# Formative and Summative Evaluation

## Formative Evaluation

Formative Evaluation is a bit more complex than summative evaluation. It is done with a small group of people to "test run" various aspects of instructional materials. For example, you might ask a friend to look over your web pages to see if they are graphically pleasing, if there are errors you've missed, if it has navigational problems. It's like having someone look over your shoulder during the development phase to help you catch things that you miss, but a fresh set of eye might not. At times, you might need to have this help from a target audience. For example, if you're designing learning materials for third graders, you should have a third grader as part of your Formative Evaluation.

Here are some different author's definitions of Formative Evaluation that will help you understand the difference.

### Scriven, (1991)

Formative evaluation is typically conducted during the development or improvement of a program or product (or person, and so on) and it is conducted,often more than once, for in-house staff of the program with the intent to improve. The reports normally remain in-house; but serious formative evaluation may be done by an internal or an external evaluator or preferably, a combination; of course, many program staff are, in an informal sense, constantly doing formative evaluation.

### Weston, Mc Alpine, and Bordonaro, (1995)

The purpose of formative evaluation is to validate or ensure that the goals of the instruction are being achieved and to improve the instruction, if necessary, by means of identification and subsequent remediation of problematic aspects.

### Worthen, Sanders, and Fitzpatrick, (1997)

Formative evaluation is conducted to provide program staff evaluative information useful in improving the program.

### Robert Stakes

"When the cook tastes the soup, that's formative; when the guests taste the soup, that's summative."

### Scriven, (1996)

- "is research-oriented vs. action-oriented"

- "evaluations are intended - by the evaluator - as a basis for improvement"
- "the summative vs. formative distinction is context dependent"

Formative evaluation helps you find out if you are achieving your goals and objects in the formative stages of your project, instead of waiting until you are have finished the project.

Formative evaluation provides:

- A chance to test your ideas and implement changes before completing the entire course (*i.e.,* a way to avoid finding out it could have been done more effectively if only you had known...)
- A way to test both the form and content of your course

An important part of formative evaluation for online courses are 'usability studies.' "In Don't Make Me Think: a common sense approach to Web Usability" Steve Krug defines the importance of user testing:

"If you want a great site, you've got to test. After you've worked on a site for even a few weeks, you can't see it freshly anymore. You know too much. The only way to find out if it really works is to test it.

Testing reminds you that not every one thinks the way you do, knows what you know: uses the Web the way you do."

Ease of Use

Three questions developed by Robert Mager in "Making Instruction Work" help define key items evoked by formative evaluations:

- Does your course accomplish what it is supposed to?
- Is it of value?
- Does it impose minimum obstacles between the student and the learning?

Additional testing of ease of use ensures the interface doesn't get in the way of the learning:

- Are students able to understand instructions?

- Are students able to navigate through the course?
- Is it obvious where they are supposed to click first?
- Are the headings and buttons clearly labeled?
- Is it visually appealing and easy to read?

Evaluating the Content

One of the most powerful reasons to conduct formative evaluation is to determine that your message is clear.

In designing your plan for conducting formative evaluation, strive to develop test methods that will answer the following questions:

- Did I achieve my goals and objectives?
- What were the obstacles?
- Did the students benefit in the way I intended?
- What were the unintended consequences?

Unintended Consequences

"The true worth of a researcher lies in pursuing what he did not seek in his experiment as well as what he sought."

–Claude Bernard, (1813–1878) French physiologist

During formative evaluation, be alert to notice the unintended consequences, which can be more powerful than the intended consequences. Unintended consequences are when you had the intention of providing one service or message and users interpret and practice it in ways you didn't think of.

Examples of unintended consequences

1) During the early days of anti-drug films, young people learned new-and-improved ways to actually use drugs by watching movies that depicted behavior that was, ironically, designed to discourage drug use.

2) Under the Americans with Disabilities Act, television broadcasters are required to provide closed-captioned programming to make broadcasts accessible to people who are hearing impaired or deaf. As it turned out the appeal of closed-captions goes far beyond its intended use and is also enjoyed by people learning to read, people learning a second language, and people who

want to watch TV but don't want to disturb others with the sound in places like restaurants, bars and bedrooms.

The unintended consequences often have more significance than one might think.

When to Test

Formative evaluation should be conducted at every stage of development to ensure the course is achieving your goals and objectives. The earlier you begin formative evaluation the more likely you'll be to have an effective course in the end. Dr. Gerold Lesser, from the Harvard Graduate School of Education, recommends allocating ten to twenty percent of your project budget to formative evaluation.

What a Test Plan Should Include

- Define what you are testing for. Asking students to perform certain tasks can help you determine the effectiveness of the instruction. For example, ask the student to find a particular piece of information or respond to a particular question.
- Define when and how long each testing session will be.
- Define where you'll test and on what equipment.
- Define your target audience members, including how you plan to select them.
- Define your testing methods.
- Ideally, you should test people who offer diversity in terms of age, experience, geographical location, educational level, gender, income, race, and ethnicity. Who you select as your group of testers is important. Even within a narrow target audience people differ greatly.
- What type of computers will you use? What type of Internet connection and browser will you select?

Examples of commonly used testing methods:

- Asynchronous discussion–via listserv or bulletin boards.
- Focus groups–Have several people test the product and then observe their conversation about the product. You begin by giving them a set of general questions, and then observe and write down their behavior and comments. If they get off track you can redirect their conversation, but your input should be minimal.

- Observation–This is usually done by observing two or more people using the product. Record where they click, how long things take to finish, at what point they ask questions, etc. Ask questions upon completion.
- Online chat interview–Conduct one-on-one, or group discussions for direct feedback.
- Pre- and post-tests–Test people before and after viewing a unit of instruction to measure strengths and weaknesses, as well as whether the instruction is necessary.
- *Example*: Can the person write HTML tags before the class? Can they write the tags after the class? What were the problem areas?
- Surveys–Users can complete a survey to rate the instruction. These are useful only if you are going to do mass testing of the product and are looking for trends (*e.g.,* in things like appeal of the course, graphic design, or first impressions). In general, people do not give serious consideration or thought to answering survey questions.

Resources

Designing & Conducting Formative Evaluation
A Powerpoint presentation summarizing the Dick and Carey model of formative evaluation.
http://www.courses.psu.edu/trdev/trdev518_bow100/D_C10present/

Examples of Formative Evaluation Reports

This reports are much broader in scope than the reports required for this course but provide some examples of why people have conducted formative evaluation and how they write up their results.

A formative evaluation of distance education:
experiences of learners and instructors
http://node.on.ca/tfl/notes/herbeson.html

Formative Evaluation of a Multimedia Program (pdf file)
ttp://www.educ.sfu.ca/narstsite/conference/97conference/maor.pdf

## Summative Evaluation

Summative evaluation provides information on the product's efficacy ( it's ability to do what it was designed to do). For example, did the learners learn what they were supposed to learn after using the instructional module. In a sense, it lets the learner know "how they did," but more importantly, by looking at how the learner's did, it helps you know whether the product teaches what it is supposed to teach.

Summative evaluation is typically quantitative, using numeric scores or letter grades to assess learner achievement.

**So what is the difference between a Summative Evaluation and Learner Assessment?**

Although both might look at the same data, a Learner Assessment generally looks at how an individual learner performed on a learning task. It **assesses** a student's learning -- hence the name Learner Assessment. For example, you might assess an entire class of students, but you are assess them individually to see how each did.

A Summative Evaluation, on the other hand, looks at more than one learner's performance to see how well a group did on a learning task that utilized specific learning materials and methods. By looking at the group, the instructional designer can **evaluate the learning materials and learning process** -- hence the name Summative Evaluation. For example, here you may find that, as a group, all of the students did well on Section A of some instructional materials, but didn't do so well on Section B. That would indicate that the designer should go back and look at the design or delivery of Section B.

# Formative Evaluation

Coined by Michael Scriven (1967), the term "formative evaluation" is used for the type of evaluation that occurs during the developmental stage of the instructional design process (Seels & Glasgow, 1990). The instructional designer evaluates materials **during** the process of instructional development to determine where there are weaknesses in the instruction so that revisions can be made (Smith & Ragan, 1999). There are several variations of formative evaluation.

| Dick & Carey | Flagg |
| --- | --- |
| Clinical (one-to-one evaluation)<br>Small group evaluation<br>Field trial | Needs assessment<br>Pre-production formative evaluation<br>Production formative evaluation<br>Implementation formative evaluation |
| **Seels & Glasgow** | **Smith & Ragan** |
| Internal review<br>Tutorial and small-group tryouts<br>Operational tryout | Design reviews<br>Expert reviews<br>Learner validation<br>Ongoing evaluation |
| **Tessmer** | **Evaluation for Technology-based Instruction** |
| Expert review<br>One-on-one evaluation<br>Small group evaluation<br>Field test | Rapid prototype<br>Alpha testing<br>Beta testing |

## Summative Evaluation

Instructional designers may be involved in evaluating the effectiveness of instructional materials after the materials have been implemented into the instructional contexts for which they were designed (Smith & Ragan, 1999). However, Seels & Glasgow (1990) find that "summative evaluation is seldom carried out by the designers responsible for developing the original instruction and is not an integral part of our system model" (p. 199). Nevertheless, it is important to be familiar with the process. Like formative evaluation, there are several variations of summative evaluation.

**Explore the models below to learn more.**

| Dick & Carey | Smith & Ragan | Kirkpatrick |
|---|---|---|
| Expert judgement<br>Field trial | Determine goals of evaluation<br>Select indicators of success<br>Select orientation of evaluation<br>Select design of evaluation<br>Design or select evaluation measures<br>Collect data<br>Analyze data<br>Report results | Level 1 - Reaction<br>Level 2 - Learning<br>Level 3 - Transfer<br>Level 4 - Results<br>Note: In Kirkpatrick's four-level model, evaluation always begins with Level 1 and each successive level builds on information gathered from the previous level. |