The TLT Group
Teaching, Learning, and Technology
A Non-Profit Corporation

The Flashlight Approach to Evaluating Educational Uses of Technology

As new technologies tumble over one another, and our options for using them multiply, confusion multiplies, too.

Faculty, librarians, IT staff, facilities planners and others need to use evaluation to study the recent past, in order to make wiser, safer choices about what to do next. But how can such studies be designed so that the findings really help to improve outcomes and reduce costs and other stress?

Naïve approaches to evaluation of educational uses of technology often:
- Focus on the technology itself
- Measure changes in goals (outcomes) that are the same for all users (e.g., are all students in this French class learning grammar better as a result of this software? Do all faculty who participated in this workshop now use our course management system better?)
- Focus (only) on hoped-for gains in outcomes, relative to doing nothing;
- Wait to start doing evaluation until the technology seems to be working smoothly;
- So that there's a good chance that the findings will be positive, and can be used to elicit rewards.

The Flashlight approach differs from this, in almost every way. Consider these nine ideas when designing a formative evaluation of some educational use of technology.

1. **Above all, do no harm.** When you get a survey, or are asked to respond to an interview, what's your first impulse? Hit the delete key? Hang up the phone? Is that because past surveys seemed of little benefit to you personally or to people you care about? You're probably in an educational institution. You need to ask yourself if your study is going to make participants feel, in the long run, more likely, or less likely, to participate in educational research in the future. What will they learn about research by helping you out?

So design studies that will help the people who are helping you: the people you are surveying, interviewing, or testing. And make it clear from the start how this is going to be a valuable, rewarding way from them to use their time. For more on this topic, see Section IV.C of this *Flashlight Evaluation Handbook*.

2. **Design the evaluation in order to discover how to improve results.** An evaluation to discover and document success or failure ("summative") asks different questions than an evaluation to discover how to improve results ("formative"). The following ideas should help you design a formative evaluation of an educational use of technology. By the way, it's a lot easier to implement principle #1 ("do no harm") if you're doing a formative evaluation.

3. **Study *what* users did repeatedly with the technology ("activities").** This is the single most important idea in this handout. If you want to improve outcomes of technology use, it makes sense to focus on an *activity* that a technology supports. IT departments must think in terms of technologies: which ones to roll out, which ones to support, how to support them. But educational outcomes are influenced by activities (supported by those technologies).

For example, the Web is touted in part because of its value as a medium for research. But the ways in which faculty create research assignments, and the skills with which students carry out that research, are influenced by many more factors than just the Web or even specific research sites found on the Web. In general these activities change far more slowly than technology. (The good news: if you had focused on a technology, you'd need to wait until that technology was implemented. And your findings would become irrelevant as soon as that technology changed. But when you focus on an activity, you can start any time, and your findings are likely to be useful for many, many years to come. See principle 8.)

Whether you organize your inquiry around an activity or a technology, you'll need to study how the technology was actually used if you want to learn how to improve outcomes. In the end, the *intended* uses don't produce outcomes. But the *actual* uses do. So you will need to discover what those actual uses have been.

Why focus on "repeated" activities? Because evaluation studies the past in order to learn how to improve the future. If you use a particular software module only once a year, then an evaluation of this year's use can't help you until next year. If, on the other hand, you study factors affecting whether students do the online readings, you can apply what you've learned immediately, and next week, and next year. The more often a technology is used in a particular way, the more quickly you can exploit your evaluation findings in order to improve future activities and results.

So, for example, if you want to discover how to get more value from "clickers" (a form of student response system), the first step is to discover what students and faculty are actually doing with clickers. For instance, are some faculty using clickers to ask challenging conceptual questions and then to foster peer debate about those questions? Or to test memorization? Or to help the faculty member decide what to do next? Or to take attendance? Each of those activities has different educational consequences. And what about students? If the faculty use clickers to grade by attendance, do some students send their clickers to class with friends?

You can't discover how to improve outcomes from technology unless you first learn how the technology was actually used.

3. Study *why* they acted that way. Knowing that people acted in a certain way helps you understand what they learned (principle 2). Discovering why they acted that way can help you figure out whether and how to change that activity.

For example, suppose you are studying the value of the online discussion capabilities of a course management system. The value of that system will be influenced in part by whether, and how, students take part in online discussion. So it makes sense to study the factors affecting whether and how each student engage in discussion online. For more on this topic, see this Flashlight workshop on studying and improving online discussion: http://tinyurl.com/37orqz

4. Remember that education does *not* work like a well-oiled machine. Even when the same faculty member teaches two sections of the same course, and has taught them for years, what students do and what they learn will differ. Add technology that increases options for faculty and students: that variation in learning among students will probably increase. So your study design, and your interpretation of its findings, both need to treat human variation and choice as fundamental, not as 'noise' that clouds your findings.

5. Recognize that different students (and faculty) have different goals, needs, and (therefore) outcomes ("Unique uses"). The beneficiaries of any program or innovation differ, but the goal in most instances is to help them all. For example, students using the same innovation may have qualitatively different skills and preparation, career goals, backgrounds, abilities/disabilities. An evaluation design that begins with the assumption that the students (and faculty) may use an innovation, technology, or program differently, and with different consequences, is called a "unique uses" design.

In contrast, an evaluation design that begins with the assumption that certain outcomes should be studied for all students (and faculty) is called a "uniform impact" design.

A typical program needs to be evaluated both ways. For more on unique uses and uniform impacts and how to evaluate each of them, see the *Flashlight Evaluation Handbook,* Chapter III, Section A.

6. Confront the dark side. Too many evaluations organize their inquiries only around the hoped-for gains from technology use. But loss and damage are inevitable, too. *Every* shift in the use of technology results in losers as well as winners. And each participant usually experiences losses as well as gains. These losses can't be reduced or compensated unless they are recognized. An evaluation can help improve results by minimizing loss, just as it can by discovering how to increase gains. And evaluations that explicitly recognize loss are more credible and compelling than those that focus purely on gain.

7. Maximize study value and impact by maximizing engagement of both subjects and the users of the findings.
        In order to improve response: One author or set of authors (A) often studies a set of student or faculty subjects (B) without any particular intention to benefit B. Even more often, subjects who respond never learn whether their time was well spent. In order to increase the response rates of B (not just the number of B's who respond, but how thoughtfully they respond), design the study in order to benefit B, no matter what the study finds. In other words, design a study that can influence action, *no matter what you discover*.
        In order to increase the influence of study findings: In the worst situations, A only publishes the findings if they confirm the original hypothesis. Bad idea. C knows that any study that is submitted is merely a device to persuade C to go against his or her better judgment. So the study findings are often given little attention. They either confirm what C already 'knows' or the study must be flawed. In order to increase the chances that C pays attention to the finding, *involve C in the design of the study and the analysis of the data*. Here too it's crucial to *design a study that can produce persuasive findings, no matter what you find.*
        A final step to improve response: Once the decision makers (C) have made use of the findings, make it a habit to report back to your subjects what actions have been aided by the time they invested in responding to your study.
        In general, a study is likely to either alienate its subjects and users ("what a waste of my time this is! Next time a researcher seeks my attention, I won't be suckered!) or to increase the chance of their engagement in future studies ("This empowered me!") For more on increasing engagement of respondents and decision makers, see Section IV of the *Flashlight Evaluation Handbook* ("Improving Engagement in Your Study").

8. Start evaluating now! For example, if you're interested in using clickers to foster conceptual learning, start evaluating conceptual learning (including faculty development) now, whether or not you're using clickers yet, whether or not you're experienced in their use yet, whether or not you're considering replacing one kind of clicker with another. And if there's a chance you might buy a different kind of student response system next year, your findings can help you
  • choose products;
  • remove barriers to effective use even before the technology becomes available,
  • provide baseline data for measuring the impact of the new technology.

The whole *Flashlight Evaluation Handbook* is designed to flesh out these and related ideas using examples and specialized guides. The TLT Group can also provide coaching (send us your draft plans and we can talk), collaborate with your staff in designing such studies, or we can even do them for you. So take a look at the *Handbook* and contact us if you'd like to talk (301-270-8311; flashlight@tltgroup.org).