calculated

Outcome measure: Star Early Literacy Enterprise assessment

Scale of outcome measure: Continuous

Normed or state test: Yes

Same outcome measure in treatment and comparison groups: Yes

Confirmatory question 6: Outcome Measure 2

Outcome domain: Student Achievement - Literacy

Minimum detectable effect size: not calculated

Outcome measure: Star Reading Enterprise assessment

Scale of outcome measure: Continuous

Normed or state test: Yes

Same outcome measure in treatment and comparison groups: Yes

Comments:

Star Reading is used as an outcome measure only when students have Star Reading scores but not Star Early Literacy scores within the spring testing window.

Section VII: Analysis Plan

Baseline data collected prior to start of intervention:

Yes

Description of baseline data:

Student demographics are culled from roster and student assessment databases that are populated prior to students' receipt of the intervention. As a necessary precondition for receiving Burst:Reading supports, each student must take a Dibels:Next examination, making the first Burst:Reading score recorded for each student a pre-treatment characteristic also.

Covariates you plan to include in the model:

English Language Learner Status, Free and Reduced Lunch Status, Gender, Grade, (Please see Optional Comments for Section VII.), Race, Special Education Status, Student Pretest

Covariates you plan to include in the model:

Fixed effect for state.

Analytic model:

-

Plan to handle cases with missing outcome data:

Delete cases with missing data for the outcome being analyzed

Planned multiple comparisons adjustment, confirmatory question 1 (Student Achievement - Literacy):

Yes

Number of planned comparisons to adjust, confirmatory question 1 (Student Achievement - Literacy):

Please see attached document analysis-plan.pdf

Correction for multiple comparisons, confirmatory question 1 (Student Achievement - Literacy):

other

Planned multiple comparisons adjustment, confirmatory question 2 (Student Achievement - Literacy):

Yes

Number of planned comparisons to adjust, confirmatory question 2 (Student Achievement - Literacy):

Please see attached document analysis-plan.pdf

Correction for multiple comparisons, confirmatory question 2 (Student Achievement - Literacy):

other

Planned multiple comparisons adjustment, confirmatory question 3 (Student Achievement - Literacy):

Yes

Number of planned comparisons to adjust, confirmatory question 3 (Student Achievement - Literacy):

Please see attached document analysis-plan.pdf

Correction for multiple comparisons, confirmatory question 3 (Student Achievement - Literacy):

other

Planned multiple comparisons adjustment, confirmatory question 4 (Student Achievement - Literacy):

Yes

Number of planned comparisons to adjust, confirmatory question 4 (Student Achievement - Literacy):

Please see attached document analysis-plan.pdf

Correction for multiple comparisons, confirmatory question 4 (Student Achievement - Literacy):

other

Planned multiple comparisons adjustment, confirmatory question 5 (Student Achievement - Literacy):

Yes

Number of planned comparisons to adjust, confirmatory question 5 (Student Achievement - Literacy):

Please see attached document analysis-plan.pdf

Correction for multiple comparisons, confirmatory question 5 (Student Achievement - Literacy):

other

Planned multiple comparisons adjustment, confirmatory question 6 :

Yes

Number of planned comparisons to adjust, confirmatory question 6 :

Please see attached document analysis-plan.pdf

Correction for multiple comparisons, confirmatory question 6 :

other

Comments:

1. Additional covariates for student-level model: Cubic spline expansion of student age, separately by state and following 10% Winsorization within state and grade; interaction of grade with indicator for missingness of age; fixed effect for year in which student joined study; fixed effect for whether posttest is Star Early Literacy or Star Reading.
2. Our final choice of outcome analysis procedure was made in advance of any comparison of treatment to control on the posttest measure. It was informed by baseline comparisons of treatment versus control students and schools; comparisons of the two groups on attrition/loss to follow-up; a study of uptake of the intervention within treatment group schools; and analysis of the SEL outcome among students within the control condition.

Section VIII: Additional Information

Links:

<https://github.com/benbhansen-stats/burst-analysis>

Github repository with R code used in the analysis. (As of this writing this is a non-public repository, meaning that it can only be viewed with express permission from an administrator.)

Files:

File Name: [473.pdf](#)

Description: What Works Clearinghouse registry entry

File Name: [stu-achievement-anal-plan.pdf](#)

Description: Detailed analysis plan, including model description, specification of tests to be conducted and detail for p-value calculations and accompanying multiplicity corrections. This is the version of our analysis plan document that was published in version 1 of this registry entry; it is no outdated. This PDF was generated from inst/stu-achievement-anal-plan.Rmd in the burst-analysis repository, commit b9959d5.

File Name: [power_analysis.pdf](#)

Description:

File Name: [stu-achievement-anal-plan-v2.pdf](#)

Description: Detailed analysis plan, including model description, specification of tests to be conducted and detail for p-value calculations and accompanying multiplicity corrections. This is the version of this document that was published in version 2 of this registry entry, removing an outdated passage that was inadvertently included in version 1 and in the same section adding sentences to clarify the intended plan. (It is now outdated.) New material appears in bold and excised material appears in crossout. It was generated from inst/stu-achievement-anal-plan.Rmd in the burst-analysis repository, commit 4bec0d2.

File Name: [stu-achievement-anal-plan-v3.pdf](#)

Description: Detailed analysis plan, including model description, specification of tests to be conducted and detail for p-value calculations and accompanying multiplicity corrections. This is the version of our analysis plan document that was published in version 3 of this registry entry; it is current. New material appears in bold and excised material appears in crossout. This PDF was generated from `inst/stu-achievement/analysis-plan.Rmd` in the burst-analysis repository, commit 47c03fe.

Comments:

Version 3 of the registry entry's additional materials incorporates several adjustments and corrections of oversights that became apparent shortly after outcome analysis began; a summary of these revisions follows. (Version 2 differed from version 1 only in the addition of the updated document `stu-achievement-anal-plan-v2.pdf`.)

Detail of changes for second revision of analysis plan (Nov 14, 2018)

1. In year 2 of the study, some schools administered Star Reading (SR) rather than Star Early Learning (SEL), our intended student achievement outcome, to a subset of pupils. As we prepared our original analysis plan we had calculated that less than 20 student-year records were affected; so we opted simply to exclude these observations from our analysis. We recently found that calculation to have been in error: another 430 records that we had thought to contain SEL scores turn out to in fact contain SR scores. In order to retain these records in our analysis, this revision adds to the covariance adjustment a fixed effect for the distinction between the two tests, at the same time adding back in to the analysis sample the small number of records that we had always known to contain SR rather than SEL scores.
 2. A miscommunication within the analysis team had led to the omission of a subgroup from the list of subgroups that would be examined for the presence of a treatment effect. Prior analyses conducted with a different data set had created an implementation fidelity prognostic index, composed of baseline school and district characteristics. While our analysis plan had noted our intention to examine in `_exploratory_` analysis whether this index associates with school-level treatment effects, it omitted a `_confirmatory_` test of whether the program was efficacious within schools with moderate or high prognoses of implementation fidelity. The PI had always intended for such a test to be conducted. The present revision addresses this omission by adding a corresponding confirmatory hypothesis test.
 3. The student-level covariate adjustment we had initially planned adjusted for student age by creating age "bins," separately by state and grade, and including dummy regressors for each bin. Upon simulating the permutation distribution of our main effect estimate, we found its finite-sample bias (under the null of strictly no effect) to be larger than expected; this bias traces in part to the complexity of the age adjustment, which consumed relatively many degrees of freedom (200). The revised data plan replaces that binning scheme with another combining Winsorization of the age variable, to mimic the leverage-limiting aspect of binning, with a spline expansion of the age variable, to address potential nonlinearity in the partial regression of student achievement on age. After interaction with State, this age covariate consumes 40 degrees of freedom.
 4. We had planned for our outcome analyses to include a fixed effect for district. Once outcome analysis began we realized that this would lead to the exclusion of one district with only a single participating school. That district happens to be one of two districts from a single state, the only state with multiple participating districts. It also happens that this state contributes only 3 participating schools across its two districts. So we decided to make the minor adjustment of including a fixed effects for state rather than fixed effects for district.
-