Impact Study Design Overview

Contents

Citing this Resource:

To reference this work, please cite:

Digital Promise (2020, January 30) Impact Study Design Overview. Every Learner Everywhere.

About the Supporting Organizations

Every Learner Everywhere is a network of twelve partner organizations with expertise in evaluating, implementing, scaling, and measuring the efficacy of education technologies, curriculum and course design strategies, teaching practices, and support services that personalize instruction for students in blended and online learning environments. Our mission is to help institutions use new technology to innovate teaching and learning, with the ultimate goal of improving learning outcomes for Black, Latinx, and Indigenous students, poverty-affected students, and first-generation students. Our collaborative work aims to advance equity in higher education centers on the transformation of postsecondary teaching and learning. We build capacity in colleges and universities to improve student outcomes with digital learning through direct technical assistance, timely resources and toolkits, and ongoing analysis of institution practices and market trends. For more information about Every Learner Everywhere and its collaborative approach to equitize higher education through digital learning, visit www.everylearnereverywhere.org.

Digital Promise is a nonprofit organization that builds powerful networks and takes on grand challenges by working at the intersection of researchers, entrepreneurs, and educators. Our vision is that all people, at every stage of their lives, have access to learning experiences that help them acquire the knowledge and skills they need to thrive and continuously learn in an ever-changing world. For more information, visit the Digital Promise website and follow @digitalpromise for updates.

Impact Study Design Overview

What is an impact study?

Gathering meaningful evidence on instructional effectiveness is challenging for any institution, which is why there's so little trustworthy evidence to guide decision-making. The best way to gather meaningful evidence about what happens when a course is changed in a specified way is to conduct an impact study. An impact study is a systematic, objective method for measuring the effects of some change, for example, a new approach to teaching or introduction of a new curriculum. Impact studies involve setting up a fair ("apples-to-apples") comparison between a new way of doing something and the way things are usually done.

An impact study involves measuring one or more outcomes we care about for two groups, one of which has the experience we want to study (the "treatment") and the other of which does not. Ideally, the only thing that varies between the two groups is the experience of the treatment, enabling us to conclude that any difference in outcomes was caused by the treatment.

Impact studies tell us not only whether or not the treatment made a difference, but also how large the difference was (and whether it was positive or negative). They can also tell us whether the treatment affected some types of students more than others. Decision makers look to past impact studies for evidence about whether a new approach or intervention is worth trying, and also may want to run impact studies on changes their institution is trying to determine whether the change should be sustained, scaled up, or dropped. Impact studies are an essential part of institutional continuous improvement initiatives.

The need for evidence

A good example of the demand for impact evidence can be found in the spread of digital learning tools, including adaptive courseware. Adoption of adaptive courseware is on the rise, but many campus leaders and faculty want to see evidence of its positive impacts before putting serious resources into implementing adaptive courseware on a broad scale. The only way to know with confidence that a course redesign incorporating digital learning is improving student outcomes is to conduct an impact study.

Imagine you're the dean of the math department at a college or university. Four years ago, your department started implementing adaptive courseware to reduce drop-out rates in Math 100, the first-year math course with the highest enrollment rates. Some of the Math 100 course sections are fully online and use courseware exclusively. Other sections use it in conjunction with face-to-face instruction. Still other course sections are using courseware to some degree, but you're not sure how—probably for assigning and receiving homework assignments. And then there are course sections that aren't using the courseware at all.

Impact Study Design Overview

After four years of using adaptive courseware, the Vice Provost of Academic Affairs asks you, "What is the impact of adaptive courseware on Math 100? Can you show me solid evidence that it is improving our Math 100 success rates?" You're given 12 months to evaluate adaptive courseware's impact on Math 100 sections before a report of your findings is due to the VP. Where do you start?

Laying the foundation of your project

Identify the goals of your study

Even when a question about impact seems clear-cut, e.g., "Can you show me solid evidence that courseware is improving our Math 100 success rates?" The answer is not as simple as comparing the proportion of students earning As, Bs, or Cs from year to year (the usual definition of "course success"). There is still complexity both in defining the groups to be compared and in figuring out how impact will be measured. A well-designed impact study will tell you whether observed differences are "real" and reveal factors contributing to student outcomes. Impact studies can also be designed to yield data that can suggest refinements of the treatment to make before trying it with new groups of students. Before designing your impact study, make sure you know what questions you're trying to answer and why. Returning to the Math 100 example, questions might include:

- Does our use of adaptive courseware improve Math 100 success rates?

- Does our use of adaptive courseware improve Math 100 success rates for specific subsets of students? For example, what impact does it have on traditionally underserved student populations?

- Do some instructors have higher success rates with the adaptive courseware than others?
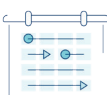
Notice that these study questions refer to "our use of adaptive courseware" rather than just "adaptive course" or the name of a courseware product. This is deliberate. **Any digital learning product can be used in many different ways, and student outcomes will be a function of the way in which the product is used ("implemented") rather than something that is inherent in the product itself.**

You'll want to make sure that you and the Vice Provost are in agreement that these are the key questions for the study to address. After such an agreement has been reached, you'll still have the challenge of specifying the data you need to answer the study questions. Adaptive courseware is being used in different ways and to a different extent in various sections of the Math 100 course on your campus. You don't need to study every possible use to address the Vice Provost's questions. You might, for example, decide that comparing course sections using courseware in conjunction with face-to-face class sections (treatment) to course sections relying solely on face-to-face sessions and conventional homework without using the courseware at all would be the cleanest comparison.

Establish roles and responsibilities

Impact studies require the involvement of more than just those working with the data. The success of a study can be insured by having the right players involved in the study's set-up and implementation as well as data analysis and interpretation. Prior to designing your study, identify who will work in the following roles:

Who	What they do
Study lead	The study lead is responsible for ensuring all the project's moving parts and players are working together in an effective manner, from resourcing to design, implementation, data capture, analysis, and data interpretation.
Leadership champion	Someone at the institutional leadership level needs to be invested in this work and willing to allocate support and resources for completing the research and acting on the results.
Faculty	Faculty who teach the course are on the front-line in its delivery. The way in which they have their students use the digital learning product is the treatment you're evaluating.
Institutional research	• Institutional Research staff are crucial to accessing needed information on student characteristics and past experiences and have responsibility for safeguarding data privacy. Additionally, IR staff are likely to be the ones conducting the impact analysis, so it's important to work closely with them to make sure the analysis addresses your questions.
Institutional Review Board (IRB)	• Impact studies that education institutions conduct with the sole purpose of improving instruction do not need IRB review. But if faculty or staff intend to publish research articles describing the study and its findings, IRB review and either approval or certification that the study is exempt from IRB review is necessary.

Start designing your impact study early

Learning is a complex activity. When measuring learning outcomes, the variability in instructors' experiences or students' baseline knowledge can make it hard to tell whether any observed differences in student performance are attributable to courseware use or to external factors or pre-existing differences between the students in different course sections. To be able to interpret course outcomes appropriately, it's important to begin designing your impact study early so that it will provide data that reflect the effects of courseware use in your classrooms. To yield accurate and reliable data, an impact study design should meet the following requirements:

1. **A comparison course condition**. To measure impact, the new or "treatment" version of the course you're evaluating must be compared to an alternative or "business as usual" (BAU) version of the course. The two or more conditions to be compared should be carefully defined, and each instructor's agreement to implement the instructional condition assigned to him or her should be obtained during the impact study planning phase, before the course has begun.

2. **Minimum condition size of 30 students.** Courses should be large enough to yield a sample size of at least 30 students in each of the conditions being compared (e.g., courseware used to offer a fully online version of the course versus a face-to-face class with courseware used only for homework).

3. **Student equivalence**. Measuring the impact of a new version of the course that incorporates digital learning requires having conditions that are equivalent in terms of other variables that might influence course outcomes. Chief among these are characteristics of the students enrolled in the course sections (conditions) being compared. Clearly, you could not say anything about the effectiveness of a new approach to using adaptive learning if all the students experiencing that approach were students who had to take developmental (remedial) mathematics while most of the students in the comparison course sections were in the college honors program.
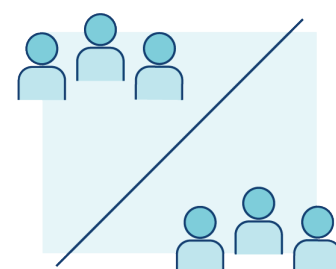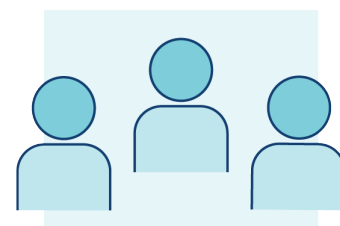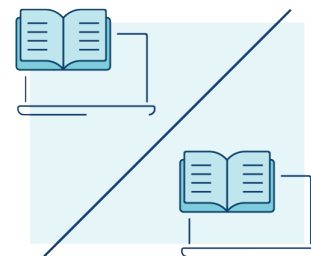
The student characteristic with the biggest influence on course outcomes is typically prior achievement. If students in the conditions being compared are fairly similar but not identical in terms of prior achievement, statistical techniques can make the groups equivalent. Other student characteristics affecting course outcomes (such as whether the student is enrolled full time) can be handled in the same way. Offices of institutional research maintain these kinds of student data and can provide it for use in statistical models that balance the conditions being compared in terms of student characteristics.

Student characteristics that should be balanced across conditions:

1. Prior achievement (e.g., placement test score, grade point average)
2. Course repeater status (yes or no)
3. Enrollment status (full or part time)
4. Gender
5. Race/Ethnicity
6. Income/Pell eligibility
7. Age

4. **Recommended: Instructor equivalenc**e. Because student learning is influenced by many characteristics of instructors, including their knowledge of the subject matter, their grading practices, and their enthusiasm for the subject, studies should be designed to disentangle the influence of different teaching techniques from those of instructor characteristics. Instructor equivalency is especially important when there are just a few instructors teaching the course.
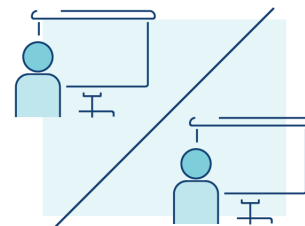
Ideally, there would be a large pool of course instructors and instructors would be assigned at random to implement either the treatment (e.g., adaptive learning) or the comparison condition (e.g., lecture plus the same online homework for every student). When this design is used, we can be fairly confident that, on average, the two conditions have equally effective instructors. An alternative when there are few instructors is to use a study design in which each instructor teaches sections in the different ways being compared. However, most colleges and universities are uncomfortable assigning instructors to teach in different ways, and few instructors like teaching multiple sections of a course using different methods during the same academic term.

The best feasible alternative often involves comparing student outcomes in sections taught by multiple instructors teaching the new version of the course with outcomes of students taught by the same instructors in the past without the new digital learning tool or innovation. Data from previous terms can be compared to data from the current term in the impact analysis. Except in cases where policies around course prerequisites or the nature of the student body changed significantly from year to year, a study design with instructor equivalence will usually be preferable to one using data from different instructors teaching in the same academic term.

5. **Recommended: A common course learning assessment.** Course grades are important and are usually one of the outcomes a college wants to analyze in an impact study (measured either as average grade for the course or as the percentage of students who "succeeded" by earning an A, B, or C). However, different instructors use different grading practices: One section's grades could be based on essays, attendance, participation, and quizzes while another section's grades could be based on three exams and a final. In addition, the commonly used five-value A-F scale does not make fine distinctions between different degrees of learning. (For example, students with quite different levels of learning may all end up with a "B.") It is recommended that impact studies use a second, more precise measure of student learning to increase the likelihood of being able to detect differences between treatment and comparison conditions in student learning (if there are any). For courses that prepare students for a standard licensure or certification exam, scores on that exam can be used as the impact study's outcome measure. For other courses, there may be a final end-of-course examination covering all the course material. If the same final examination is used in both the treatment and the comparison course sections, scores on this exam can be analyzed to ascertain whether student learning differed. Academic leaders can facilitate collecting evidence on the impact of changes in course design by requiring use of a common final assessment in all course section.
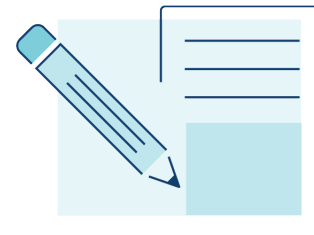
In the ideal impact study, students and instructors would be assigned to treatment and comparison conditions at random (by chance) because this is the best way to establish group equivalency. However, students and faculty don't like being assigned at random and impact study designs using random assignment are rare in postsecondary education research. The two designs colleges and universities use most often for course impact studies are retrospective and concurrent comparison

Retrospective comparison group design with same instructor(s)	Concurrent comparison group design with different instructors
Financial considerations	Ensure that the terms of payment (e.g. per student, per course, flat fee) are best suited for your institution. • Could better pricing be obtained by adjusting payment metrics, such as an institution-wide flat fee instead of per user charges? • Are discounts available for particular circumstances (e.g. Pell Grant funding)?
Scope of agreement	Ensure that the scope is sufficiently broad to meet your needs. • Does it cover all intended types of users (e.g. students, instructors, administrators, visiting scholars) who will need to use it (sometimes, these are called "Authorized Users")? • Are you allowed a sufficient number of logins? • Will students be able to access the courseware from all the devices they might use (e.g., types of computers, types of browsers, mobile devices/tablets)?
Deliverables	Ensure that the license agreement is clear regarding what the vendor will furnish. • Is the courseware going to be off-the-shelf, semi-customizable, or an adaptive platform that can be used to build a custom course? • Is the length of access suitable (e.g. single semester, multiple quarter use, year-long)? • How much work is the vendor accountable for in ensuring deliverables meet your institutions accessibility guidelines and work well with assistive technologies?
Acceptance period	Ensure that there is an adequate acceptance period that gives you sufficient time to inspect and test that the courseware is acceptable and works properly.

Impact Study Design Overview

Retrospective comparison group design with same instructor(s)	Concurrent comparison group design with different instructors
Description Student outcomes from course sections taught by one or more instructors implementing a new approach, such as adaptive learning, are compared to student outcomes from sections of the same course taught by the same instructors using a different approach in prior academic terms.	**Description** Student outcomes from sections of the course taught using a new approach, such as adaptive learning, are compared to student outcomes from other sections of the course taught by different instructors using a different approach during the same academic term.
Requirements and Conditions • Instructors implementing adaptive courseware who have taught the course in prior semesters. • Total of at least 30 students in sections using the new approach and at least 30 students in the course in prior semesters when a different approach was used. • No significant changes from year. To year in prerequisite requirements (e.g., placement policies) • When collecting data from previous years, it should be from the equivalent academic term (i.e., spring semester outcomes should be compared to the prior spring, not to fall outcomes).	**Requirements and Conditions** • A common course final and/or standardized grading practices across sections. • Total of at least 30 students in sections using the new approach and at least 30 students in sections where the new approach is not being implemented. • The instructional approach used in comparison sections that are not using the new approach should be documented.

Benefits	**Benefits**
• Having the same instructor for treatment and comparison course sections typically means the same instructor quality. • Data on student grades and student characteristics for the comparison sections has already been collected by the Office of Institutional Research.	• Course and institution prerequisites (e.g., admissions criteria) would be the same for all students during the study implementation. • If you have a large enough number of instructors, differences in instructor quality are likely to even out across conditions. • If course sections use a common final exam, differences in the evaluation of student learning are minimized. • More accurate data on faculty practices in treatment and comparison sections can be collected based on the present rather than on recall of past teaching practices.
Drawbacks	**Drawbacks**
• Any changes in institution or department policies that affect student characteristics or course performance will be an alternative explanation for any differences between treatment and comparison condition course outcomes. • It may be more difficult to collect accurate data on instructor practices (e.g., whether all students received the same homework problems) because instructors in the comparison condition are being asked to recall practices from prior semesters. • Instructors may have used different examinations and grading criteria in the past.	• Instructor differences introduce additional variation into the statistical testing and weaken the strength of conclusions regarding the impact of the treatment. • Even with a common final exam, instructors may use different criteria for determining course grades.

Data collection

You will need a dataset with information about individual students in the treatment and the comparison versions of the course. The data you need fall into three major categories.

1. Common measure of course outcome. An impact study involves one or more outcomes measured for both treatment and comparison students. Using scores on a common end-of-course assessment allows us to ensure that the student outcomes we're comparing for the two course conditions are the same. Institutional Research offices will have course grades and can generate measures of course success (earning an A, B or C), but often will not have final examination grades for students who took the course in prior academic terms. Exam grades usually have to be collected from course instructors. If the conditions being compared used the same digital learning product but used it in different ways, it may be possible to get learning system data from the product vendor to support the analysis of more specific outcomes (such as learning in particular skill areas or number of concepts mastered). Many vendors are happy to support research on their product and can help arrange data access.

2. Student characteristics. While course examinations and grades give some indication of learning outcomes, they're not the whole picture. Once you've designed your study, you'll need to collect information about students and instructors for both the courseware-using groups and the comparison groups—whether those are "business as usual" sections of the course taught in the same term or historical data from previous years.

- A prior achievement measure for each student is very important because prior achievement is typically the strongest predictor of future achievement. In a statistical model, it explains most of the differences in outcomes. Comparing course outcomes for students in different course conditions without controlling for their prior achievement can lead to wrong decisions about the effectiveness of a change in the course pedagogy (either suggesting a positive impact when there isn't one or hiding a real benefit).

- Academic information about students' past experiences, enrollment, and study plans can be useful in making sense of their course outcomes. The most predictive of these variables is whether or not the student has attempted the course previously. Those who enrolled in a course and failed to complete it successfully in the past are likely to struggle.

- Student demographic information is what will help you disaggregate your data to see the courseware's effect for different kinds of students. It is important to address the equity goals of this work.

- The left-hand side of the table below lists pre-existing student characteristics that are related to student course outcomes and should be equated statistically for the different conditions being compared. On the right-hand side are student characteristics that are less critical to include in the analytic model but still important and generally available from Institutional Research offices.

Impact Study Design Overview

Essential Student Data	Useful Student Data
• **Prior achievement measure:** There are many possible choices for this measure. You could use SAT or ACT score, GPA, or score on a placement exam. If the course you're studying has a prerequisite course, grade or final exam score from that course could be used. • **Academic information:** Whether student has taken the course before • **Demographic information:** Race/ethnicity, gender, Pell Grant eligibility	• **Academic information:** Enrollment status (full- or part-time), major, transfer status, number of college credits • **Demographic information:** Age, First-time-in-college status

3. **Context and course implementation.** For organizations committed to continuous improvement processes, measuring impact on student outcomes is necessary but not sufficient to drive improvement. Courses can always get better, and a measure of impact (be it positive, negative, or null) by itself does not tell you how to improve what you're doing. Especially when you've invested effort in redesigning a course to incorporate technology and have found no impact on student outcomes (or worse yet, a negative impact), you're going to want to know what happened. Further, if you want to add to the knowledge base around effective integration of technology, you'll need to have documented the implementations practices and context for your technology use well enough that others will be able to tell whether your findings are relevant to their circumstances and your evaluation can be included in research syntheses.

Recommendations for information you should document about the context for your impact study and the way the course was implemented in the conditions being compared is shown in the table below. You'll be asked for this information if you seek to publish your study in a research journal, and it is much easier to compile the information while the course is being implemented rather than trying to collect it later.

Impact Study Design Overview

Contest	Implementation in Both Treatment and Comparison Sections	Additional Information About Implementation in Treatment Sections
• Course subject • Section size • Course prerequisites or corequisites	• Modality of instruction (face-to-face, online, blended or hybrid) • Course length (in weeks) • Use of mastery learning (yes/no) • Primary instructional approach (lecture, discussion group, project-based learning, flipped classroom, emporium, other) • How course grade determined • Weekly contact hours with instructor • Required course materials (textbook, courseware, other)	• Main purpose for which students used the digital learning tool • Amount of time students spent using the digital learning tool in a typical week • Percentage of grade based on student's work with the digital learning tool • Instructor use of data from digital learning system to shape future instruction

In addition, it is useful to gather information on the experiences and perspectives of students and instructors using the new approach. They can tell you how they used (or failed to use) the digital learning tool and what they liked and what they disliked about it. This kind of information is very useful when you plan future uses of digital learning. Student surveys and focus groups are two techniques for capturing students' experiences and opinions. We recommend having someone other than the course instructor conduct these data collections and that information identifying individual students be removed from their responses prior to sharing with faculty.

Analyzing and interpreting your data

Once you have collected your data, you will need someone from your Institutional Research office or a statistical consultant to help you set up an appropriate analytic model and determine:

- Whether or not there were statistically significant differences between conditions after equating them in terms of pre-existing student characteristics

- The size and direction of the difference between conditions

- Whether or not there were significant differences between treatment sections taught by different instructors or between different kinds of students in the treatment condition

The IR office can run statistical models to explore differences in student performance between pre-implementation and post-implementation while controlling for prior achievement, student demographics, and academic information. Additionally, these models can be used to explore differences in performance based on race/ethnicity, gender, and Pell Grant status to examine the impacts of courseware implementation for underrepresented minorities and low-income students.

These results will be presented in terms of statistical significance and effect size. With each analysis, the tested assumption is that there is no difference between the "business as usual" course and the course where digital learning is being implemented in the way you've specified. So, when a statistical test returns a significant result, this means there is evidence to support the inference that any observed difference between average outcomes for the two groups is not due to chance.

If you're running your impact study with instructors who are new to the kind of teaching and use of digital learning in your treatment condition, remember that deep change takes time. Based on its experience with colleges introducing adaptive courseware, APLU estimates that it takes about three semesters to start seeing substantive improvements in student success. With multiple iterations, faculty get more comfortable and familiar with the software, learn how to better onboard students to the courseware, and how to best prepare a course that is supported by adaptive courseware.

The **effect** size of an impact study is the difference between the average outcome for the treatment condition and the average outcome for the comparison condition, divided by the standard deviation (a measure of the amount of variation) for the entire sample. If the treatment and comparison conditions do not differ, the effect size is 0. Positive effect sizes mean the treatment group did better; negative effect sizes mean the comparison group did better.

Statistical significance indicates that the observed difference in outcomes between treatment and comparison groups would be very unlikely to occur by chance if the treatment had no effect. The concept of statistical significance should not be confused with importance or practical significance.

Impact Study Design Overview

Revisiting our earlier example about Math 100 and its evaluation questions, here are some possible impact study outcomes and their interpretation:

1. **Does courseware improve Math 100 success rates?**

Positive Impact

Course success rates are higher in Math 100 sections implementing the courseware.

The way instructors used the courseware led to significantly better student outcomes. Such a positive result is unusual the first time that instructors use a new instructional approach but certainly justifies using the new approach in future terms, perhaps involving more instructors.

Null Impact

Course success rates are the same for Math 100 sections implementing the courseware as for those that don't use it.

Even without a difference in course outcomes, you might want to continue using the courseware if students and instructors responded to it positively and/or it saved students the costs of textbooks. You can also explore differences in outcomes for different course sections to look for clues as to what to do next. If some instructors get very good outcomes, what did they do differently? If students of a particular demographic (e.g., Latinx) had poor outcomes with the treatment, what was their experience of it?

Negative Impact

Course success rates are lower in Math 100 sections implementing the courseware.

This result is disappointing, but it should be remembered that this is the first time the new approach was tried. Possible areas for exploration are negative student sentiment, lack of onboarding on how to use the courseware, or misalignment between the courseware and the course final examination. You might want to address one or more of these in a second try implementing the courseware prior to abandoning its use altogether.

2. Does use of courseware improve success rates for specific kinds of students? What impact does it have on traditionally underserved student populations?

Positive Impact

Implementation of the courseware improved student success rates for underserved students.

This would be a very positive outcome for the first implementation of a new instructional approach. You would not only want to continue using the digital learning approach but also want to find out more about why it's effective, for underserved students and consider spreading the approach to other courses.

Null Impact

Implementation of the courseware had no impact on student success rates for underserved students — achievement gaps haven't changed.

Such a result is common the first time a new approach involving digital learning is used, and especially if your student sample is small. By itself, this finding would not warrant dropping the new approach; at least you have done no harm. You should explore why the new approach was not effective for underserved students, particularly if it appeared to improve learning for other students.

Negative Impact

Implementation of the courseware had a negative impact on success rates for underserved students.

Supporting underserved populations and reducing achievement gaps is an institution-wide goal. If introducing courseware had a significant negative impact for underserved students, the institution will want to understand why and either cease using the courseware or change the way it is implemented.

3. What can we learn from instructors whose students have high success rates with the courseware?

Significant Differences

Course success rates are significantly different for different instructors implementing the software.

Variation in results for different course sections the first time a new approach is tried is to be expected. And this variation can be a source for new insights. Data analytics and reports from faculty and students can be used to identify differences in implementation practices associated with differences in course outcomes. Practices used by the most successful instructors can be documented and shared with other instructors.

Null Impact

Success rates are equivalent for different instructors implementing the courseware.

If success rates are consistently high, it's time to open the champagne. If success rates are consistently low, interviews with instructors may reveal issues with the courseware design or its compatibility with the local curriculum and instructor preferences.

These results will be presented in terms of statistical significance and effect size. With each analysis, the tested assumption is that there is no difference between the "business as usual" course and the course where digital learning is being implemented in the way you've specified. So, when a statistical test returns a significant result, this means there is evidence to support the inference that any observed difference between average outcomes for the two groups is not due to chance.

If you're running your impact study with instructors who are new to the kind of teaching and use of digital learning in your treatment condition, remember that deep change takes time. Based on its experience with colleges introducing adaptive courseware, APLU estimates that it takes about three semesters to start seeing substantive improvements in student success. With multiple iterations, faculty get more comfortable and familiar with the software, learn how to better onboard students to the courseware, and how to best prepare a course that is supported by adaptive courseware.

Revisiting our earlier example about Math 100 and its evaluation questions, here are some possible impact study outcomes and their interpretation:

Are you an Every Learner cohort institution? If so, Digital Promise will work with your team to design impact studies, gather the required data, and run statistical models to evaluate the impact of your digital learning innovation.

Advancing equity and continuous improvement

Understanding aspects of the way you are implementing digital learning that are working well and aspects that are falling short is key to continuous improvement and advancing equity in college classrooms. If some student groups do worse with the new digital learning approach than they do in conventional classrooms, you'll want to understand why. Similarly, differences in the student outcomes that different instructors obtain with the new approach incorporating digital learning should be explored to look for indications of best (and worst) practice.

The student characteristics information you collected from your institutional research office can be used in analyses that pinpoint student vulnerabilities. For example, do full-time students get more benefits from adaptive courseware use than part-time students do? Does courseware use have larger or smaller effects for low-income students? Differences for these student groups might lead you to think about course time requirements and the cost of instructional materials in treatment and comparison versions of the course, respectively. Your hunches about possible barriers for different student groups can be explored through interviews or focus groups with students from these groups who took the course. Note that if you want to make statistical comparisons such as these for different kinds of students, you'll need at least 30 students in each condition with the characteristic (e.g., part-time enrollment or low income) you're concerned with. That means using data from larger courses or combining data from multiple academic terms to get a large enough sample to test the impacts of courseware on underserved groups.

Impact Study Design Overview

Another way to gain insight into the way in which the treatment can be improved is to use data collected through student surveys and focus groups. These data collections can address issues known to be important for first-generation, low-income, and underserved-minority students such as:

Do students have ready access to adequate computing devices and internet connections? If internet connections get dropped, pages take a long time to download, or the screen doesn't display important content, learners can easily get discouraged and give up.

Do students feel they have reasonable odds of succeeding with the courseware? Adaptive systems are designed to give each learner what he or she needs, and for students entering the course with weaker preparation, this can be an overwhelming number of content modules to work through.

Do students in sections using courseware feel isolated? Many students, especially women and students of color, enjoy the social aspects of classroom interactions. If each student is working individually on a computer, students can feel a lack of connection with their fellow students and the instructor.

It is worthwhile also to look at course outcomes for different instructors and different kinds of instructors:

Are outcomes different in sections taught by instructors with different amounts of experience implementing the new digital learning approach? There may be instructors who have never used an online component in their classes and are only doing so in response to a department requirement. There may also be instructors who have been using courseware in one way (for example as a homework tool) for a long time and find it challenging to start using courseware in a different way (for example, using information on student struggles within the courseware to decide how to allocate their time during face-to-face class sessions).

Do outcomes differ for regular faculty members and part-time adjunct instructors? It is not uncommon for regular faculty to teach a few course sections a semester while adjunct or part-time instructors may teach up to six course sections—often in different institutions. Faculty members may have received professional development on how to implement new instructional techniques while adjunct instructors did not.

Do sections taught by some instructors have particularly good outcomes? Exploring the details of what the most successful instructors did in terms of introducing their students to the courseware and coaching them on how to use it, helping students make connections between what they learned using courseware and in-class activities and examinations, and leveraging information on how students are doing in the courseware to make mid-course corrections in their lesson plans are techniques to look for. Your most effective instructors can be great resources for improving your implementation design and for coaching others who teach the same course in the future.

The important thing to remember is that improving teaching and learning and closing equity gaps are both continuing challenges. No one impact study will surmount these challenges, but if such studies become part of business as usual, impact studies can be an essential tool for addressing equity and improvement.
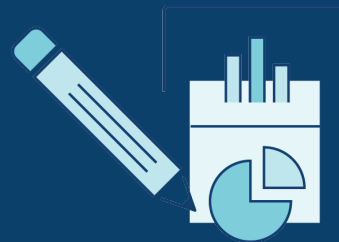
Helpful links

For more resources, visit our website

For questions, contact us.