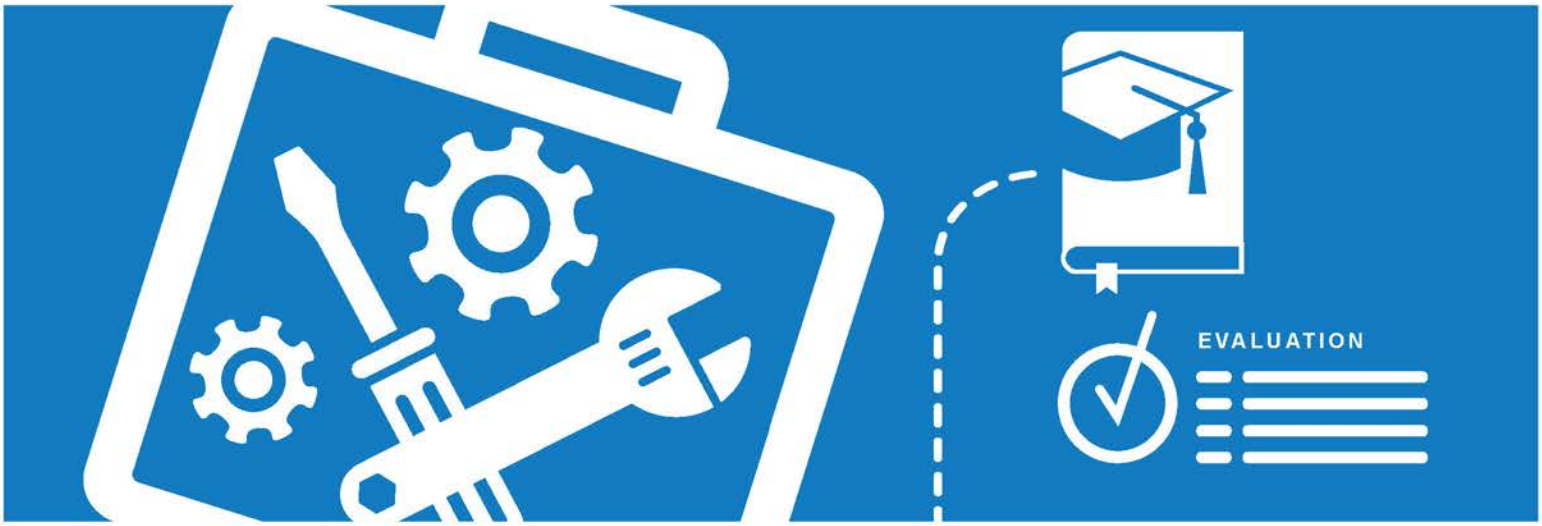


Example Evaluation Plan for a Quasi-Experimental Design

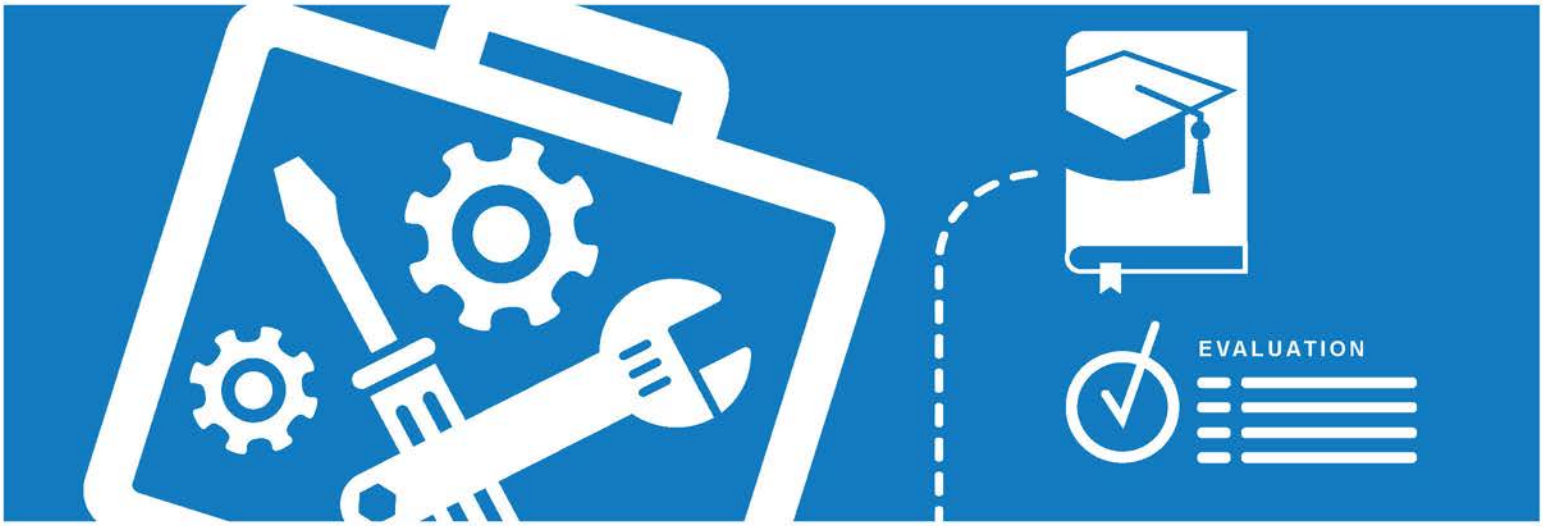
The Evaluation Plan Template identifies the key components of an evaluation plan and provides guidance about the information typically included in each section of a plan for evaluating both the effectiveness and implementation of an intervention. Evaluators can use this tool to help develop their plan for a rigorous evaluation, with a focus on meeting *What Works Clearinghouse*^{TM} evidence standards. The template can be used in combination with the Contrast Tool, a tool for documenting each impact that the evaluation will estimate to test program effectiveness.

This document provides an example of a detailed evaluation plan for evaluating the effectiveness of an intervention. Developed using the Evaluation Plan Template, the plan is for a quasi-experimental design (QED). The example illustrates the information that an evaluator should include in each section of an evaluation plan, as well as provides tips and highlights key information to consider when writing an evaluation plan for a QED. Accompanying this example plan is the Example Contrast Tool for a QED, which lists each impact that the example evaluation will estimate to test program effectiveness. The example Evaluation Plan and the example Contrast Tool can be reviewed side-by-side.

Prepared by:

Abt Associates
Anne Wolf
Barbara Goodson
Katie Gan
Cristofer Price
Beth Boulay

August 2016



The Institute of Education Sciences (IES) has made this tool publicly available as a courtesy to evaluators. However, the content of this tool does not necessarily represent IES's views about best practices in scientific investigation.

This tool was developed under U.S. Department of Education Institute of Education Sciences (IES) contracts ED-IES-10-C-0064 with Abt Associates and ED-ODS-12-A-0019/0031with AEM Corporation and its subcontractor Abt Associates. These contracts provided evaluation technical assistance for evaluations of interventions funded by the Investing in Innovation and First in the World programs. Tools, webinars, and other materials were developed to help grantees engage in good scientific practice and produce evaluations that meet *What Works Clearinghouse*^{TM} evidence standards.

CONTENTS

Grantee: Southeast Technical University (SETU)

1. Evaluator Information

1.1 Contact Information

Evaluator: Research123
100 Babel Fish Blvd.
City, State 11677

Lead Researcher: Baruch Spinoza, Ed.D.
bspinoza@research123.email
669-221-6251

1.2 Independence

Research123 is not affiliated with State Community College, played no role in the development or implementation of the ModMath intervention, and will independently conduct all key aspects of the evaluation, including collection of any data used in the impact analysis, execution of the impact analyses, and reporting of study findings. The findings reported will not be subject to the approval of the project director or any staff developing or implementing the intervention

1.3 Confidentiality Protections

Research123 has secured IRB approval for this study from Ethical and Independent Review Services. The research team will conduct this research in a manner consistent with the provisions of the Family Educational Rights and Privacy Act (FERPA).

To protect confidential data, Dr. Spinoza will ensure that data are accessible only to personnel and consultants associated with this study or to authorities legally authorized for access. Dr. Spinoza will create and maintain a list of unique, random id codes for the student data, allowing team members to review student performance without names or other identifiers appended. All transfers of confidential data between the company and college occur through our secure website, which uses VeriSign security. Any hard copy data will be stored in locked file cabinets. Hard copies and data files containing individually identifiable student or teacher data will be destroyed one year after the study has been completed. All State Community College administrative data provided for this research will remain the property of the participating community college, even while stored in a database hosted by Research123.

2. Summary of Interventions

Improving student success in developmental education classes is one of the key challenges that community colleges face as they strive toward increasing student graduation rates. This issue is particularly acute in mathematics.

To address this challenge, the developmental and credit mathematics faculties at State Community College (SCC) have collaborated on *ModMath*, a redesign of the community college's traditional three-course developmental math sequence, which includes pre-algebra, elementary algebra, and intermediate algebra. At State Community College, students are required to complete at least one college-level, credit bearing mathematics class in order to graduate. Students who enter the college without the necessary preparation for college-level math must successfully complete intermediate algebra as a prerequisite. The goal of *ModMath* is to improve students' successful completion of developmental math courses in order to prepare them to be able to meet core course requirements for college-level mathematics. For many students, passing required college mathematics is an obstacle to persistence in community college and, ultimately, degree attainment. Beyond enabling students to complete their college mathematics requirements, *ModMath* is intended to improve overall student academic achievement, progress in developmental education, credit accumulation/persistence, and successful exit from community college.

ModMath is a redesigned approach to the three-semester developmental mathematics sequence. Students' learning is structured by the *ModMath* online learning system, in coordination with an in-classroom instructor who guides and supports students through their work. Students have access to the online *ModMath* system both in and outside of the classroom. As in traditional classes, students in *ModMath* attend regularly scheduled classes with an instructor (albeit in a computer lab), complete homework, and take quizzes and tests. However, the *ModMath* software personalizes the content and pace of the instruction they receive so that students stay engaged in their coursework and find success. Additionally, because *ModMath* spans the full sequence of developmental math at State Community College, students benefit from a consistency and continuity of instruction as they advance.

Students enter SCC with varied levels of math preparation, as determined by their score on the Accuplacer assessment, administered prior to the start of their first semester. Some students only need to take intermediate algebra before they are ready for college-level mathematics; however, others need all three developmental math classes (pre-algebra, elementary algebra, and intermediate algebra). The *ModMath* intervention is designed for students who need to take the full three-semester developmental math sequence – specifically, it will be offered to students who need to begin with pre-algebra. Currently, students who enter the college ready to take elementary or intermediate algebra will be enrolled in the business-as-usual developmental math classes, without an option to take *ModMath*. All developmental math students will be offered *ModMath* elementary algebra by fall 2018, and *ModMath* intermediate algebra by fall 2019.

3. Impact/Effectiveness Evaluation

The impact study will used a quasi-experimental design to assess the impact of *ModMath* on a core set of student outcomes during the four years of the grant (2016-2019, covering SY 2016-17 through SY 2018-19). The evaluation will focus on students who enter State Community College needing instruction in pre-algebra. The study will compare students who choose to enroll in the *ModMath* sequence of developmental math courses with a matched group of students who choose to enroll in the traditional developmental math sequence offered by SCC, examining the successful completion of the developmental math sequence, math credit accumulation, and two-year degree attainment/transfer to a four-year university.

3.1 Research Questions

The general question for this evaluation is, "Does *ModMath* increase community college success for developmental math students, who enter college needing instruction in pre-algebra, compared to the usual developmental math sequence?" Within this larger question, we've specified the following more specific research questions for this study. Research questions are also listed in the accompanying contrast tool.

TIP!

In your evaluation plan…

- ☑ Outline specific, narrowly defined research questions that will be addressed by the study
- ☑ Have a research question for each specific test of the intervention effect.

The first three research questions focus on the ***successful completion*** of each developmental math course:

- What is the effect of ModMath on students' *successful completion of pre-algebra* at the end of their first semester of college relative to business-as-usual developmental math?
- What is the effect of ModMath on students' *successful completion of elementary algebra* at the end of their second semester of college relative to business-as-usual developmental math?
- What is the effect of ModMath on students' *successful completion of intermediate algebra* at the end of their third semester of college relative to business-as-usual developmental math?

The study will address two research questions about ***credit accumulation in college-level math***, which is necessary to meet core course requirements at State Community College.

- What is the effect of ModMath on students' *credit-accumulation in college-level math* by the end of their fourth semester relative to business-as-usual developmental math?
- What is the effect of ModMath on students' *credit-accumulation in college-level math* by the end of their fifth semester relative to business-as-usual developmental math?

Finally, the study will address one research question about whether *ModMath* enables students to ***successfully exit community college.***

- What is the effect of ModMath on students' *attainment of an Associate's degree or transfer to a Bachelor's degree-granting institution* within five semesters relative to business-as-usual developmental math?

3.2 Comparison Condition

The comparison group will be formed from students enrolling in the traditional developmental math sequence (without *ModMath*). They will receive the usual developmental pre-algebra, and then be offered the usual developmental elementary and intermediate algebra courses. All instructors use the same curriculum and texts for each of these courses; however, individual instructors may vary in their pace or use of supplemental supports and materials to fit their preferences or students' needs. These classes generally include an instructor presenting material in front of the entire class, with some group work or individual practice, during class time. While these classes take place in a regular classroom (i.e., not a computer lab), instructors may refer students to computer resources to supplement their teaching or as part of homework assessments.

3.3 Study Sample and How Intervention and Comparison Groups are Selected/Assigned

The evaluation will take place at State Community College, which has an entering class of approximately 8,000 new students each fall and approximately 2,000 new students each spring. Approximately 20% of new students require remedial math education prior to enrolling in college-level math. The study will focus on the students who need the entire three-course developmental math sequence, who represent approximately 60% of students who enroll needing developmental math (i.e., approximately 960 students each fall and 240 each spring). Student placement into pre-algebra is based on the student's score on the Accuplacer assessment, administered prior to the start of the first semester. Information about study samples is described in this section and also shown in the accompanying contrast tool, on the "samples" tab.

Selection of Instructors

Initially, *ModMath* will be taught by the 7 developmental education faculty members that helped create the new program and will roll out to an additional 12 instructors over time, as shown in Table 1. The first 7 instructors will begin implementing *ModMath* in fall 2016. The other developmental math instructors at SCC will continue teaching the traditional mathematics course sequence through summer 2018. In summer 2018, all developmental math instructors will be trained to teach the *ModMath* modules for pre-algebra and elementary algebra. In fall 2018, all pre-algebra and elementary algebra classes will be taught using *ModMath*. In summer 2019, all instructors will be trained to teach the *ModMath* modules for intermediate algebra. In fall 2019, all developmental math classes will be taught using *ModMath*.

Table 1. Schedule for implementation of *ModMath*.

	Fall 2016 – Summer 2018 *(Fall 16, Spring 17, Summer 17, Fall 17, Spring18, Summer 18)*	Fall 2018 – Summer 2019 *(Fall 18, Spring 19, Summer 19)*	Fall 2019
ModMath	7 instructors • Pre-algebra • Elementary algebra • Intermediate algebra	All 19 instructors • Pre-algebra • Elementary algebra 7 instructors • Intermediate algebra	All 19 instructors • Pre-algebra • Elementary algebra • Intermediate algebra
Traditional Developmental Education	12 instructors • Pre-algebra • Elementary algebra • Intermediate algebra	12 instructors • Intermediate algebra	None

The schedule for offering *ModMath* course sections in any given semester will be determined by computer lab and instructor availability. Instructors teach an average of two sections of pre-algebra each semester, with approximately 25 students per section.

Selection of Students

Student Eligibility. The study will be conducted with three cohorts of students enrolling in pre-algebra for the first time in fall 2016, spring 2017, or fall 2017 at State Community College (SCC). Table 2 summarizes the eligibility and exclusion criteria for all students in the evaluation sample, both treatment students and comparison students. Only students enrolling in pre-algebra when they enter State Community College will be included in the evaluation. Students who start by taking elementary or intermediate algebra in their first semester will not be included in the evaluation. Students in the evaluation sample may be enrolled in State Community College full-time or part-time, but they must be enrolled in at least one other course besides pre-algebra.

TIP!

- ☑ Clearly define eligibility criteria that will be applied in both the treatment and comparison groups.
- ☑ Describe how students obtain access to the intervention.
- ☑ Try to avoid selecting a comparison group with students who have been offered the intervention but elected not to participate. In that situation, differences in outcomes could be the result of pre-existing differences between students who do and do not choose to participate in the intervention, rather than the result of the intervention itself. Note that this situation will *not* jeopardize a QED's potential to meet WWC standards with reservations.

Table 2. Summary of Eligibility and Exclusion Criteria for Students in the Evaluation Sample

Eligibility Criteria	• Placed into pre-algebra, based on Accuplacer score at time of college entry • Enroll in pre-algebra in first semester at SCC • Enrolled in at least one other course at SCC in addition to pre-algebra
Exclusion Criteria	• Enrollment in developmental math in the opposite condition in any subsequent semester (will be excluded from that semester forward). This includes treatment students who enroll in business-as-usual developmental math (non-*ModMath*) and comparison students who enroll in *ModMath*.

Students will register for pre-algebra following the normal course selection process. They will have an opportunity to select either a *ModMath* or a business-as-usual pre-algebra class. Although students' scheduling needs are most likely to drive their choice, their selection of *ModMath* or business-as-usual pre-algebra also may be based on their preference for the specific instructional format or for a particular instructor.

Once students enroll in *ModMath* or business-as-usual pre-algebra, the college will attempt to keep students enrolled in the same condition for subsequent developmental math classes. Enrollment in *ModMath* pre-algebra will be a prerequisite for enrollment in *ModMath* elementary algebra, and both *ModMath* pre-algebra and elementary algebra will prerequisites for *ModMath* intermediate algebra. In other words, students who select a business-as-usual pre-algebra class will only be able to enroll in the traditional developmental math courses in subsequent semesters. However, students who select a *ModMath* pre-algebra class may potentially enroll in a traditional developmental math course in a subsequent semester. They will be encouraged to continue with *ModMath* for elementary and intermediate algebra, by having *ModMath* sections offered at the same times as traditional elementary and intermediate algebra sections, but they will not be precluded from enrolling in traditional elementary algebra and intermediate algebra classes.

TIP!

☑ Make sure students selected for the comparison group have little opportunity to participate the intervention.

☑ Avoid a design where comparison group access to the intervention may result in having no untreated comparison group. If comparison group students in a QED have access to the intervention (e.g., in other courses and/or in subsequent semesters), consider excluding them from the analytic sample. If many of the students in the comparison group have are exposed to the intervention, the study will not yield a test of the intervention's effects.

In this example, State Community College has implemented course enrollment prerequisites to preclude comparison students from enrolling in intervention classes in subsequent semesters. In addition, students who do enroll in classes in the opposite condition will be excluded from the evaluation sample.

Any student who, in a subsequent semester, enrolls in a developmental math class in the opposite condition will be excluded from the sample from that semester forward. In other words, if a student who enrolled in *ModMath* pre-algebra later enrolls in a traditional elementary algebra class or a traditional intermediate algebra class, the student will be excluded from the sample for analyses of outcomes measured in that semester and all future semesters. Similarly, students who enroll in business-as-usual pre-algebra will be excluded from the sample if they were to enroll in *ModMath* in a subsequent semester, although they should be precluded from doing so.

Students who stop taking developmental math before completing the sequence will remain in the sample, regardless of condition. For example, students who take pre-algebra but do not take developmental math in their second semester will remain in the sample. Similarly, students who take pre-algebra in their first semester and elementary algebra in their second semester, but do not take intermediate algebra in their third semester will still be kept in the evaluation sample. Finally, students who take a semester off of developmental math and then return to it (e.g., pre-algebra in their first semester and elementary algebra in their third semester) will also remain in the sample.

Table 3 provides a summary of the students that will be included in the treatment and comparison conditions.

TIP!

☑ Clearly define how students who do not fully participate in the intervention will be handled in the evaluation sample.
- Include all students provided access to the intervention, regardless of participation, to address research questions about the effect of *offering* students the intervention (analogous to an intent-to-treat analysis in an RCT).
- Includes only students that complete the full intervention to address research questions about the effect of *participation* in the intervention (analogous to a treatment-on-the-treated analysis in an RCT).

In this example, students who do not complete the three-semester developmental education sequence (in either condition) are included in the sample.

Table 3. Description of Students in the Treatment and Comparison Conditions

ModMath Treatment Group	• Enrolled in *ModMath* in 1st semester AND • In either of the following in 2nd and 3rd semesters: o Enrolled in *ModMath* developmental math o Not enrolled in any developmental math course
Business-as-Usual (BAU) Comparison Group	• Enrolled in BAU developmental math in 1st semester AND • In either of the following in 2nd and 3rd semesters: o Enrolled in BAU developmental math o Not enrolled in any developmental math course

Selection of Treatment Group Students. All full-time and part-time students enrolled in a *ModMath* pre-algebra course in their first semester at State Community College (and enrolled in at least one other course besides pre-algebra) will be included in the treatment group, as long as they do not subsequently enroll in a traditional developmental math class.

Selection of a Matched Comparison Group of Students. From the pool of full-time and part-time students enrolled in traditional pre-algebra in their first semester at State Community College (and enrolled in at least one other course besides pre-algebra), the evaluation team will select a matched comparison sample of students similar to those enrolled in *ModMath*, using propensity score matching. Students in traditional pre-algebra will be matched to students in *ModMath*, within cohort (fall 2016, spring 2017, or fall 2017).

Specifically, the following student-level data, obtained from the college administrative records, will be used in the matching process for each cohort.

- *Prior achievement*. SCC tests students on the Accuplacer as part of the enrollment process. Accuplacer results are used to place students into developmental courses (or indicate students are ready for college-level courses). Typically, the Accuplacer is administered during enrollment, which occurs just before the start of each semester.
- *First-generation to college status.* The registrar's records (from student applications) indicate whether either of a student's parents ever enrolled in postsecondary education.
- *Age*. The registrar's records indicate the age at which students first enroll in college. Students will be matched based on whether they are younger than 20 when they first enroll in college, 20 – 25, 26 – 30 or older than 30.
- *Full-time/part-time status in their first semester*. Students choose to enroll either full-time (at least 12 credits, pay tuition fee) or part-time (fewer than 12 credits, pay per credit). While choices about full or part-time enrollment in subsequent semesters may be influenced by experiences at SCC, students make decisions about their first semester prior to being enrolled in SCC.
- *Gender.* Data on whether the student is male or female will be included.
- *Race/Ethnicity.* Students' race/ethnicity is identified as Hispanic, Black, White, or Other
- *First-time to College Status.* Records indicate whether or not students have previously been enrolled in postsecondary education.

Using logistic regression, we will estimate propensity scores for all students enrolled in pre-algebra (both *ModMath* and business-as-usual) based on this set of matching variables. We will compare the distribution of propensity scores in the treatment group and the comparison group to see if the two groups span a similar range (i.e., have similar propensity to enroll in *ModMath*). After identifying the area of common support, we will divide the propensity scores into quintiles to create five strata. Comparison group students will be matched to treatment group students in the same stratum. Given that the number of students in the comparison group is larger than the number of students in the treatment group, this approach to matching will maximize the total sample size.

Expected Sample Sizes

Table 4 presents expected student sample sizes overall, by cohort, and by condition. These sample size estimates are based on the assumption that, in a given semester, each instructor teaches an average of two pre-algebra sections with approximately 25 students. With 7 *ModMath* instructors, there are 14 *ModMath* pre-algebra class sections offered each fall, serving approximately 350 students. With 12 instructors teaching traditional pre-algebra, there are 24 sections each fall serving approximately 600 students. For the spring semester, there are fewer entering students and a correspondingly lower number of students enrolling in pre-algebra. Therefore only 6 sections of *ModMath* pre-algebra, serving 140 students, and 6 sections of traditional pre-algebra, serving 140 students, are offered in the spring. In total, across all there cohorts, the expected student sample size is 848 students in *ModMath* and 1,352 students in the comparison group.

Actual analytic samples sizes will be smaller, given that some students will drop out of State Community College during the course of the evaluation. Also, students who enroll in *ModMath* pre-algebra, but subsequently enroll in a traditional developmental math class, will be excluded from the evaluation, which will further reduce the size of the treatment group sample.

Table 4. Expected Number of Class Sections and Students in the Evaluation

	Full Sample		Treatment Group		Comparison Group	
	Class Sections	Students	Class Sections	Students	Class Sections	Students
Cohort A (Fall 2016)	38	960	14	354	24	606
Cohort B (Spring 2017)	12	240	6	140	6	140
Cohort C (Fall 2017)	38	960	14	354	24	606
Total	88	2,160	34	848	54	1352

Progression of Students and Cohorts Across Multiple Years

Table 5 illustrates the semester-to-semester progression of students over time for the three cohorts that will be included in the evaluation of *ModMath*. Although students are expected to progress through the entire three-semester developmental math sequence by the end of their third semester at State Community College, the evaluation will measure effects of the intervention on outcomes up through students' 5th semester of enrollment or until they graduate or transfer to a four-year institution.

The three cohorts will be combined for analyses, and the effects of *ModMath* will be examined ***at the end of each semester from students' first semester through their fifth semester***. The timing of outcome measurement will vary by cohort, as shown in Table 3. Spring 2019 will be the last time student outcomes are measured. This schedule means that the study will be able to follow cohorts A and B through their fifth semester and cohort C through their fourth semester. The samples for the analyses of *ModMath* effects are as follows:

TIP!

- ☑ If an intervention spans multiple semesters or multiple years, and/or if the study includes multiple cohorts, clearly describe the progression of students (and cohorts) over time.
- ☑ Indicate when students will receive the intervention, and when data on outcomes, baseline measures, and covariates will be collected.
- ☑ Include a chart, table, or other graphic (like Table 5) to clearly show how students (and cohorts) progress from year to year (or semester to semester) relative to the timing of the intervention and collection of outcome data.
- ☑ State when impacts will be assessed in relation to (a) the amount of intervention exposure/length of follow-up; (b) student "grade," and/or (c) how long the intervention has been in place.

- End of first semester is fall 2016 for Cohort A, spring 2017 for Cohort B, and fall 2017 for Cohort C (cells A1, B1, and C1).
- End of second semester is spring 2017 for Cohort A, fall 2017 for Cohort B, and spring 2018 for Cohort C (cells A2, B2, and C2).

- End of third semester is fall 2017 for Cohort A, spring 2018 for Cohort B, and fall 2018 for Cohort C (cells A3, B3, and C3).
- End of fourth semester is spring 2018 for Cohort A, fall 2018 for Cohort B, and spring 2019 for Cohort C (cells A4, B4, and C4).
- End of fifth semester is fall 2018 for Cohort A and spring 2019 for Cohort B (cells A5 and B5). Cohort C will not reach the end of the 5th semester before the end of the evaluation.

TIP!

☑ Clearly state whether the cohorts will be combined or analyzed separately.

Combining cohorts will increase the sample size and improve statistical power for detecting intervention effects. However, if there are differences in the intervention for different cohorts, you may want to analyze cohorts separately. But be aware, the WWC may adjust for multiple comparisons if cohorts are analyzed separately. For more information, see WWC Procedures and Standards Handbook (version 3.0), p. 25-26 and Appendix G.

Table 5: Progression of Students over Time, by Cohort

Academic Year	Semester	Cohort A	Cohort 2	Cohort 3
2016-17	Fall 2016 (Semester 1)	A1* 1st semester *1 semester exposure*		
	Spring 2017 (Semester 2)	A2* 2nd semester *2 semesters exposure*	B1* 1st semester *1 semester exposure*	
2017-18	Fall 2017 (Semester 3)	A3* 3rd semester *3 semesters exposure*	B2* 2nd semester *2 semesters exposure*	C1* 1st semester *1 semester exposure*
	Spring 2018 (Semester 4)	A4* 4th semester *1 semester post-intervention*	B3* 3rd semester *3 semesters exposure*	C2* 2nd semester *2 semesters exposure*
2018-19	Fall 2018 (Semester 5)	A5* 5th semester *2 semesters post-intervention*	B4* 4th semester *1 semester post-intervention*	C3* 3rd semester *3 semesters exposure*
	Spring 2019 (Semester 6)		B5* 5th semester *2 semesters post-intervention*	C4* 4th semester *1 semester post-intervention*

Note: Cell labels indicate the cohort (A, B, or C) and students' semester in State Community College (1-5). Shading indicates the timing of the intervention for each cohort. Asterisks (*) indicate the timing of outcome data collection for each cohort.

3.4 Key Measures and Plan for Obtaining Data

In this section, we describe data collection and the variables that will be examined in the analysis.

Data Collection

Data will be collected from two sources: administrative data from State Community College and the National Student Clearinghouse.

Administrative data. Most measures will come from State Community College administrative data. SCC will provide all outcome data for the study sample to Reseach123 in summer/fall 2019, once the data are complete. SCC will send data to Research123 using a secure file transfer protocol and data will be de-identified, but include study-specific student identifier so that individuals may be tracked between semesters. The dataset will include all students who were enrolled in developmental math classes in fall 2016 – fall 2018, regardless of whether they enrolled in *ModMath* or traditional courses. These administrative data will include data held by the registrar from the students' applications/enrollment (e.g., first-generation-to-college status, demographics, Accuplacer score) as well as data tracking students' progress at SCC each semester (e.g., full-time/part-time enrollment status, course enrollment, course grades, credit accumulation, and degree attainment).

National Student Clearinghouse data. State Community College has a data-sharing agreement with the National Student Clearinghouse. These data will be used to supplement State Community College registrar data on student enrollment, allowing the evaluation team to obtain enrollment data for students who have transferred to another institution of higher education. This will allow us to differentiate between students who drop out of college from those who transfer to another institution, so these data can be used to measure successful exit from SCC. Data from the National Student Clearinghouse will be provided to Research123 by SCC using the same data security and privacy protections described above.

Analytic Measures

Below we describe the outcome measures, baseline measures, and other independent variables that will be used in analyses of the impacts of *ModMath*. Information about the planned analytic measures is also provided in the accompanying contrast tool, on the "outcomes" and "baseline measures" tabs.

Outcome Measures. The evaluation will examine the effects of *ModMath* on five outcomes – successful completion of each developmental math course; credit accumulation in college-level math; and attainment of an Associate's degree or transfer to Bachelor's degree-granting institution. Table 6 summarizes the domains, outcomes, measurement timing, and the baseline measures to be used to assess the equivalence of the treatment and comparison group in each analytic sample.

TIP!

☑ Clearly define outcome measures for students who leave a study institution before the outcome data are collected, especially for studies of postsecondary interventions.

In this example, the evaluator distinguishes between when outcome data will be treated as missing and when the outcome will be defined based on the last data point available (i.e., the last semester completed).

☑ Treat data as missing for outcome measures that would have a different value for students that leave a study institution if data are available from sources outside the study institution (or attempt to obtain the data from other sources).

In this example, a student who leaves SCC before his/her fourth semester may go on to earn college math credits at another institution – without such data, his/her outcome for "total credits in college math" would be missing.

The following three academic achievement outcomes will be constructed:

- ***Successful completion of each developmental math course/module set.*** Three outcomes will be constructed to measure whether students pass each developmental math course – pre-algebra by the end of the 1st semester; elementary algebra by the end of the 2nd semester; and intermediate algebra by the end of the 3rd semester (or the corresponding set of *ModMath* modules). Students who earn a C or better will be classified as passing the course. Those who receive a grade of D or F, who withdraw from the course, or who never enroll in the course (after completing an earlier course) will be classified as not having passed the course. Students who withdraw or transfer from State Community College before the end of their third semester will have missing data for these outcomes.

One measure of credit accumulation and persistence will be constructed (for two time points):

- ***Total credit accumulation in college-level math.*** The total number of credits earned in college-level math will be constructed for each of the two semesters after the developmental math sequence is expected to be complete – i.e.., in students' 4th semester and 5th semester. Students who leave State Community College will have data for this outcome through their last semester at SCC and will have missing data for this outcome after their withdrawal or transfer.

Finally, one measure of attainment will be constructed:

- ***Successful exit from community college.*** Students will be classified as successfully exiting community college if they have attained an associate's degree (as determined by SCC administrative data) or are enrolled in a bachelor's-degree granting institution at the end of the 5th semester after their enrollment in State Community College (as determined by National Student Clearinghouse data). Students will be classified as not successfully exiting community college if they are still enrolled at SCC and have not completed sufficient credits to attain an associate's degree or if they are no longer enrolled in SCC and are not enrolled in a bachelor's-degree granting institution (as determined by NSC data).

Table 6. Outcome domains, measures, timing of measurement, and associated baseline measures

Domain	Outcome Measure	Timing of Measurement	Baseline Measures
Academic Achievement	Successful completion of (i.e., passing grade in) developmental Pre-Algebra (*binary*)	End of 1st semester (cohorts A, B, C)	Prior to 1st Semester: Accuplacer Elementary Algebra score First-generation-to-college status
	Successful completion of (i.e., passing grade in) developmental Elementary Algebra (*binary*)	End of 2nd semester (cohorts A, B, C)	
	Successful completion of (i.e., passing grade in) developmental Intermediate Algebra (*binary*)	End of 3rd semester (cohorts A, B, C)	
Credit Accumulation/ Persistence	Total credit accumulation in college-level math courses (*continuous*)	End of 4th semester (cohorts A, B, C) End of 5th semester (cohorts A, B)	
Attainment	Successful exit from community college (*binary*), either: - Attainment of AA/AS degree - Transfer to BA/BS-granting institution	End of 5th semester (cohorts A, B)	

KEEP IN MIND…
Although you may choose to define your outcome domains differently than the WWC does, it's important to be aware of how your outcomes will be classified by the WWC, because the WWC will apply multiple comparisons adjustments for multiple impacts estimated in the same domain. For more information, see the relevant topic area review protocol: WWC Topic Area Review Protocols.

Baseline Measures

For each outcome, baseline equivalence will be assessed for one measure of academic achievement and one measure of student socioeconomic status. The baseline measure of academic achievement will be student scores on the Accuplacer Elementary Algebra test, administered to all students at the time of enrollment in SCC to determine their initial math placement level. Accuplacer scores range from 20-120.

The baseline measure of student socioeconomic status will be status as a first-generation-to-college student. Students will be classified as having a parent who ever enrolled in postsecondary education or not having a parent who enrolled in postsecondary education.

Other Covariates

The other covariates used in the analysis will be those described earlier as used for matching. In addition to Accuplacer Elementary Algebra score and a first-generation indicator, these include: age, gender, race/ethnicity, first-time-to-college status, and full-time/part-time status in the first semester. To reflect the fact that matching and sample selection will take place for each of three cohorts, we will also include terms for cohort and matching strata.

3.5 Statistical Analysis of Impacts

The impact analysis will examine outcomes for students who enrolled in a *ModMath* developmental pre-algebra class compared to those who enrolled in a traditional developmental pre-algebra class. Hierarchical linear modeling will be used to estimate the impact of the *ModMath* on student outcomes, adjusting for cohort, matching strata, and baseline student characteristics. See the accompanying contrast tool for information about each test of the intervention's effect that will be estimated; they are shown on the "contrasts" tab.

TIP!

- ☑ Account for cluster assignment in the analytic model. Outcomes for students grouped together in the same cluster (e.g., class, school) are likely to be correlated. If models do not adjust for clustering, standard errors may be underestimated.
- ☑ Use a method such as multilevel modeling (HLM), Huber-White Sandwich estimator, or GEE (e.g., Stata's "cluster" option) to adjust standard errors for clustering. If you do not, the WWC will apply a post-hoc correction to the standard error of your impact estimate, which will likely be more conservative (i.e., resulting in a larger p-value) than the adjustment you apply based on your sample data.

Impact Analysis Model

The hierarchical linear model specified below will be used to estimate the impact of *ModMath* and address the research questions.

Level-1 (student-level):

$$Y_{ij} = \beta_{0j} + \sum_{d=1}^{D} \beta_{dj} MatchStratum_{dij} + \sum_{k=d+1}^{K} \beta_{kj} X_{kij} + \varepsilon_{ij}$$

Level-2 (course-section level):

$$\beta_{0j} = \gamma_{00} + \gamma_{01} MODMATH_j + \sum_{m=2}^{M} \gamma_{0m} Cohort_{mj} + \mu_j$$

TIP!

- ☑ Include terms in the analysis model to represent matching strata and cohorts. Doing so will:
  - Improve the precision of the impact estimate.
  - Adjust for unequal assignment probabilities.
  - Adjust for unequal rates of attrition across blocks.

In this example, there are five matching strata, based on propensity score quintiles within each of three cohorts.

Where:

Y_{ij} = outcome score for student *i* in course section *j*
β_{0j} = outcome score in comparison course section *j*
β_{dj} = vector of coefficients for *d* matching strata
$MatchStratum_{dij}$ = indicator for matching stratum *d* for student *i* in course section *j*
β_{kj} = vector of coefficients for individual characteristics
X_{kij} = set of variables for individual characteristics, including age, full-time/part-time status in the first semester, and each of the measures used to establish baseline equivalence (Accuplacer Elementary Algebra score and a first-generation indicator)
ε_{ij} = error term for student *i* in course section *j*
γ_{00} = average outcome score in comparison courses

γ_{01} = difference in outcome between treatment sections and comparison class sections
$ModMath_j$ = treatment status for course section *j*
γ_{0m} = vector of parameters for the effects of cohort
$Cohort_j$ = set of dummy variables indicating cohort A, B, or C for course section *j*
μ_j = error term for course section *j*

KEEP IN MIND…
Evaluators may choose to analyze impacts on binary outcomes using multilevel logistic regression or multilevel linear regression. Either approach will yield unbiased estimates of the intervention impact. Logistic regression models are designed for binary data. However, linear models may be simpler to estimate and interpret, and they yield standard error estimates that are approximately correct even when the underlying data generating process is nonlinear (Judkins & Porter, 2015).

The coefficient, γ_{01}, provides a covariate-adjusted estimate of *ModMath*. In other words, it represents the average difference in outcomes between students in *ModMath* and non-*ModMath* courses after controlling for the covariates in the model. The hypothesis test for γ_{01} will determine whether or not the intervention has a statistically significant impact on the given outcome. A standardized effect size will be calculated by dividing the impact estimate (γ_{01}) by the pooled standard deviation derived from the unadjusted sample standard deviations for the outcome in the intervention and comparison groups.

We will estimate this HLM model for all student outcomes – those on a binary scale as well as those on a continuous scale. For both binary and continuous outcomes, the linear model yields unbiased estimates of the intervention impact.

The contribution of covariates for student characteristics and baseline performance will be assessed for inclusion in the model. If the coefficient term for a covariate has a p-value less than p = 0.20, we will consider that covariate to be contributing to the precision of the impact estimate, and will include it in the model. Research has demonstrated that this approach is effective for identifying covariates to retain and those to drop in order to minimize the standard error on the impact estimate (Budtz-Jorgensen et al, 2001; Maldonado & Greenland, 1993; Price et al, 2007). The matching strata and cohort variables will be included in the model regardless of coefficient significance, in order to account for the selection of class sections within cohorts and the selection of students within matching strata.

TIP!

- ☑ Develop criteria for which covariates to include/exclude from the analysis model. The WWC rating of the study will not be affected by the approach used to include/exclude covariates – as long as you are careful not to include any covariates that could have been affected by the intervention.
- ☑ Use literature in the field to guide the selection of covariates. There may be covariates that should be included based on theory or prior empirical research, leading you to include certain covariates regardless of p-value or any other criteria.
- ☑ Consider backward selection or another empirically-based approach if you do not have a substantive basis for selecting covariates.

Treatment of Missing Data

TIP!

☑ Do not impute missing outcome data or missing baseline data in a quasi-experimental design. A QED cannot meet WWC standards if missing data are imputed.

For each analysis, the sample will be composed of all matched students with observed data for the outcome measure and both baseline measures used to assess baseline equivalence (i.e., Accuplacer score and first-generation-to-college status). Missing data will be handled by casewise deletion; no missing outcome measures, baseline measures, or other covariates will be imputed.

Adjusting for Multiple Comparisons

TIP!

☑ Adjust for multiple comparisons to lower the chance of a false positive finding.

As indicated in Table 6, two of the outcome domains of interest in the study – academic achievement and credit accumulation – will be tested using multiple outcome measures. Within each outcome domain, we will apply Benjamini-Hochberg adjustments to any statistically significant findings. This approach is consistent with WWC practice, and is intended to account for inflated chance of a Type I error (i.e., finding a statistically significant effect in the sample when one does not exist in the population).

Following the WWC Postsecondary Education review protocol, which indicates that the longest follow-up period should be treated as primary, we will apply the Benjamini-Hochberg correction for the impacts on the longest follow-up period for outcomes in the same domain: (1) at the end of the first semester Applied Mathematics course for the math achievement domain and (2) at the end of students' 4th semester in community college for the credit accumulation domain. We will not adjust for tests of impacts in earlier semesters.

3.6 Baseline Equivalence Testing

In addition to using propensity score matching to select a matched comparison group of students from traditional pre-algebra sections, we will assess the equivalence of the treatment and comparison students in each analytic sample for each outcome. Because the outcomes do not have "natural" baseline measures (i.e., the same measure at baseline), we will assess equivalence using a measure of baseline academic achievement (Accuplacer Elementary Algebra test) and a measure of SES (first-generation-to-college student status). The analytic sample for each outcome will be defined as the students who have non-missing post-test data and non-missing data for both baseline measures. Baseline equivalence will be established for each outcome using the analytic sample for that outcome; across outcomes, analytic samples may vary somewhat for different outcomes (and the same outcome measured at different points in time) due to differences in missing data.

TIP!

☑ Assess baseline equivalence for the analytic sample (or samples). Do not include any student who is missing the outcome measure in tests of baseline equivalence. In QEDs, the WWC requires that baseline equivalence be assessed for the sample of students that have both non-missing baseline data and non-missing outcome data.

☑ Assess baseline equivalence for *each* analytic sample. Remember that the analytic sample may differ from one contrast to another, depending on what data are missing.

We will assess baseline equivalence of each analytic sample using a multilevel model reflecting the structural features of the design (i.e., treatment and comparison groups formed at the course section level; blocking by cohort; students matched within propensity-score strata). Specifically, we will use a modified version of the model described above for testing intervention impacts. However, we will move the baseline measure to the left-hand side of the model, retain the treatment indicator, cohort indicators, and matching strata on the right-hand side, and omit all other covariates. The parameter estimate for the treatment variable (γ_{01}) will provide an estimate of the magnitude of the baseline mean difference between the treatment and comparison students in the scale of the baseline measure.

KEEP IN MIND…
In this example, baseline equivalence is assessed using a statistical model, accounting for the structural features of the design (i.e., matching strata, cohorts). The WWC will also accept a comparison of unadjusted baseline sample means for the intervention and comparison group to establish baseline equivalence.

Level-1 (student-level):

$$Y_{ij} = \beta_{0j} + \sum_{d=1}^{D} \beta_{dj} MatchStratum_{dij} + \varepsilon_{ij}$$

Level-2 (course-section level):

$$\beta_{0j} = \gamma_{00} + \gamma_{01} MODMATH_j + \sum_{m=2}^{M} \gamma_{0m} Cohort_{mj} + \mu_j$$

Where:

Yij = baseline score for student *i* in course section *j*
β_{0j} = baseline score in comparison course section *j*
β_{dj} = vector of coefficients for *d* matching strata
$MatchStratum_{dij}$ = indicator for matching stratum *d* for student *i* in course section *j*
γ_{00} = average baseline score in comparison courses
$ModMath_j$ = treatment status for course section *j*
γ_{01} = baseline score difference between treatment sections and comparison course sections
γ_{0m} = vector of parameters for the effects of cohort
$Cohort_{mj}$ = set of dummy variables indicating cohort A, B, or C for course section *j*
ε_{ij} = error term for student *i* in course section *j*
μ_j = error term for course section *j*

For continuously-scaled measures (e.g., Accuplacer), we will calculate the standardized baseline difference (Hedges' *g*) by dividing the parameter estimate (γ_{01}) by the pooled standard deviation derived from the unadjusted sample standard deviations for the intervention and comparison groups.

For binary measures (e.g., first-generation-to-college status), we will report the percentage of students in the comparison group who are first-generation college students. Using the same modified model described above for estimating the magnitude of the baseline difference, we will calculate and report the model-adjusted percentage of students in the treatment group who are first-generation college students. Both of these percentages, as well as the number of students in each condition, can be used to calculate a Cox index (an effect size for binary measures) instead of Hedges' *g*.

$$Cox\ Index = \left[ln\left(\frac{p_t}{1-p_t}\right) - ln\left(\frac{p_c}{1-p_c}\right)\right]/1.65$$

Where, p_t is the probability that a student in the treatment group is a first-generation college student, and p_c is the probability that a student in the comparison group is a first-generation college student.

The treatment and comparison students will be considered to be equivalent on a given measure if the baseline difference is ≤ 0.25, given that we will control for the baseline measure in the impact analysis model in that circumstance, regardless of the p-value.

Also, unadjusted comparison group means, adjusted treatment group means, and standard deviations at baseline will be reported for the Accuplacer Elementary Algebra test and for first generation to college status.

4. References

Budtz-Jorgensen E, Keiding N, Grandjean EM, Weihe P (2007). *Confounder selection in environmental epidemiology*. Assessment of health effects of prenatal mercury exposure. *Annals of Epidemiology*; 17:27-35.

Deke, J. & Puma, M. (2013). Coping with missing data in randomized controlled trials. Evaluation Technical Assistance Brief for OAH & ACYF Teenage Pregnancy Prevention Grantees.

Judkins, D. R., & Porter, K. E. (2015). *Robustness of ordinary least squares in randomized clinical trials.* Statistics in Medicine. *Advance online publication. doi:10.1002/sim.6839*

Maldonado, G., Greenland, S., (1993). *Simulation study of confounder-selection strategies*. American Journal of Epidemiology 138(11); 923-936.

Price, C., Goodson, B., Stewart G. (2007). *Infant Environmental Exposure to Thimerosal and Neuropsychological Outcomes at Ages 7 to 10 Years Technical Report Volume II*. Bethesda: Abt Associates, Inc.; Prepared for Immunization Safety Office Centers for Disease Control and Prevention.

Puma, M., Olsen, R., Bell, S., & Price, C. (2009). What to do when data are missing in group randomized controlled trials (NCEE 2009-0049).

What Works Clearinghouse (March, 2014). *WWC Procedures and Standards Handbook, Version 3.0.* http://ies.ed.gov/ncee/wwc/pdf/reference_resources/wwc_procedures_v3_0_standards_handbook.pdf

What Works Clearinghouse (March, 2016). *WWC Cluster Design Standards.* http://ies.ed.gov/ncee/wwc/pdf/reference_resources/wwc_cluster_standards_030416.pdf

What Works Clearinghouse (March, 2016). *WWC Reviewer Guidance for Use with the Procedures and Standards Handbook, Version 3.0.* http://ies.ed.gov/ncee/wwc/pdf/reference_resources/wwc_reviewer_guidance_030416.pdf

What Works Clearinghouse (August, 2015). *WWC Review Protocol for Individual Studies in the Postsecondary Education Topic Area, Version 3.1.* http://ies.ed.gov/ncee/wwc/pdf/reference_resources/wwc_pe_protocol_v3.1.pdf