Center for Technology in Learning

SRI International

Measuring Learning: A guidebook for gathering and interpreting evidence

Center for Technology in Learning, SRI International

This guidebook has been developed to help you think about how to evaluate the early outcomes of your technology-supported course redesign.

This is a first draft, and we welcome your feedback. If you're willing to answer a few questions to help us improve the guidebook, please go to: http://www.surveymonkey.com/s.asp?u=102883208988

Thank you.

Sponsored by hp invent

ACKNOWLEDGEMENTS

PURPOSE

As an instructor incorporating new technologies into your practice, you are bound to notice changes that occur, both in your classroom environment and in your students' performance, attitudes, and behaviors. This guidebook is intended to help you decide what changes are important to notice and how to measure them. The evidence that you collect will enable you to tell a compelling story about your course redesign project and its impact on your teaching, your classroom, and your students.

Conducting a high quality evaluation that yields valid and valuable information about what is taking place in the classroom does not always require trained evaluation researchers. This guidebook will lead you through the evaluation process, enabling you to pull together information to create an emerging picture of the changes that have occurred in your classroom and for your students. In particular, your evaluation can:

"Conducting a high quality evaluation that yields valid and valuable information about what is taking place in the classroom does not always require trained evaluation researchers."

- Help you to understand whether you made the changes you were hoping to make, and whether you achieved the outcomes you were hoping to achieve;

- Help you make decisions about what's worth doing again, and what you can improve next year; and

- Build a compelling story about the value of your project that you can present to stakeholders and potential funders.

In education, changes made in teaching and learning often take many years to fully manifest. Nonetheless, educational innovations do lead to some more immediate outcomes. These outcomes are often beneficial in their own right; they also provide evidence that the project is having an effect and that further transformations may follow. This guidebook focuses on measuring "emerging evidence" — outcomes that can be realistically observed and measured within your first year of implementation. If you have the resources for a longer-term evaluation, the same thought process will apply.

Designed to be clear, concise, and practical, this guidebook will walk you through the steps of the evaluation design process, and serve as a workbook to help you take those steps: defining key objectives, linking project activities to project goals, identifying sources of evidence, setting up comparative analyses, collecting and analyzing data, and reporting findings to internal and external stakeholders.

OVERVIEW OF THE EVALUATION PROCESS

What observable or quantifiable things can illustrate the changes that have occurred in your course, your teaching, and your students? This can include types of evidence you already record (including attendance, retention rates, and student performance on homework and tests) as well as other sources of evidence that can strengthen your story (such as student behaviors and the quality of student interactions in class).

How will you collect and measure the evidence that you have identified? What instruments will you use, and how will you design the evaluation so that you can measure improvement? A strong comparison design will enable you to make strong claims about the impact of your tablet project.

How will you coordinate data collection to ensure that you obtain the information you need? How will you make sure that your data are credible and that they accurately represent your students and your project?

How will you analyze and interpret data in a way that will enable you to make strong claims about your project? How will you present a compelling story about your project to stakeholders?

STEP 1

ARTICULATE THE GOALS OF YOUR COURSE REDESIGN PROJECT

Why are you doing this project?

Planning an evaluation begins with a clear understanding of why you have undertaken the project to begin with. Project goals should be defined as relatively broad objectives that your project aims to realize. A strong goals statement is a concise articulation of your ambitions that reflects the overall purpose of your course redesign project.

Goals for educational technology projects often fall into a few main categories, such as student learning, academic engagement, classroom environment, or equity. Student-focused goals are likely to include improvements to student learning and students participating more deeply in academics. Instructor professional learning goals may include acquiring skills in a new pedagogical style (for example, leading student-centered activities), or you may hope that through your project you will experience the value of asking open-ended questions. Goals for changes to the classroom environment might include increased collaboration or improvements to *formative assessment* (gathering information about student understanding in order to inform instruction) and *feedback* (providing constructive reactions and suggestions to students in response to their work). Ultimately, teacher learning and classroom culture goals tie into goals for student learning: the reason that we hope to see certain changes in teaching practices and classroom environment is because we expect those changes to be beneficial for student learning as well. Your project may also aim to have impacts beyond the classroom; for example, you may hope to increase the numbers of minority students pursuing a major in your subject area.

A clear statement of project goals will be the basis of defining your project or refining its design (**step 2**), selecting a focus for your evaluation (**step 3**), and ultimately deciding exactly what you want to measure (**step 4**). For example, if one of your main goals is to increase student participation, you'll want to make sure to include measures of student participation (such as counts of the numbers of students participating in discussions) in your evaluation design.

"A clear statement of project goals will be the basis of defining your project or refining its design, selecting a focus for your evaluation, and ultimately deciding exactly what you want to measure."

Sample Goals Statements
1. The goal of this project is to improve the effectiveness of the course for building students' understanding of geologic processes and students' skills in field geology.
2. This project is designed to improve teaching and learning by creating an interactive classroom environment and integrating formative assessment into classroom practices. Through this project, the course should more successfully improve student understanding of calculus concepts.
3. The goal of this project is to use technology to link theoretical instruction, hands-on experience, and real-world applications in chemistry in order to create a course that more effectively fosters student understanding of chemistry concepts.
4. The course redesign is intended to create a richer learning experience that is better able to improve students' abilities to solve practical engineering problems and complete real-world engineering projects.

Step 1 To Do: What are the main goals of your course redesign project?
1. _____ _____
2. _____ _____
3. _____ _____
4. _____ _____

STEP 2

DEFINE HOW YOUR COURSE REDESIGN PROJECT WILL ACHIEVE YOUR GOALS

How does the new design of your course map to the goals you set out?

Once you have defined your project goals, the next step is to articulate how your course redesign is going to achieve them. You may already have a good idea of what you plan to do in your classroom, or perhaps the project is already under way. Think now about linkages: how exactly will your project design lead to accomplishment of the goals you set out in **step 1**?

For projects that introduce the use of technologies, it is important to lay out the specific ways that you expect using those technologies to lead to your goals for teaching and learning. For example, if your project uses tablet PCs, you should consider the specific affordances of tablets, such as:

". . . if your project uses tablet PCs, you should consider the specific affordances of tablets. . ."

- *Tablet PCs are personal and portable.* These characteristics enable students to do hands on work collecting data, running simulations, and accessing resources, and they make it possible for students to work on their tablets anytime, anywhere. As a result, students might be more engaged in their work, and spend more time on task.

- *Tablet PCs have pen-based writing capabilities.* The ability to write directly on the computer screen offers students new opportunities for taking notes in class and recording observations in the field. If your project includes drawing chemistry models or writing mathematical equations as a specific project activity, you might expect students to improve in this skill.

- *Tablet PCs are often networked.* Networking enables students to interact wirelessly with one another and with the instructor. Individual wirelessly networked devices offer a whole range of possibilities for enhancing classroom collaborations and changing the dynamic of classroom interactions overall. For example, using quick polls to gather formative information about student understanding, and in turn providing immediate feedback to students to clarify misconceptions, can be a powerful use of networking capabilities. This use of tablet PCs might link to your goals for formative assessment.

Of course, for any technology-enhanced educational project, the hope is to change teaching and learning in ways that go beyond direct application of the technology itself. For example, if you are using the tablets to integrate real-world applications into coursework, perhaps you want to extend this goal by taking students out into the field more frequently. When you map out what you're doing to achieve your goals, be sure to think about things that are specifically enabled by the technology and things that aren't.

Geologic Field Mapping with Tablet PCs: Project Design

One of the main goals of this project is to help students better understand the processes and patterns that shape geographic field areas. In the redesigned courses, students will use tablet PCs and other technology—such as geographic information systems (GIS) and global positioning systems (GPS)—to map geologic formations while they are out in the field. The instructor believes that using mobile technology in the field will enable students to visualize the landscape at multiple scales and with multiple data layers, and to compare what they see physically before them with what they see on the tablet. The instructor hopes that these opportunities for visualizing the geologic phenomena at hand will greatly assist students in understanding landscape processes.

Tablets for Classroom Interaction and Calculus Learning: Project Design

The goals of this project fall into two main categories: student learning (specifically, to improve student understanding of calculus concepts) and classroom environment (to increase active student participation and peer interaction, and to improve feedback processes). A key component of the project is the use of student response systems on tablet PCs. Because of the anonymity afforded by the tablets, instructors hope that students will feel comfortable sharing their ideas with classmates. Further, the ease of soliciting responses from students will enable instructors to assess student understanding frequently and in a variety of innovative ways, thereby making the instructor aware of student understanding and enabling the instructor to target teaching accordingly. The instructor expects that with these classroom practices in place, student learning will increase.

Step 2 To Do: For each of your main project goals, describe how your redesigned course is intended to achieve it.

1.

2.

3.

4.

STEP 3 CREATE YOUR EVALUATION QUESTIONS

What are the fundamental questions that you hope to answer through your evaluation?

Evaluation questions are a beacon to guide the evaluation planning process, from identifying data sources to selecting an evaluation design. Your evaluation questions should reflect the goals that you have set for your course redesign project and for your evaluation. If your project aims to increase student interaction with real-world science, for example, then an evaluation question might be about whether more of your course assignments use real-world contexts than they did before your course was redesigned.

Carefully posed evaluation questions can help to ensure that you design and conduct an evaluation that stays true to your goals.

Guidelines for evaluation questions:

- *Find a good balance between generality and specificity.* Questions such as *"Did the project work?"* are too vague to be particularly useful, because they don't point to any particular outcomes or criteria of interest. Conversely, questions such as *"Did the project teach students how to create graphs on a Tablet PC?"* are probably too specific, because focusing solely on detailed technology outcomes can cause you to overlook the larger results of your project for teaching and learning.

- *Ask questions that you can answer within the time frame of the evaluation.* It is important for you to have some evaluation results within the one-year timeframe of your project grant. Consequently, a question about post-graduation employment patterns (e.g., *"Do more students from the technology-enhanced class obtain jobs in scientific fields as compared with students from the traditional class?"*) is likely to be beyond the scope of this particular evaluation.

- *Consider asking questions about both implementation and outcomes.* The evaluation can certainly give you good information about *whether* your course redesign helped students to learn more; but that question alone doesn't help you to learn *how that happened.* If your goals include making fundamental changes to your course, be sure to include evaluation questions about *implementation*: the degree to which those changes were successfully implemented in your first year.

- *Ask comparative questions that reference a relevant control group.* The best evidence shows a shift in learning relative to a group that has not experienced the course redesign (called a "*control*" or "*comparison*" group, as opposed to the "*treatment*" group that went through your redesigned course). When asking these comparative questions, make sure that you are clear about who and what is being compared. Asking simply, *"Are students more engaged?"* is not sufficient because it begs the question, *"More engaged than what?"* A better way to word this question would be to ask: *"Are students in the technology-enhanced course more engaged in the academic course content than students in the traditional version of the*

"Carefully posed evaluation questions can help to ensure that you design and conduct an evaluation that stays true to your goals."

same course?" Since we're focusing on emerging evidence, the comparison group and experimental design do not have to be perfect—just good enough to indicate that a measurable and important shift has occurred.

Examples of Evaluation Questions
1. Are students gaining a deeper conceptual understanding of scientific processes in structural geology than they did before the course was redesigned?
2. Has active student participation in class increased as compared with previous, non-technology-enhanced versions of the same course?
3. Are formative assessment and feedback more common and more rapid in the redesigned course than in the traditional course? Does the instructor utilize the information gained through real-time formative assessment to tailor instruction to meet student needs?
4. What were the biggest challenges to successful implementation? How can those be overcome next year?

Step 3 To Do: Write down your evaluation questions. Note that as you progress through the evaluation planning process, you will have a number of opportunities to revisit your evaluation questions and to refine them to reflect more detailed statements of project outcomes.

1. _____

_____

2. _____

_____

3. _____

_____

4. _____

_____

STEP 4 | DEFINE SPECIFIC LEARNING OUTCOMES

If your project is a success, what specific changes will have occurred?

Fast forward to the end of the year. You have been teaching your course in new ways all semester, and both you and the students think the changes have been a great success. What do you notice? What are you doing differently? In what ways is your classroom environment different? What are students doing that they weren't doing, or weren't able to do, before?

Whereas *goals* are overarching ambitions for a project, *desired outcomes* are the specific changes or improvements that you hope to see as a result of your course redesign project. In light of the broad goals you have established (**step 1**), think now about concrete outcomes associated with those goals.

Two important distinctions about outcomes:

- *Short-term vs. long-term*. In any education project, some outcomes will be seen almost immediately, while others require more time to come about. For example, teaching students to take notes on their tablet PC and giving them opportunities to do so may lead to improved note-taking skills rather quickly. Research has shown that over time, improved note-taking skills can lead to increased student learning. It is important to recognize, however, that student learning (as traditionally measured by test scores) is a longer term outcome that may not be immediately visible. It is important to know which outcomes you expect to see right away; which might take most of the semester to develop; and which might not be realistic to achieve in one year's time.

- *Proximity*. Is each outcome a direct or indirect result of your project? You can expect the greatest change to occur on indicators that the design of your project addresses directly. For example, if your project requires students to use tablets for various class activities, you can expect that students will use tablets more frequently than they did before. However, if your course redesign does not require students to access real-world data for their projects, while it's possible that they might start doing this on their own, it's less likely that data on course connections to the real-world will show improvement. It is a good idea to base your selection of outcomes on the specific activities of your project.

Although the outcomes of each project will vary depending on the project itself as well as on the instructor, students, course, school, and a host of other contextual factors, there are some outcomes that are commonly associated with technology projects in STEM courses. The following is a sample list of target outcomes for tablet-based projects in college-level science, math, and technology courses. The most important outcomes you will need to measure, of course, are those that describe improvements in student learning. This list also suggests other outcomes you might want to study, in part because they are expected precursors to ultimate student learning outcomes.

"You can expect the greatest change to occur on indicators that the design of your project addresses directly."

Student Learning

- Students learn more _____ (fill in the blank based on your specific course topics) and have a better understanding of related concepts, e.g., students have a better understanding of the relationship between derivatives and integrals, and acquire mathematical habits of mind

- Students improve their _____ skills, e.g., developing a proof or analyzing data

- Student improve their 21st century skills, e.g., communication and problem-solving

Student Affect and Behavior

- Students are more focused on the academic content of the course

- More students attend class on a regular basis

- More students actively participate in class activities and discussions

- More students continue through to the end of the course; fewer students drop out

- Student perceptions of their own abilities improve; they have more confidence in the subject area and do not feel alone when they don't understand something

Classroom Environment

- The classroom is more interactive: students learn from one another, students work collaboratively, students communicate with the instructor

- A wider range of students actively participate in class activities

- Students have more autonomy and exhibit more responsibility in directing their own learning

- Assessment is used formatively and students do not feel threatened by routine assessments of their understanding

- Feedback cycles are shortened: instructors assess student understanding, provide immediate feedback to students, correct misunderstandings, and refine instruction to meet students' needs

Instructor Professional Learning

- Instructors learn to design a course based on specific student outcomes

- Instructors understand how to use technology to transform instruction and enrich learning

- Instructors increase their repertoire of pedagogical strategies: they ask open-ended questions, increase classroom equity by soliciting participation from all students, or allow students more autonomy in directing their own learning

- Instructors improve student assessment and feedback processes: they garner information about student understanding and refine instruction more nimbly in order to meet student needs

- Instructors create a community of practice to tap into shared resources and work collaboratively to identify best practices

Geologic Field Mapping with Tablet PCs: Target Outcomes
In keeping with the project goals, which centered around student acquisition of knowledge and skills, the instructors identified a number of specific learning gains they hope students will achieve. Two of these target outcomes are: 1. Students improve their field mapping skills; they are more facile in the field and they are better able to identify the geologic processes that shaped an area. 2. Students have a better understanding of geologic concepts that are listed in the course syllabus (such as rocks and minerals, plate tectonics, geologic time, and surficial processes).

Step 4 To Do: List your target outcomes. Think about all the outcomes you're hoping to see. Which are direct outcomes of your course design, and which are more long-term or distal? Make sure you have a path of logic linking each outcome to the features of your project that make it reasonable to expect the outcome will come about.
1. _____ _____ 2. _____ _____ 3. _____ _____ 4. _____ _____

STEP 5 IDENTIFY SOURCES OF EVIDENCE

How exactly will you know that your desired outcomes have occurred? What will you be able to see?

Based on your research questions (**step 3**) and your more specific target outcomes (**step 4**), the next step is to select indicators that will provide evidence of progress and success. In education, as in science, *evidence* refers to concrete information that tells you that something is or is not happening. One way to think about evidence is to ask: What will I be able to see that shows that this outcome is occurring in my classroom or for my students?

Gathering evidence always involves a process of compiling artifacts or observing phenomena, and then recording and systematically measuring those phenomena. Sometimes the process is common and straight-forward. For instance, gathering attendance data involves observing and recording which students are in class, and then tallying the counts of students. Gathering evidence that shows that students engage in higher quality discussions, however, requires a more conscious decision about how to observe, record, and measure "quality" in a discussion. One option is to develop a set of criteria for high quality discussions and to then apply that rubric to your recorded observations about what occurred during class discussion.

"Gathering evidence that shows that students engage in higher quality discussions, however, requires a more conscious decision about how to observe, record, and measure "quality" in a discussion."

One relatively straight-forward way to collect data about a number of outcomes is to ask students what they think. *Self-report data* can come from a variety of sources, including surveys, interviews, and focus groups. While self-report data is relatively easy to collect and aggregate, it does have a number of significant drawbacks. Most problems stem from the fact that self-report data doesn't tell you what has actually happened, but instead provides *opinions* of what has happened. It is easy for respondents to confound factors: for example, enthusiasm for a new device may translate to a false sense of having mastered the content. Furthermore, unless surveys have gone through rigorous processes for reliability and validity, the data can be skewed due to individual differences in interpreting questions and rating responses. The examples of evidence provided here offer more tangible indicators for a selection of commonly-desired outcomes.

Evidence of student learning in the content area

Student content area learning subsumes multiple facets of student learning including content knowledge, conceptual understanding, and subject-related skills (such as titration or creating graphical representations). The most common source of evidence for student content area learning is student performance on regular course assessments. There are also other valid sources of evidence that can offer different insight into student understanding and the types of thinking and reasoning skills that students may have acquired.

Examples of evidence of student content learning:

• Students ask probing questions, share thought processes, and challenge ideas

• Students perform better on course assessments, including homework, labs, presentations, and exams

• Students get better course grades

It may also be useful to look in your classroom for factors that the literature reveals to be pre-cursors to student learning. Although these factors are not measures of student learning per se, they can add to the picture of what is currently happening in the classroom and what implications that holds for future learning. For instance, time on task does not inherently provide evidence of student learning. However, research has shown that the amount of time that students spend working with a particular concept is associated with how much they learn about it. Thus, if students in the technology-enhanced class are spending more time in class, and more class time grappling with content, than students in traditional classes, then it is likely that student learning will ultimately increase in the technology-enhanced course.

Examples of evidence of pre-cursors to student content learning:

• Students spend more time on task

• Students take better notes

Evidence of student academic engagement

Increased student engagement in course content is a common goal for many technology initiatives in education, but it can be difficult to measure. Many evaluations use student responses on surveys as a primary gauge of engagement. Yet asking students whether they *like* the course as currently designed is very different than looking for *behaviors* that suggest academic engagement. There are a number of other options for collecting more rigorous and more compelling evidence about student engagement in learning. Specifically, research shows that when students are more engaged in their learning, their behaviors change in a number of concrete ways. Students who are interested in the subject at hand are more likely to exhibit "good" student behaviors, such as attending and participating in class, completing homework assignments, and even doing reading and other work beyond what is minimally required.

". . .asking students whether they like the course as currently designed is very different than looking for behaviors that suggest academic engagement."

Examples of evidence of student engagement:

• Students spend more time in class: i.e. they attend class more often, arrive early and leave late, and persist through to the end of the course

• Students participate more actively in discussions and other course activities

• Students more commonly attend office hours to extend classroom conversations

• Students complete assignments more regularly

Evidence of changes to the classroom environment

Many of the changes to your classroom environment are directly related to the implementation of your technology project. For instance, student uses of technology, the content and structure of course activities and assignments, and the

instructor's use of formative assessment may all be components of your project. As such, information about these elements of the classroom environment provide evidence of project *implementation* more than they provide evidence of project *outcomes*. Evidence of successful implementation can in turn serve as confirmation of a changed classroom environment (e.g. showing that the classroom has become more student-centered).

Examples of evidence of classroom environment changes

- A greater percentage of course activities and assignments require students to use technology to do hands-on work with multiple representations, simulations, and data
- Students have more opportunities to choose which methods or resources to use for different tasks
- The instructor solicits information about student understanding more frequently and uses a greater variety of formative assessment strategies
- Feedback cycles are more rapid: there is less time between gauging student understanding and providing feedback and/or refining instruction to meet student needs
- More activities and assignments involve group work and collaboration
- More activities and assignments involve connections to the real world: e.g. field trips, use of real data, interactions and/or collaborations with professionals, etc.
- Activities and assignments call for more use of visual representations and require students to make links across multiple representations
- A greater range of students participate actively in class activities

Geologic Field Mapping with Tablet PCs: Sources of Evidence
In order to answer the evaluation questions which pertain to student learning of content and skills, the primary sources of evidence for this project come directly from student work products (including student assignments, tests, and even student discourse). Some sources of evidence include:

1. Students make fewer errors in their geologic maps.

2. Students produce high quality geologic maps that are thorough, insightful, and clear.

3. Students ask probing questions in the field that demonstrate an initial understanding of the geologic history of the area.

4. Students perform better on course assessments. |

Tablets for Classroom Interaction and Calculus Learning: Sources of Evidence

In keeping with this project's emphasis on changes to classroom practice, many of the sources of evidence that this project will document relate to project implementation. With the hope that the changed classroom environment will lead to improved student learning, some sources of evidence are also intended to document student learning. Sources of evidence include:

1. A greater percentage of students actively participate in class activities.

2. The instructor solicits information about student understanding more frequently and uses a greater variety of methods to do so.

3. The instructor provides more feedback to students and does so in a more timely manner.

4. The instructor refines instruction (alters pacing, revisits material, addresses misconceptions) based on feedback from students.

5. Student homework assignments reflect conceptual understanding: they are accurate, and they show logical thought progressions in finding solutions to problems.

6. Students perform better on course assessments.

Step 5 To Do: Describe the evidence that you will use to gauge progress on each of the target outcomes that you identified.

1. _____

_____

2. _____

_____

3. _____

_____

4. _____

_____

STEP 6

DESIGN YOUR EVALUATION: DATA, INSTRUMENTATION, AND COMPARISON DESIGNS

How will you capture and measure evidence? How will you structure your evaluation?

In the previous step you thought about sources of evidence that can help you answer your evaluation questions and illustrate progress toward the outcomes that you hope to see. Evidence must be systematically collected and measured in order to serve your evaluation needs. In this section we will describe how to measure indicators and collect data to garner the evidence required for your evaluation. In the next section we will describe the selection, adaptation, and development of the instruments that you will use for data collection.

As an instructor working in an educational institution, you already collect data that will be useful for your evaluation. You may also already have some measurement methods in place for your school accreditation process. For example, you may already have defined Student Learning Outcomes (SLOs), or other ways to meet accreditation requirements. These existing data provide a great foundation for your project evaluation and will save time and energy in the evaluation process. Using these established measures should enable you to see if you progressed further on your target outcomes in the technology-enhanced course than you did in traditional versions of the same class.

In addition to making use of existing measures and indicators, you may also decide to design and develop some new measurement tools to collect additional information that will help you build a more robust picture of whether your project goals and outcomes are being achieved.

Tablets for Classroom Interaction and Calculus Learning: Data Sources and Instrumentation

For this evaluation design, the project leader decided to make use of the Student Learning Outcomes (SLOs) her team had defined for school accreditation purposes. In particular, the department had identified three specific learning outcomes for calculus students: a) analyze an integral to determine the correct method of integration; b) apply theorems to determine if a series converges or diverges; and c) understand the concept of a limit. The department designed a new assessment to measure each of these concepts and skills. All Calculus I students (in all sections in both the fall and spring semesters) would take the new assessment, providing parallel data from both technology-enhanced and traditional sections of the course and thus enabling the project leader to compare student performance on the end-of-term assessment. In addition, the project leader asked her colleagues to administer a pre-test at the beginning of the year. This test will not be identical to the SLO assessment developed by the department, but it will nonetheless provide a baseline to establish that the different course sections all started out with students of comparable ability.

While some evidence is inherently quantifiable (such as attendance rates), other sources of evidence require special measurement techniques. Some of these measures you probably already use as part of your common practice. The most familiar examples of educational measurements are student assessments of learning, including curriculum-based exams and standardized achievement tests. Indeed, even student learning is not inherently quantifiable, and there is much debate around the best way to measure it. For other outcomes, you may need to develop your own basic measurement tools. For example, if you want to evaluate student projects according to specific tablet-related criteria, you may need to develop a simple rubric that enables you to do so.

Collecting data on student learning

Your evaluation will almost certainly need to answer the question, "*How much are students learning?*" Some of the most common measures of student learning include student performance on homework, labs, projects, and exams, as well as final course grades. These data are convenient because, as an educator, you probably collect them as part of your common practice. Further, because many of these measures remain consistent from one year to the next, or among different sections of the same course, they provide valuable information for doing comparative analyses of student learning.

"Assessing student achievement in ways that are appropriate to your new course is an essential, albeit perhaps technically difficult, thing to do."

Assessing student achievement in ways that are appropriate to your new course is an essential, albeit perhaps technically difficult, thing to do. Sound assessments provide an accurate measure of student learning. When implementing a new educational project, you may be teaching new concepts or highlighting new subject-related skills. Existing assessments (including both curriculum-based assessments and standardized achievement tests) may not address the knowledge and skills that are currently emphasized in your course. Consequently, you may find it useful to use new assessment strategies that address the focus of your re-designed course.

Key strategies to consider for measuring student learning:

- *Align assessments of student learning with course content.* At a minimum, it is useful to revisit your existing course assessments and make sure they are capturing the depth of learning you hope to see in your redesigned course. Revising assessments to align with course content will allow you to obtain a more accurate picture of student understanding and to detect more nuanced improvements in student learning.

- *Use an evidence-centered approach to design your assessments.* Ask yourself: What knowledge, skills, and abilities do I want to measure? What behaviors do we need to elicit in order to measure the knowledge, skills, and abilities? What kinds of assessment items and tasks will elicit those behaviors?

- *Consider the range of item and assessment task formats when designing your assessments.* Multiple-choice, short open-ended, and expanded essay questions provide evidence of different kinds of students understanding. For example, multiple choice items are often used to provide evidence of whether students can recall or recognize facts and principles, whereas open-ended items and tasks are often used to provide evidence of understanding at a deeper level by

requiring students to express relationships, provide explanations, or interpret data.

- *Measure student understanding through concept tests.* Concept tests are short questions that focus on the key concepts of a lesson. The instructor queries the students, who respond via technology. After they have responded, students discuss the question and explain their answers to one another. After a period of discussion, the students send in responses to the question again. Eric Mazur of Harvard University (http://mazur-www.harvard.edu/ education/educationmenu.php) has shown that student understanding (as measured by the number of correct answers) generally goes up after this opportunity for peer learning. A number of pre-made concept tests in physics and other science disciplines are available for use from Mazur's Project Galileo website (http://galileo.harvard.edu/).

- *Design rubrics to assess student work.* These rubrics would include criteria for success along dimensions specific to your course redesign, enabling you to explicitly evaluate the target competencies for any given assignment. Rubrics are a good way to judge student performances (e.g. a presentation or a debate) to ensure that you are capturing a reading of component skills.

- *Use technology to ask more open-ended questions.* For example, tablets make it easier to collect and assess student work in the form of a diagram or concept map.

Collecting data on student academic engagement

Regularly recorded statistics regarding attendance, assignment completion, retention, and enrollment in subsequent subject-area courses provide helpful data that can serve as evidence that students are engaged in their course work.

Student behaviors such as arriving on-time or early and actively participating in class activities are also indicators of student engagement that can be valuable to capture. One way to obtain this information is simply to record the numbers of students arriving early, actively participating in class, and staying late. Keeping track of the information will be easier if you take a systematic approach to the data collection.

Finally, you might want to corroborate your observations of student behaviors with evidence from a student survey. Surveys often use *scales* in order to understand more complex notions, such as "student academic engagement". A scale is a set of questions that address different aspects of a single complex concept. All of the questions that comprise a scale are analyzed collectively to provide a more robust picture. There are a number of standard scales that are available for measuring student engagement. Some of the resources in the appendix contain examples of validated scales on this and other topics. In particular, you might want to look through the resources on the Online Evaluation Resource Library (OERL, http://oerl.sri.com/) for tried and true survey scales pertaining to student engagement.

Collecting data on the classroom environment

Your own course materials—such as syllabi, lesson plans, and course assignments—are a useful source of data for examining course content and characteristics of course activities and assignments. For example, if one of the goals of your project is to increase connections to the real world, you might count the number of field trips taken throughout the semester and the number of assignments that require students to make connections to the real world. If your project aims to capitalize on the unique affordances of tablets by having students work with visual representations, then you might count the number of activities and assignments that involve the use of technology for visual representation.

Of course, the actual events and interactions taking place in your classroom provide the most direct information about changes to the classroom environment. While anecdotal comments and observations can be telling, it is also a good idea to collect more methodical accounts of classroom practices. For example, if one of the components of your project is to give students more opportunities to work one-on-one with technology, then you could record the amount of time that students spend using the new technology during class. If your project aims to improve instruction through the use of formative assessment, you could record information about practices related to formative assessment and feedback.

Examples of data and accompanying measurement instruments

Data you already collect

- Student grades on course assessments, including labs, homework, exams
- Final course grades
- Attendance
- Assignment completion
- Retention / persistence / course completion
- Enrollment
- Syllabi, lesson plans, course assignments

Potential new data and associated measurement instruments

- Rubrics for evaluating student work: e.g. projects and presentations based on tablet-specific criteria
- Assessments aligned with skills being taught and emphasized
- Worksheet for recording participant behaviors: e.g. numbers of students arriving early or staying late; numbers of students attending office hours (and reason: are they there because they missed class or to pursue a deeper discussion?); numbers of students responding to questions
- Worksheet for recording time spent on an activity: e.g. student time on task and time using tablets
- Worksheet for recording data on feedback cycles: e.g. instructor frequency of soliciting information about student understanding; types of probes used; numbers of students responding; time elapsed between gathering assessment information and providing feedback / refining instruction

"Your own course materials — such as syllabi, lesson plans, and course assignments — are a useful source of data for examining course content and characteristics of course activities and assignments."

Comparison Evaluation Designs

A strong evaluation design is essential for creating a compelling story about your project. The design of your evaluation will shape the statements that you make about the relationships between your project, changes in your classroom, and the outcomes that were in evidence by the end of the year.

The most basic form of a comparison design study is a simple pre/post design in which data is collected for the treatment class at the beginning of the semester and again at the end. The problem with this design is that there is no way to know how the learning gains observed compare to learning gains for a traditional course. That is, we would expect that the students would learn something from the beginning of the course to the end; but did they learn more than they would have in a traditional version of the course?

This example illustrates the importance of *baseline* data. In particular, the strongest evaluations use pre-measures from both a *treatment* and *control* group, thereby enabling a comparison of the relative changes within each group. For example, by administering a pre-test and a post-test to students in both the technology-enhanced course (the treatment) and the traditional course (the control), it is possible to compare the learning gains made by each group.

"... the strongest evaluations use pre-measures from both a treatment and control group, thereby enabling a comparison of the relative changes within each group."

Another major issue in evaluation is the need to show that any improvements in teaching and learning were in fact due to the project itself, and that the changes were not a result of other factors. The best way to show that the course redesign caused the results is to minimize the other differences between the two groups. To this end, it is important that the students, the instructor, the course content, and the course assessments are all reasonably comparable in both the redesigned course and the traditional comparison course.

There are a number of reasons why the students in the two courses might not be entirely comparable. For example, it is important to think about how the students selected into the particular course that they attend. At community colleges, many students work while also attending school. Is one of the courses a night course, thereby attracting many of the working students, while the other course is a daytime course, attracting many of the full-time students? Were there any scheduling conflicts, such that all of the students taking advanced mathematics courses had to enroll in one section of the course, while the other section is largely comprised of students who are not pursuing higher level math? If these sorts of situations exist, it is important to consider the differences between the two groups of students, and how those differences might impact the results of your evaluation.

Here are several common comparison designs:

Comparison across instructors. In this design, two instructors teach the course in the same year, one in the traditional format and the other in the redesigned format. This evaluation design requires that the two instructors are teaching the same course, with the only difference being the components of the course redesign itself. Students in both courses should be comparable in terms of demographics, working status, prior coursework, and achievement levels (as measured by a pre-test given to both groups, if possible).

By virtue of the fact that this design compares two classes both being taught in the same year, the evaluation does not need to rely on historical data in order to conduct comparative analyses. Instead, the evaluators can collect matching data throughout the year, including baseline measures from both groups. Further, for both groups, evaluators can collect data that has been specifically chosen and designed with this evaluation in mind.

The primary limitation of this design stems from having two different instructors in the different classes, opening the possibility that different results were due, at least in part, to the different instructors, rather than to the redesign of the course.

Comparison across sections. This design involves one instructor teaching two or more sections of the same course in the same year. The instructor teaches at least one section in the traditional format and the other(s) with the technology-enhanced redesign. The requirements for this design are much the same as those of the previous design: same course, comparable students, and equivalent measures. As with the previous design, comparing two classes taught in the same year provides numerous opportunities for collecting extensive, appropriate, and equivalent data from both groups. The fact that both classes are taught by the same instructor has both benefits and limitations for the study. On the one hand, having the same instructor lead both classes means that there are no teacher-level differences to consider. On the other hand, it can be difficult for a single teacher to implement two different types of courses, particularly if the two courses require different pedagogical styles from the teacher. As a result, elements of the redesigned course might creep into the control course (or vice versa), thereby watering down the measured effects of the project.

Comparison across years. In this design, an instructor who taught the course in the traditional manner last year teaches the same course with the technology-enhanced redesign in the current year. As with other comparison studies, this design requires that the two groups have the same course (except for the redesign components), comparable students, and equivalent measures. Because this design compares a treatment class with traditional classes from previous years, it also requires that the appropriate historical data are available. Further, it is not possible to compare indicators from assessments and other measures designed to align with this evaluation. The primary benefit of this design is that the same instructor teaches both the comparison and the treatment courses, but without the risk of contamination, because the courses are taught in different years.

Compare across topics. This design can enable you to build a comparison within a single run of the course and with only one instructor. The design involves one instructor, teaching some topics within a single course in the traditional format, and other topics according to the course redesign. This unique design benefits from the fact that the instructors and the students are the same in both the treatment group and the comparison group. However, the design is complicated by the fact that it requires both the instructor and the students to switch gears mid-way through the course. If the new version of the course is implemented first, the implementation of the traditional format may be "contaminated" (i.e. affected by the treatment condition) since it can be hard for instructors and students to "forget" new and better ways of teaching and learning that they've recently adopted.

Regardless of the sequence, adopting new practices in teaching and learning takes time; reducing the duration of the project may also reduce its chances of success. Finally, in this design, the content of the different units being compared is not the same. If certain units are more difficult than others, the results of the evaluation may be confounded. Nevertheless, if it is not possible to implement any of the other comparison designs, this method can at least provide some clues to the impact of your course redesign project.

Guidelines for designing your evaluation

There are a number of guidelines to keep in mind when thinking about your evaluation design. First, the evaluation design and the data sources go hand and hand; in order to ensure that the design and the data support one another and make a feasible combination, they need to be considered in tandem. For example, if it is important that you use new assessments of student learning because existing assessments are not sensitive to the new types of learning now possible in your classroom, then your comparison design cannot use this measure to compare this year's course to last year's, because last year's students will not have taken the same assessments.

Regardless of the design of your study—but particularly if you are unable to use a comparative design—it is a good idea to provide multiple sources of evidence for each outcome. Findings are stronger when they can be supported with corroborating evidence from other data sources.

The process of developing a complete evaluation design involves many interrelated considerations. By this point, you have thought about a number of important factors: What data will you collect and where will it come from? How will you measure the outcomes and evidence that you have targeted? What type of comparison design will you employ? Having defined answers to these questions, you should look back to make sure that a) your evaluation questions and your evaluation plan still align with one another and b) both the questions and the plan meet your evaluation needs. Are you able to refine your evaluation questions now that you have specified each of the elements of your evaluation plan?

Geologic Field Mapping with Tablet PCs: Comparison Design

For this study, the project leader used a comparison design involving two sections of the same course, offered in the same semester, and taught by two different instructors. Students in both sections were determined to be comparable to one another as far as background, demographics, and previous experience with the subject. Students in both sections took a pre-test, and did one field mapping assignment without any technology. After that, students in the traditional class continued in the same manner, while students in the tablet class began using tablet PCs equipped with the necessary aerial photos and other information, as well as GPS and GIS systems, in the field. The instructors designed a special rubric for evaluating student maps, with a particular focus on content and clarity of presentation. The instructors also developed a simple instrument for systematically recording information about the quality of students' questions in the field. In addition, the instructors administered a survey to garner student opinions about the mapping process (with and without technology) and about resources, tools, and practices that facilitated their acquisition of mapping skills.

Step 6 To Do: Sketch out your evaluation design.

1. What will your evaluation compare to see if students are learning more now that they did before? What about your other outcomes?

    _____

    _____

    _____

2. List your data sources and note instruments that need to be developed.

    _____

    _____

    _____

STEP 7
IMPLEMENT YOUR REDESIGNED COURSE, CREATE INSTRUMENTS, AND COLLECT DATA

What steps will you take to ensure that you obtain the information that you need?

By this point you have done a lot of work to think through each of the elements of your project evaluation. You have defined project goals and drawn a clear path between your project design and those goals. Based on your project goals, you defined evaluation questions that in turn guided the identification of specific outcomes and associated sources of evidence. You also constructed a comparative evaluation design that will yield strong findings about your project. What remains now is to put all of your planning into action.

The implementation of your project evaluation should take place in parallel with your project implementation. It is best if data collection can take place throughout the course of the semester and be as integrated as possible with what you're doing in the classroom. In **step 6** you determined which types of data you would use to gather evidence about your target outcomes. Now you need to obtain the actual data collection instruments and determine the logistics of how you will collect the data.

"It is best if data collection can take place throughout the course of the semester and be as integrated as possible with what you're doing in the classroom."

Guidelines for data collection and instrumentation:

- *Keep careful documentation of data relevant to your evaluation.* Think about the sources of evidence that you identified in **step 5**: Will you have the documentation that is necessary to yield that evidence? For example, if you hope to show that a greater percentage of course activities require students to do hands-on work with technology, then you will need to be able to compare syllabi from traditional and redesigned versions of the course. Do you have a written syllabus that accurately reflects the activities that you did in class? You may already be documenting certain information with respect to your school accreditation requirements. If so, it makes perfect sense to use this information as part of your evaluation data as well.

- *Select, adapt, or develop new instruments.* For each of your data collection instruments, you will need to decide whether: a) an appropriate instrument already exists; b) you will need to modify an existing instrument to meet your needs; or c) you will need to develop a new instrument altogether. For example, if you have decided to create new student assessments that are aligned with your new course focus, you may need to spend some time adapting and/or creating assessment items. If you plan to conduct a student survey, you can probably borrow existing scales that deal with student engagement and elements of the classroom environment, but you may need to develop some of your own questions that address specific components of your project.

- Researchers take numerous factors into consideration when putting together instruments for an evaluation study. For instance, in order for instruments to yield credible data, the instruments themselves must be valid and reliable. A *valid* instrument measures what it claims to measure. A *reliable* instrument is one that provides equivalent results even if it is administered at different times, in different places, or by different people.

- For more information about instrumentation you can refer to the recommended resources described in the Appendix. The Standards for Educational and Psychological Testing (within the Resources for Researchers section) might be particularly useful. You may also want to consult with the relevant research group at your institution. One possibility is to bring an education graduate student on board for your evaluation, so that they can provide assistance throughout the course of your study.

- *Consider the logistics of your instrument administration and data collection.* Usually there are multiple options for any data collection, including both traditional paper-and-pencil methods and electronic or digital methods. When deciding which methods make sense for you, it is a good idea to look into the tools that you may be able to access through your institution. It is possible that your institution already uses electronic or online tools for administrative records such as enrollment, attendance, and grades. In addition, your institution may have scantron technology that would facilitate processing of survey response sheets. Online survey tools (such as Survey Monkey or Zoomerang) can greatly simplify the process of collecting and organizing survey results; these programs can also generate some basic analyses of the survey data.

  "One possibility is to bring an education graduate student on board for your evaluation."

- *Create a data collection schedule.* The evaluation design that you laid out in **step 6** has important implications for data collection. If you are planning to do a pre/post comparison design with a treatment and control group, then you will need to make sure to obtain or collect data to fulfill each of these required data points. To get accurate "pre" measures of any sort, it is better to gather data before you have started implementing your new course, as opposed to once the semester is underway. Similarly, you will get better evidence of the project's impact if the time separation between the pre- and post- measures gives the project enough time to make an impact. Thus, it is better to wait until the end of the semester to gather post data, rather than gathering post data mid-semester.

- *Provide training on how to administer or use the instruments.* For the most part, you will do most data collection yourself. However, in order to gather data from a comparison group, you may be working with a colleague who is collecting parallel data in a traditional version of the course. You may also have asked a colleague to observe your class and to record their observations on an observation protocol. It is always important to train data collectors with regard to the specific focus of your data collection and the intricacies of the specific instruments they will use.

- *Collect data from each member of your sample.* Having data from as many participants as possible means that your data are more representative of your student sample, leading to stronger evaluation findings. Conversely, missing data can have significant implications for the interpretation of your results. For

instance, if you administer a survey to your chemistry class on a day when many students are away on a field trip for a calculus class, then your results will not reflect the opinions of that important group of students enrolled in higher math courses. One strategy for getting the highest response rate possible is to make the data source in question a required part of the class. For instance, you may decide to require that students complete the student questionnaire in order to get their grade for the course. Data that are integral to the course (such as homework assignments and course assessments) are ideal because they are likely to yield a high response rate without lots of hassle.

Geologic Field Mapping with Tablet PCs: Collection Schedule

In keeping with the study design, the general course assessment and the student survey were administered to all students both at the beginning of the semester and at the end, thus providing pre- and post-measures of student performance and student attitudes. The instructors decided that course assessments from previous years were still relevant to the target content and skills of the redesigned course, so they used the existing assessment measures. In addition, geologic maps from each student were collected periodically throughout the semester. The newly-designed instrument for documenting student questions was completed by the instructors of both the technology-enhanced and traditional sections after each class trip to the field.

Step 7 To Do: Write out your data collection timeline.

STEP 8 ANALYZE DATA AND REPORT FINDINGS

How will you analyze and interpret data in a way that will enable you to make strong claims about your project? How will you present a compelling story about your project to stakeholders?

Extracting interesting and relevant findings from raw data is a process that requires careful forethought as well as craftsmanship. We offer a few pointers for working with your data and reporting your findings to key stakeholders, but we do not go into extensive detail on these topics. For more information about data analysis and interpretation and for information about reporting and disseminating your work, refer to the annotated resources list in the appendix. You may also want to consider partnering with a research group at your institution that can provide support throughout the evaluation process.

"Data that do not fit with the rest of the information might indicate that something went wrong somewhere in the process."

Steps in working with your data:

• *Conduct quality assurance by reviewing and cleaning the data.* The process of sorting through your data should be ongoing throughout the data collection period. Above all, it is important for you to be familiar with the data that you are collecting so that you will know right away if something looks amiss. Looking through the data as they come in will also enable you to identify any outliers. Data that do not fit with the rest of the information might indicate that something went wrong somewhere in the process—perhaps a survey item was uniformly misinterpreted by participants; maybe a single participant did not follow instructions properly; or perhaps something was simply entered incorrectly into a database. A deep familiarity with your data will enable you to avoid confusion when it comes to data analysis and interpretation.

• *Analyze the data.* Your evaluation questions (**step 3**) and comparative design (**step 6**) should guide the analyses that you conduct. In all likelihood, the process of collecting data and conducting initial analyses will lead to as many new questions as it does answers. After an initial round of analysis, it may be appropriate to conduct some additional analyses to explore the questions that arose over the course of the implementation.

• Data analyses vary greatly in complexity. Some data analyses are relatively straightforward—in fact, you are probably already accustomed to conducting many of these analyses as part of your routine practice. For instance, finding course averages, comparing the average in one class section that of another class section, and looking at learning gains over the course of a semester are likely all things that you already do to some degree. In addition, you may know how to do some statistical analyses with your quantitative data or methodological analyses of qualitative data. If you are interested in doing more complex data analyses, one option is to pair up with the education or social science researchers

at your institution. Careful data analysis can reveal strong findings about your project.

- *Interpret the data.* The process of interpreting data goes beyond the findings themselves to try to understand how and why the findings might have come about. A key consideration when interpreting data is to think about alternative explanations for the findings that emerged from your data analysis. The comparative evaluation design that you established in **step 6** should help to rule out some of the possible confounding explanations of your findings, but there are invariably others to bear in mind. For example, your data might show that there was not a particularly large difference in the performance of students in the traditional redesigned courses on a series of mini-assessments that you designed to test conceptual understanding. However, if a large number of struggling and low-motivation students had dropped out of the traditional course because they felt frustrated, then the students remaining in the traditional course would be primarily the stronger students. Consequently, the results from the mini conceptual assessments would reflect the performance of stronger students and suggest that the redesigned course was less effective than it actually was. Because any number of factors can influence your data, careful data interpretation is critical for constructing a complete and accurate picture of the impact of your project.

Remember, too, that the data collected in this evaluation is *emerging evidence*—the trends that you see are indications of "changes under development" that may take even stronger hold and result in larger impacts as your project matures.

- *Summarize your findings.* Now that you have analyzed and interpreted your data, you should be able to articulate a number of emergent, evidence-based findings about the impact and effectiveness of your course redesign project. What did you learn that was in line with your expectations? What did you learn that was surprising or unanticipated? Did you meet the goals that you laid out in **step 1**? When you write down your findings statements, make sure that you have evidence to back up each finding.

- *Assess the implications of your findings.* One of the most valuable outcomes of a project evaluation is the opportunity for continuous improvement to your course redesign. Specifically, the findings that emerge from your evaluation will have implications for how you want to conduct the course in the future. In some instances your findings may indicate areas for project improvement. If there are elements of your course redesign which were not as effective as you had hoped, you will probably want to reconsider how to achieve those desired outcomes. On the other hand, positive findings that are strongly connected to your course redesign may lead you to continue to deepen these new pedagogical practices in the future.

". . .if a large number of struggling and low-motivation students had dropped out of the traditional course because they felt frustrated, then the students remaining in the traditional course would be primarily the stronger students."

Guidelines for reporting and disseminating your project story:

The insights you gained from your evaluation acquire newfound value when you share them with other members of the community. A number of stakeholder groups are likely to be interested in your evaluation research. For example:

• Your current funder

• Potential funders

• Other STEM faculty considering technology use

• Other members of your institutional community, including administrators and Other decision-makers

• Other members of the business or political community

Think broadly about the story that you would like to tell about your project. You designed the project to achieve specific goals for teaching and learning. As you implemented your project, you also set out to see whether or not it was effectively achieving the goals you established. Having completed one evaluation cycle, think back to your evaluation questions (**step 3**) and to the concrete evidence that you collected (**step 7**), and consider what you learned. How did your project impact teaching and learning in your class? How did it change the course experience for your students? How was your redesigned course really different from a traditional version of the same course? When you report to stakeholders, you will want to tell a comprehensive story, that links goals to project design to findings and conclusions in a coherent and convincing way.

As you prepare to present your evaluation research to others, it is also a good idea to think about the specific audience that will be viewing your work. What information or emphasis is relevant to that particular audience? What is the best way to present that information so that it will be accessible to your audience? Typically, the following sections are included in evaluation reports:

• Background, including project description and overall goals

• Evaluation questions

• Evaluation procedures, including measurement and evaluation design

• Data analysis

• Findings

• Conclusions (and recommendations)

Within each of these sections, you will want to walk your readers through the logic and process of your evaluation. Thus, you should include descriptions and explanations for each step of the evaluation, including a discussion of your comparative design, the instruments that you used, and your training procedures for administering the instruments, among other particulars. This will allow the reader to understand and appreciate the basis for your claims.

Geologic Field Mapping with Tablet PCs: Findings and Conclusions

The project found that students in the redesigned course did produce higher quality geologic maps with fewer errors than the student maps from the traditional course. However, the instructors were initially surprised when they looked at the data gathered from the instrument they developed to capture student questioning. While students in the technology-enhanced course seemed to be asking rather mundane questions about technical assistance, a greater number of questions from students in the traditional course focused on geology content. Closer inspection of the questions from both classes that focused specifically on geology content, however, revealed that those questions from the technology-enhanced course represented more conceptual depth than those from the traditional course.

Overall, the instructors were greatly encouraged by their findings. In order to minimize confusion around the technical operations of the tablets in the field, the instructors decided that in the future, before delving into field work, they would set aside some time to provide technical training to the students with a particular focus on using the tablets for geology.

Step 7 To Do: Once you have completed your data analysis and interpretation…

  a) What are your 3 main findings?

  b) What 3 things will you do differently next time you run your redesigned course?

RESOURCES

I. EXAMPLE STUDIES

This list represents a collection of research studies related to Tablet PCs in education. The studies focus on a variety of factors, from implementation of Tablet PCs, to perceptions and attitudes toward technology integration, to learning outcomes related to Tablet PC programs. The studies employ a variety of data collection tools and methods, including surveys, interviews, classroom observations, assignment analysis, and assessments of student learning. We hope that this collection of reports will give you ideas about outcomes that might be of interest to target and measure as well as data collection methods that could be relevant for your project evaluation.

Singapore Tablet PC Report: Executive Summary

http://ctl.sri.com/publications/downloads/TabletPC_ExecSummary.pdf

The Singapore Tablet PC project examined programs at two Singaporean high schools in which all students received a tablet PC to use both at school and at home. The research describes the Tablet-PC-based teaching and learning environment and identifies factors that affect implementation at the classroom and school levels. The study emphasizes descriptive information about teachers' and students' use of tablets, attitudes toward tablets, and perceptions of teaching and learning related to tablet use. The mixed-method study includes teacher and student surveys and interviews as well as qualitative analyses of Tablet PC applications and lesson plans. The report presents findings regarding perceptions of convenience and efficiency, changes in classroom practices, effects on learning and motivation, teachers' views of effective professional development for Tablet PC use, ease of technology use, barriers to integration of the Tablet PCs into curriculum, changes in classroom practices and management, and the effects of the Tablet PC on learning and motivation for students of different achievement levels.

Tablet PCs in Schools: Case Study Report, BECTA

http://www.becta.org.uk/corporate/publications/documents/tabletpc_report.pdf

This report summarizes findings from case studies representing a wide range of tablet PC projects in 12 primary and secondary schools in England. The research employed qualitative methods including interviews, classroom observations, questionnaires, and tablet portfolios and usage logs. The report provides an overall analysis of all 12 projects, separately describing the impact of the tablet projects on staff (e.g. reenergized teachers), curricula (e.g. increased cross-curricular work), and students (e.g. improved critical evaluation of information). The report also presents findings about supports and barriers to implementing effective tablet programs and discusses what is needed for success. Staff and student technological training and complementary technologies that enhance tablet usage are also discussed. Finally, the report provides descriptions and findings from each of the 12 individual projects.

The Impact of Tablet PCs and Pen-based Technology on Education

Berque, D.A., Prey, J.C., & Reed, R.H. (Eds.) (2006). *The Impact of Tablet PCs and Pen-based Technology on Education*. Indiana: Purdue University Press.

In April 2006, a group of educators convened at Purdue University for the first Workshop on the Impact of Pen-based Technology on Education (WIPTE). This monograph contains papers presented at the workshop, representing experiences, evaluations, and lessons learned with regard to pen-based technology interventions in education. The papers cover both K-12 and higher education as well as a broad range of disciplinary subjects, from Japanese language to geology to computer science. Through compiling and sharing these works, the editors hope to communicate best practices in the use of pen-based digital technology and help practitioners to identify practices that would be appropriate and beneficial for their own circumstances.

Great Maine Schools Project: One-to-One Laptops in a High School Environment

http://mainegov-images.informe.org/ mlte/articles/research/PCHSLaptopsFinal.pdf

The final report of the Piscataquis Community High School Study describes findings about the one-to-one laptop initiative implemented at this Maine high school. The research used surveys of students, faculty, and parents; interviews with students, faculty, and administrators; and administrative data from the past three academic years as the primary sources of evidence for their study. The survey instruments are included in the appendix of the report. The study provides a good example of analysis and findings based on survey data within the context of an educational technology program.

II. EVALUATION GUIDES: HOW-TO RESOURCES FOR PRACTITIONERS

The guides listed here have been designed as user-friendly guides to evaluation. Many of them contain details about different approaches to evaluation and specific data collection and analysis methods that are not covered in the HP Evaluation Guidebook itself. These resources offer additional background information about theoretical approaches to evaluation. They can also serve as supplemental practical resources throughout the evaluation process, as many of them contain detailed guidelines and working tools for establishing your program model, selecting evaluation measures, and conducting data collection and analysis.

Online Evaluation Resource Library (OERL)

http://oerl.sri.com/

The Online Evaluation Resource Library (OERL) is an online repository of resources for education evaluation, providing a rich collection of sound practices in evaluation and guidelines for their application. The web site is organized by types of evaluation resources (e.g., plans, instruments, reports) and types of projects (e.g., Curriculum Development, Teacher Education). The collection of over 130 instruments contains student assessments, questionnaires, interview protocols, observation protocols and other types of instruments. The collection of 38 plans and 60 reports contains complete and excerpted versions, with accompanying

explanatory annotations. Criteria for sound evaluation practices drawn from the *Program Evaluation Standards, 2nd ed. (Joint Committee on Standards for Educational Evaluation, 1994)* are presented for each type of evaluation resource. In addition, guidelines and scenarios explain how the evaluation resources can be used or adapted and how OERL users can take advantage of the capabilities of the online, interactive environment. The collection has a powerful search tool that allows users to search for specific resources in the OERL library by attributes such as content area, project type, and name of sponsoring university. The OERL collection is being expanded and new plans, reports and instruments are added to the site periodically.

The 2002 User-Friendly Handbook for Project Evaluation (NSF)

http://www.nsf.gov/pubs/2002/nsf02057/nsf02057.pdf

This guidebook provides a highly accessible explanation of each of the steps involved in planning and conducting educational evaluations. The handbook contains four major sections: types of evaluation; steps in doing an evaluation; an overview of quantitative and qualitative data collection methods; and strategies that address culturally responsive evaluation. The first section defines formative evaluation and summative evaluation and describes the purposes of each. The second section describes the evaluation process, from developing a conceptual model of the program to developing evaluation questions, defining measurable outcomes, developing an evaluation design for sampling and statistical analysis, collecting and analyzing data, and disseminating information to interested parties. The third section presents a discussion of the theoretical and practical merits of quantitative and qualitative methods; it also provides information on a variety of specific data collection methods such as interviews, focus groups, surveys, and observations. The fourth section discusses the importance of conducting evaluation in a culturally responsive manner. Finally, the guide includes charts and templates intended to assist program managers as they work through the evaluation process, with specific tools for conceptualizing a program model, identifying key stakeholders, and articulating goals as measurable outcomes.

Using Technology Evaluation to Enhance Student Learning

Means, B., & Haertel, G. D. (Eds.) (2004). *Using technology evaluation to enhance student learning*. New York: Teachers College Press.

This practitioner-oriented volume provides a summary and analysis of the perspectives presented in *Evaluating Educational Technology*, discussed below. This handbook is intended to help identify research methods that are most useful to policy makers, articulate the need for relevant measures of student learning, highlight the importance of looking at effects over the long term, and discuss what policy makers and administrators can do to improve research and implement successful technology-enabled instructional innovations. Each section contains an overview of the arguments made in the companion volume and a series of commentary pieces designed to provide practitioners with the information needed to make research-based decisions concerning the use and evaluation of educational technologies.

Kellogg Foundation Evaluation Handbook

http://www.wkkf.org/DesktopModules/WKF_DmaItem/ViewDoc.aspx?fld=PDFFile&CID=281&ListID=28&ItemID=2810770&LanguageID=0

The Kellogg Handbook is a thorough and well-organized reference for planning and conducting evaluations. The first part of the handbook provides an historical overview of evaluation and explains the Kellogg Foundation philosophy of program evaluation. The second part of the handbook is a how-to guide for planning and implementing project-level evaluations. Project evaluation is split into three components: context evaluation, implementation evaluation, and outcome evaluation. The handbook provides detailed guidelines for identifying stakeholders; developing evaluation questions; budgeting; selecting an evaluator; determining data collection methods; collecting, analyzing and interpreting data; communicating findings; and utilizing evaluation results. Also included are worksheets for project managers to use in applying handbook guidelines to their own project evaluation.

Kellogg Foundation Online Evaluation Toolkit

http://www.wkkf.org/default.aspx?tabid=75&CID=281&NID=61&LanguageID=0

The Kellogg Toolkit is designed to provide guidance to Kellogg Foundation grantees working in a variety of fields not limited to education. The Toolkit provides explanatory information and further resources for the following evaluation topics: where to start; evaluation approaches; evaluation questions; evaluation plan; budgeting; hiring and managing evaluators; and additional resources. The toolkit provides guidelines and worksheets for defining the scope and purpose of your evaluation and for identifying key stakeholders. The toolkit also includes examples of how to link program components to the outcomes you want to see, sample evaluation questions, methods for data collection and analysis, and information about procedures and logistics, reporting and dissemination, and budget.

III. EVALUATION GUIDES: RESOURCES FOR RESEARCHERS

The resources listed here offer in-depth discussions of research designs and methodologies in the social and behavioral sciences. Intended primarily for professional researchers and evaluators, these references offer technical explanations and guidelines for a wide variety of rigorous research methodologies. The books and journals listed below also present analyses of some of the recent advances and current issues and debates pertaining to education research in particular, and social science research more broadly.

Evaluating Educational Technology: Effective Research Designs for Improving Learning

Haertel, G. D., & Means, B. (Eds.) (2003). Evaluating educational technology: Effective research designs for improving learning. New York: Teachers College Press.

In this volume, education, evaluation, and assessment experts discuss considerations and conclusions pertaining to the evaluation of educational technologies and the measurement of student learning. The chapters are organized into three sections. The first section presents arguments for both experimental research (to examine whether or not a particular innovation is effective) and quasi- and non-

experimental studies (to examine why innovations do or do not work and the contextual interrelationships that exist in educational settings). Several chapters discuss issues around defining the intervention and present contrasting views on the purposes educational research should serve. The second section considers the kinds of student learning that should be investigated and how that learning should be measured. The third section looks at the need for longitudinal research designs to detect both long- and short-term effects of technology. This is not a hands-on guide for conducting educational evaluation, but it is a valuable resource for gaining insights about the critical issues and considerations in educational technology evaluation research.

Designing and Conducting Mixed-Methods Research

Creswell, J., and Clark, V. (2006). *Designing and Conducting Mixed-Methods Research*. California: Sage Publications.

This book focuses on practical guidance for designing mixed methods studies and advice for the specific demands and issues of mixed methods research. This resource is designed to be manageable for beginning researchers, providing step-by-step descriptions of the research process, practical guidelines and exercises for researchers to use in designing their study, examples from actual mixed methods studies, and a glossary of terms.

Quantitative Applications in the Social Science, Sage Publication Series

http://www.sagepub.com/booksSeries.nav?series=Series486&desc=Quantitative%20Applications%20in%20the%20Social%20Sciences

The Quantitative Applications series is a collection of digestible books of about forty pages each, covering a wide variety of topics pertaining to quantitative research in the social and behavioral sciences. These books are a valuable source of information for those interested in pursuing a specific research topic in depth. Some titles of interest include: *Cohort analysis, Introduction to generalized linear models, Analyzing complex survey data, Multi-level modeling, Longitudinal research, 2nd Edition, and Relating statistics and experimental design, among others*.

The Sage Handbook of Qualitative Research, Third Edition

Denzin, N.K. & Lincoln, Y.S. (Eds.) (2005). *The Sage Handbook of Qualitative Research*, Third Edition. California: Sage Publications.

The Sage Handbook of Qualitative Research offers a thorough textbook-style resource for qualitative studies. The book covers current issues in qualitative research, inquiry strategies, methods of data collection and analysis, the art of interpretation and evaluation, and future directions. Each chapter focuses on a particular topic, ranging from ethics and politics in qualitative research to institutional review boards, critical ethnography, analytic perspectives, and qualitative evaluation, among numerous others.

Journal of Mixed-Methods Research, Sage Publications

http://mmr.sagepub.com/

The Journal of Mixed-Methods Research is a quarterly, international publication with contributions from the premier researchers and practitioners in the field of

mixed-methods research. The journal aims to contribute to the foundations of mixed-methods research and provide reflections on important issues in the field. The journal includes methodological and theoretical discussions on a variety of topics, including types of research questions and design, sampling issues, parameters of mixed-methods research, and logistics of conducting mixed-methods research. Also included are book and software reviews.

Standards for Educational and Psychological Testing

AERA, APA, NCME. (1999). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.

The Standards were developed to encourage the ethical use of tests and to provide a set of criteria for developing, evaluating, administering, and scoring tests, as well as for making use of test results. Broadly speaking, tests are instruments that provide a standardized process for collecting, evaluating, and scoring performance in a specified domain. Because tests have significant implications for individuals, organizations, and society, it is important for those involved in the testing process to be conscientious of sound testing practices. This book contains sections that cover the test development process (including technical information on validity, reliability, scaling, and scoring), fairness in testing, and testing applications. The book also contains chapters focusing explicitly on the rights and responsibilities of test developers, test takers, and test users (namely, those who analyze and interpret testing results and those who use test information for decision making). Each chapter contains background and explanatory text on the given topic, followed by a series of pertinent, operational standards.

IV. FOUNDATIONS OF LEARNING SCIENCE AND ASSESSMENT: FROM THEORY TO PRACTICE

These resources represent seminal works in the fields of learning sciences, assessment, and educational technology. The books provide theoretical discussions of current understandings in these fields, drawing from recent advances across diverse disciplines including cognitive science and educational measurement. The authors aim to draw connections between theory and practice, discussing the practical impact of new understandings of how people learn, new methodological capabilities, and novel conceptualizations of the role of technology in transforming learning.

How People Learn: Brain, Mind, Experience, and School

National Research Council. (2000). *How People Learn: Brain, Mind, Experience, and School,* Expanded Addition. Committee on Developments in the Science of Learning. Bransford, J., Brown, A., and Cocking, R., editors; and Committee on Learning Research and Educational Practice. Donovan, M., Bransford, J., and Pellegrino, J., editors. Washington D.C.: National Academy Press.

This book synthesizes recent research on the scientific basis of learning and links the research with actual practice in the classroom. The authors argue that new understandings of how people learn require that we update what is taught, how it is taught, and how learning is assessed. The book is organized in three broad categories: Learners and Learning; Teachers and Teaching; and Future Directions, and covers topics from the expert-novice continuum, to teacher learning, to tech-

nology supported learning. The material is presented in an accessible manner, with research conclusions and associated guidelines for practice clearly identified and explained.

Knowing What Students Know: The Science and Design of Educational Assessment

National Research Council. (2001). *Knowing What Students Know: The Science and Design of Educational Assessment*. Committee on the Foundations of Assessment. Pellegrino, J., Chudowsky, N., and Glaser, R., editors. Board on Testing and Assessment, Center for Education. Division of Behavioral and Social Sciences and Education. Washington, DC: National Academy Press.

The Committee on the Foundations of Assessment, convened in 1998 by the National Research Council and supported by NSF, engaged in a three-year study to compile information about new and promising developments in assessment research and practice. This book represents the committee's findings and conclusions. The book discusses advances in cognitive science and human learning (broadening the scope of what is considered important to assess) and educational measurement (broadening capabilities to develop and interpret better assessments). Based on these advances in cognitive science and measurement, the committee points to implications for assessment design. Chapters on assessment in the classroom and the use of technology for assessment may be of particular interest. The report is a comprehensive and somewhat technical review of the current state of educational assessment; it presents components of student achievement, knowledge, and skills that should be assessed and guidelines and critical features for assessment design.

Enhancing Undergraduate Learning with Information Technology: A Workshop Summary

Hilton, M. (Ed.) (2002). *Enhancing Undergraduate Learning with Information Technology:* A Workshop Summary. Washington, DC: National Academy Press.

This book documents a workshop on enhancing undergraduate learning with information technology, which focused on the use of information technology to transform undergraduate science, math, engineering, and technology (SMET) courses. In particular, workshop participants discussed the current state of educational information-technology (IT), current efforts to incorporate IT into SMET classes, the question of how to evaluate the impact of IT on teaching and learning, and the potential of IT to transform SMET learning in the future. Participants also considered cultural and institutional barriers to the use of IT in education as well as challenges in assessing student learning in technology-rich environments and evaluating the impact of technology programs. The book contains a set of case studies of innovative uses of technology in SMET undergraduate courses and steps to maximize the potential of IT for transforming SMET education.