Towards A Formative Evaluation Tool

Gwendolyn J. Ogle

Dissertation submitted to the Faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in
Curriculum and Instruction

John Burton, Chair
Katherine Cennamo
Glen Holmes
Mike Moore
Larry Weber

April 10, 2002
Blacksburg, Virginia

Keywords: Evaluation, Formative Evaluation, Summative Evaluation, Evaluation Models, Evaluation Approaches, Evaluation Tools

Copyright 2002, Gwendolyn J. Ogle
Towards A Formative Evaluation Tool

Gwendolyn J. Ogle

(ABSTRACT)

Evaluation is an integral part of instructional design. Formative evaluation, specifically, is a phase identified in many instructional design models and is recognized as an important step for program improvement and acceptance. Although evaluation has many models and approaches, very few deal specifically with formative evaluation. Further, no one set of guidelines has been found that provides a comprehensive set of procedures for planning and implementing a formative evaluation. Encapsulating such guidelines into a “tool” that automates the process was the author’s initial idea. The author’s intent in Chapter 2 was to find a model or checklist as a stepping off point for future formative evaluation tool development. In lieu of finding such a model, one was created (Chapter 3), pulling from several formative evaluation models and the author’s own experience. Chapter 3 also discusses the purpose behind developing a formative evaluation tool - to create an accessible, efficient, intuitive, and expedient way for instructional designers and developers to formatively evaluate their instruction or instructional materials. Chapter 4 focuses on the methodology selected to evaluate the tool, presented in prototype. Chapter 5 presents the results of the evaluation; comments received from the expert reviewers are presented and ideas for tool improvement are generated. Finally, the Appendices include the formative evaluation tool prototype as well as the documentation that accompanied the tool during its evaluation.

The initial idea behind this developmental dissertation was the creation of a formative evaluation tool. The focus of the dissertation itself, however, was on the justification for such a tool, and the literature behind the making of the model and consequently the tool. The result of this developmental dissertation was the prototype of an evaluation tool that with improvements and modifications is deemed promising by the experts who reviewed it. Although designed with formative evaluation in mind, it was generally agreed that this tool could be utilized for both formative and summative evaluation. The expert review was successful not because the tool was without fault, but because the review truly achieved its purpose – to identify areas of strength, weakness, and to suggest improvements.
DEDICATION

This dissertation is dedicated to my husband Todd for all the support, encouragement, and constancy he has given selflessly. We have journeyed together through much of our lives – the completion of our dissertations just being one phase of a very fulfilling life together. Todd, I thank you for the constant love and support you have given me, and for the pride and honor you display in all you do and all you are. I look forward to the next phase of our journey together – the arrival of our first child – and all future adventures life affords us.
ACKNOWLEDGEMENTS

I first and foremost want to thank my advisor and chairperson, Dr. John Burton, who is not only a respected professor, researcher, and advisor, but is also an incredible mentor. The confidence he instills in his students is a gift that will last long after graduation commences. The faith he has in his students and his willingness to fight the good fight on their behalf is unmatched.

I also want to thank my committee members, who each offered invaluable advice and guidance. Dr. Katherine Cennamo, your thoroughness and attention to detail have really given this dissertation a polish it otherwise might not have seen. Dr. Glen Holmes, your technical perspectives and hypothetical questions have always kept me honest and broadened my horizons. Dr. Mike Moore, your insightful questions have always led me down the right path. Dr. Larry Weber, your knowledge of evaluation and your suggestions and guidance are unparalleled. I have felt very supported by all of you during this process, and I thank you.

I want to thank Dr. Kevin Oliver for his insightful feedback. If only I had taken advantage of your knowledge sooner! Thank you for your time and suggestions.

I want to thank Jeff Seaton for his support as my NASA Graduate Student Researchers Program (GSRP) mentor. Without the GSRP position, I might never have generated this dissertation idea. Jeff, I have felt your support in the work I have done for NASA as well as the work I have done for this dissertation.

I want to thank Dr. Greg Sherman who was a very positive and energetic force early in my dissertation process. His encouragement was so positive and he was so free with ideas, one incredible thought after another. Dr. Sherman, your influence on my dissertation is strong, your influence on my knowledge of evaluation is great, and your positive comments will stay with me always.

I want to thank those who have given me opportunities during my time at Virginia Tech: Dr. Norm Dodl and Dr. John Burton and the Education Technology Lab, Dr. John Moore and the Faculty Development Institute, Dr. John Wenrich and Dr. Joy Colbert and the Institute for Connecting Science Research to the Classroom (ICSRC), Jeff Seaton and NASA, Dr. Holly Bender and Dr. Jared Danielson and the Biomedical Informatics Research Group (BIRG), and all others who provided the means by which I could learn, grow, and pay for this wonderful degree.

On a personal note, I want to thank my parents – Robert and Janet. I have really enjoyed becoming your friend, not just your daughter. I am the person I am because of the life I have had – I want you to know that I wouldn’t change a thing! I want to thank my husband’s parents – John and Linda – for their constant support in more ways than can be counted. They have raised a magnificent son and have taught me so much in the years since I have joined their family.

Last, I want to thank all other friends and family, including my brother Shawn and sister-in-law Kathy, all of whom have provided support and encouragement over the years. I hope to never take a single one of you for granted.
# TABLE OF CONTENTS

DEDICTION ........................................................................................................ III

ACKNOWLEDGEMENTS .................................................................................. IV

LIST OF FIGURES .......................................................................................... VIII

LIST OF TABLES .............................................................................................. IX

CHAPTER 1: INTRODUCTION ........................................................................ 1

CHAPTER 2: LITERATURE REVIEW ............................................................ 3

Evaluation vs. Research .................................................................................. 3

History of Evaluation ...................................................................................... 3

Evaluation in Instructional Design ............................................................... 5
  The CAI Design Model (Hannafin and Peck Design Model) ...................... 5
  The Dick and Carey Systems Approach Model for Designing Instruction ... 6

Problems and Perceptions of Evaluation .................................................... 8
  The Public Relations Problem ................................................................... 8
  Poorly-Planned and Poorly-Conducted Evaluations .................................. 9

Evaluation Types, Designs, Models, Approaches, and Tools .................... 11

Evaluation Types .......................................................................................... 11
  Planning Evaluation .................................................................................. 11
  Formative Evaluation .............................................................................. 12
    Expert Reviews ..................................................................................... 13
    One-To-One Evaluations ...................................................................... 13
    Small Group Evaluations ..................................................................... 14
    Field Trials ........................................................................................... 14
  Summative Evaluation ............................................................................. 14
  Predictive Evaluation .............................................................................. 14

Evaluation Approaches ............................................................................... 15
  Systems Analysis Approach .................................................................... 17
  Behavioral Objectives Approach ............................................................ 18
  Decision-Making Approach .................................................................... 18
  Goal-Free Evaluation Approach ............................................................ 19
  Art Criticism Approach ......................................................................... 20
  Professional Review Approach .............................................................. 21
  Quasi-Legal Approach .......................................................................... 21
  Case Study Approach ............................................................................ 21
CHAPTER 3: TOWARDS A FORMATIVE EVALUATION TOOL ..............43

Arguments For the Need of a New Formative Evaluation Model ............43

Arguments Against the Need of a New Formative Evaluation Model ............46

The Development of a New Formative Evaluation Model .........................47
  1. Clarify product/program goals and objectives ........................................49
  2. Define purpose of evaluation .................................................................50
  3. Determine evaluation questions ..............................................................50
  4. Plan the evaluation ..............................................................................50
  5. Select criteria, indicators, and data sources ..........................................51
  6. Develop instruments ...........................................................................52
  7. Submit evaluation proposal .................................................................53
  8. Make revisions to the evaluation proposal (if necessary) .......................54
  9. Schedule appointments for evaluation ..................................................54
 10. Conduct evaluation ............................................................................54
 11. Organize and analyze results ............................................................54
 12. Report results ....................................................................................54
 13. Revise instruction accordingly ............................................................55
14. Publish results ........................................................................................................55

Arguments For the Need of a Formative Evaluation Tool ...........................................56

CHAPTER 4: METHODOLOGY .................................................................................58

Evaluation Planning ....................................................................................................58
Tool Development .......................................................................................................61
Tool Evaluation ...........................................................................................................62
  Quantitative or Qualitative Methodology? ...............................................................63

CHAPTER 5: RESULTS AND CONCLUSIONS .........................................................64

Executive Summary ...................................................................................................64
Expert Reviewer’s Step-By-Step Comments ...............................................................65
Expert Reviewer’s Questionnaire Responses ...............................................................72
Meta-Evaluation .........................................................................................................77
Future Steps ...............................................................................................................78
Conclusion ................................................................................................................80

REFERENCES ..........................................................................................................81

APPENDIX A: FORMATIVE EVALUATION TOOL PROTOTYPE ..........85
APPENDIX B: EVALUATION TOOL FLOWCHART ............................................110
APPENDIX C: INTRODUCTION TO THE EXPERT REVIEWER ......................112
APPENDIX D: SOFTWARE CAPABILITIES .....................................................115
APPENDIX E: FORMATIVE EVALUATION TOOL EXPERT REVIEW INSTRUMENT .................................................................121
LIST OF FIGURES

Figure 2.1. The CAI Design Model (CDM) ................................................................. 5
Figure 2.2. The Dick and Carey Systems Approach Model for Designing Instruction..... 7
Figure 2.3. A Taxonomy of Major Evaluation Approaches ....................................... 16
Figure 2.4. The Hammond Model: A Structure for Evaluation .............................. 26
Figure 2.5. Stake’s Countenance Model .................................................................. 27
Figure 2.6. Management Model of Evaluation ......................................................... 29
Figure 2.7. The Sherman Model for Selecting Evaluation Procedures .................. 32
Figure 2.8. The Braden Formative Evaluation Model ............................................. 34
Figure 2.9. Evaluation Overview: Steps of a “Typical” Evaluation .......................... 36
Figure 2.10. The Evaluation Process: A Model for Classroom Teachers............... 37
Figure 2.11. The Evaluation Guide ......................................................................... 39
LIST OF TABLES

Table 2.1. Pre-service and in-service teachers’ responses to hearing the word evaluation 9
Table 2.2. Pace and Friedlander’s presentation of Alkin’s Evaluation Model .................. 28
Table 2.3. Flagg’s Four Phases of Formative Evaluation ................................................. 35
Table 4.4. Overview of methodology ............................................................................... 63
CHAPTER 1: INTRODUCTION

TenBrink (1974) defines evaluation as “the process of obtaining information and using it to form judgments which in turn are to be used in decision making.” (p. 8). M. C. Alkin (Alkin, 1973) defines evaluation as “the process of ascertaining the decision areas of concern, selecting appropriate information, and collecting and analyzing information in order to report summary data useful to decision makers in selecting among alternatives” (p. 150). Cronbach (1975) broadly defines evaluation as “the collection and use of information to make decisions about an educational program” (p. 244). Although some definitions focus on the attainment of objectives, the passing of judgment, or the process of scientific inquiry, deciding a course of action is a common thread in most definitions of evaluation (Stevens, Lawrenz, & Sharp, 1997).

The words evaluation, assessment, and measurement appear to be used interchangeably. Choppin (1985) attempts to “maintain the semantic distinctions” between these words and identifies the “ultimate objective” (p. 1747) of each. Measurement implies a numerical assignment of value, using instruments such as rulers, stopwatches, etc. “Measurement is rarely carried out for its own sake. It may be included in an assessment or evaluation, but is more to be regarded as a basic research procedure.” Choppin argues that the term assessment should be reserved for application to people, including grading, certifying, etc. Assessment may often utilize a test for measurement, but it rarely has “much in common with scientific measurement.” The term evaluation Choppin reserves for application to “abstract entities such as programs, curricula, and organizational variables” (p. 1748). Evaluation implies determining value and worth, and often involves making comparisons to other programs, curricula, or organizational schemes. Choppin writes, “Just as assessment may be characterized as a routine activity in which most educators will be involved, evaluation is an activity primarily for those engaged in research and development.”

Evaluation can serve several important purposes in the development of instruction, including (but not limited to) goal refinement, documentation, determination of impact, and program improvement (Hawkes, 2000). Hanson (1978) writes, “Good evaluation is central to the continued development of a profession.” (p. 97). Evaluation is becoming more mainstream as funding agencies require affirmation that their money is being spent on projects or programs that are “effective”. Stevens, Lawrenz, and Sharp (1997) quote from a 1992 report issued by the Government Accounting Office,

Over the next few years, the federal government will face powerful opposing pressures: the need on the one hand to reduce the federal deficit, and the demand, on the other, for a federal response to some potentially expensive domestic problems... (The need is for) program effectiveness evaluations which estimate the effects of federal programs using statistical analysis of outcomes...for groups of persons receiving program services compared with similar groups of non participants (p. x).

Even with growing support for evaluation, it still has yet to gain acceptance among the masses. Threats to this acceptance are: lack of evaluation knowledge, time
constraints, budgetary constraints, lack of personnel, or a negative predisposition to evaluation.

Evaluation should never be underestimated as a standardized, prepackaged process. Each evaluation will be individualized for the project or program it is addressing. Very few projects or programs will be exactly the same, or want to have the exact same questions answered. However, evaluation planning could be a standardized, prepackaged process, and the steps involved in conducting an evaluation could clearly be identified.

The following literature review (Chapter 2) contains information about evaluation and its many types, approaches, designs, models, and tools. In Chapter 3 an argument is proposed for the need of a new, more prescriptive model for formative evaluation, as well as an argument for the need of a formative evaluation tool. Chapter 4 discusses the methodology behind evaluating such a model and tool. Chapter 5 outlines the comments received from the expert evaluators and addresses the intended changes to and future evaluation of the tool.
CHAPTER 2: LITERATURE REVIEW

Evaluation vs. Research

The difference between evaluation and research may lie in their differing abilities to generalize (Worthen & Sanders, 1987) and their differing relationships to decision-making (Hanson, 1978). Research may bring about a better understanding of a topic or question, but “the information produced by research is not necessarily used to make decisions” (p. 98).

Good evaluation, like good science, utilizes measurement and observations that are accurate, reliable, and valid, gathers evidence systematically, and analyzes results objectively. When it comes to purposes, however, science seeks to discover regularities, generalizations, and laws, to test hypotheses, or to account for and explain the reasons or causes for what happens. Evaluation is more clearly pragmatic and, most important, explicitly seeks to produce judgments of value, worth, or merit, whereas science shuns such judgments. Perhaps in an effort to connect evaluation and science, the term evaluation research has come into prominence. (Pace & Friedlander, 1978, p. 9)

Rossi and Freeman (1993) define evaluation research as “the systematic application of social research procedures for assessing the conceptualization, design, implementation, and utility of … programs” (p. 5). Oetting (1978) writes, “The goal of evaluative research is to provide you with the particular information you need to make the best decisions” (p. 37).

History of Evaluation

Joseph Rice introduced systematic evaluation when he gave the same spelling test in several American schools for the sake of course improvement (Cronbach, 1975). His interest was in evaluating the curriculum of that time which included spelling drills, which he proved were worthless; this ultimately provoked curriculum revision.

There was a “measurement movement” (Pace & Friedlander, 1978) or “testing movement” (Cronbach, 1975) in education in the 1920’s and 1930’s, and evaluation “was defined as roughly synonymous with educational measurement” (Pace & Friedlander, 1978, p. 2; Worthen & Sanders, 1987). During this period administrators used achievement tests at their own discretion to assess the teacher and/or the school system. The content was considered “standard” and was not assessed during the 1920’s. The injudicious administrative testing of the 20’s and 30’s fell into disfavor and was replaced by more descriptive measures for assessing course adequacy. Also, the 1920’s belief in a “standard” curriculum gave way to the belief that the best learning experience would result from teachers and students planning their own content and curriculum (Cronbach, 1975). Pace and Friedlander discuss the definition of evaluation in the 1930’s.

Historically the term evaluation came into popular usage within education in the 1930’s to describe a territory that was broader than tests and measurements….The term evaluation, unlike the neutral word
measurement, denoted an explicit concern for purposes and values. (p. 4)

Pace and Friedlander also wrote,

The formalization of school and university accreditation procedures led to another definition: Evaluation is the exercise of professional judgment. That this definition has persisted is evident in many current evaluation practices where judgments are based on the opinions of experts, whether or not the data and criteria used in reaching those judgments are clear. (p. 2)

After 1930, tests were given to make judgments about individuals (Cronbach, 1975). Formalized school and university accreditation procedures lead to the view that evaluation was synonymous with professional judgment (Worthen & Sanders, 1987). The emphasis of test making was on the precise comparison of individuals with one another or with a standard. Cronbach wrote that precise scores for individuals are not always required (i.e. when evaluating a course), but that was the evaluation trend of the day.

Ralph Tyler’s work in the 1930’s and 1940’s helped define evaluation as something more than just testing students (Cronbach, 1975; Kinnaman, 1992; Pace & Friedlander, 1978). Tyler argued that determining whether or not a program met its stated objectives was a better way of evaluating that program than by mere student test scores alone. Prior to 1935, students were tested mostly on factual knowledge. Tyler noted that higher mental processes were not evoked by such factual testing and that the instruction that emphasizes such factual learning may be a barrier to more important educational outcomes. Tyler was able to demonstrate that tests could be developed that could measure more general educational outcomes (Cronbach, 1975).

Lee J. Cronbach later expanded on Ralph Tyler’s work by arguing that a program evaluation should look more at whether or not that program met its own desired outcomes and look less at how it compared with other programs (Cronbach, 1975). Cronbach and Scriven proposed an expansion of our concept of evaluation to include formative evaluation – collecting data during development for the improvement of the instruction (Dick, Carey, & Carey, 2001). Prior to the 1960s, instructional evaluation was usually summative (Tessmer, 1993). Formative evaluation did not become a formalized evaluation method until the late 1960s when Scriven gave the process its name. Formerly referred to as ‘tryout’, ‘developmental testing’, or other terms, formative evaluation became systematic when formal models of the process were developed and used with the programmed instruction movement (Tessmer, 1993).

During the 1960’s and 1970’s a new definition for evaluation emerged. Evaluation was defined as “a comparison of the effectiveness of an innovation with other existing products” (Dick et al., 2001, p. 283). The 1960’s brought about many developments in assessing educational programs, projects, and materials, not just assessing student performance (The Joint Committee On Standards for Educational Evaluation, 1981). Evaluation became the “process of identifying and collecting information to help decision makers choose among available alternatives” (Pace & Friedlander, 1978, p. 3).
Kinnaman (1992) feels that the biggest influence on educational evaluation was prompted by the Elementary and Secondary Education Act (ESEA) of 1965, a federal program requiring all schools receiving federal funding to submit annual evaluation reports. Kinnaman credits the growth of the assessment field to the ESEA’s major emphasis on program evaluation. “Federal, state, and local governments allocated millions of dollars each year to help evaluators address these questions… Evaluators responded with diversified designs, more elaborate procedures, better instruments, increased services, and expanded training. A great deal of time, money, and effort went into the improvement of educational evaluation” (The Joint Committee On Standards for Educational Evaluation, 1981, pp. 3-4).

From this came the push to provide “professional judgments” as to whether or not the evaluations were being conducted properly (The Joint Committee On Standards for Educational Evaluation, 1981). The late 1970’s and early 1980’s saw attempts at standardizing evaluation, a trend that arguably still exists today.

Evaluation in Instructional Design

Nearly all instructional design models contain an evaluation component, and some contain both formative and summative components. Formative evaluation is part of the instructional design process (Tessmer, 1993). “The formative evaluation component distinguishes the instructional design process from a philosophical or theoretical approach. Rather than speculating about the instructional effectiveness of your materials, you will be testing them with learners” (Dick et al., 2001, p. 302). Although the step of formative evaluation is placed after instructional development in most instructional design models, Braden (1992) recommends that formative evaluation be placed within every step of the instructional design process.

Below are two instructional design models selected for their prominence in the field of instructional design. Note that evaluation is a prominent component in both models.

*The CAI Design Model (Hannafin and Peck Design Model)*

![Figure 2.1. The CAI Design Model (CDM)](image-url)
The CAI (computer-assisted instruction) design model (Hannafin & Peck, 1988, p. 60), seen in Figure 2.1, combines its steps into phases. Although the model flows from left to right, the only way to move from one phase to the next is through “Evaluation and Revision”. This model has no set stopping point, and accentuates the importance of ongoing evaluation in design.

*The Dick and Carey Systems Approach Model for Designing Instruction*

Perhaps the most common model used for instructional design is the Dick and Carey Model (Dick et al., 2001, p.2-3). This model is based in the systems approach and favors the contemporary notion of viewing instruction as a whole, a system, in which each component is necessary for successful learning. The learners, instructor(s), instructional materials, and the learning environments are all components of this system. This “systems point of view” (p. 2) also views the instructional process itself as a system. A systems view approach integrates preparation, implementation, evaluation, and revision of instruction into one process. The Dick and Carey Model includes formative evaluation as a step within the design process; summative evaluations, although identified in the model, are conducted after the design process has been completed.

A systems approach to designing instruction requires that the learners interact with the instruction in meaningful and active ways. Practice worksheets with feedback, self-tests, and interactive instructional materials are just a few examples of instructional materials that could be designed using the systems approach. Instructional materials, instructional strategy, instructional objectives, and assessment instruments all need to be formatively evaluated; the systems approach maintains that each of these components interacts with the other to create, or not to create, effective learning.
Figure 2.2. The Dick and Carey Systems Approach Model for Designing Instruction
Problems and Perceptions of Evaluation

Evaluation has two key problems. First, it has a public relations problem. Second, it often isn’t conducted well, if at all. (Subsequently, one can see why it has a public relations problem.) Deborah Gwinner, a change management consultant at Anderson Consulting, lists 10 reasons why companies do incorrect training evaluations and consequently fail to get desired business or performance results (Abernathy, 1999, p. 9). The reasons, in order of process, are:

1. Lack of planning
2. Assumption that training is a cost rather than an asset
3. Lack of sponsorship
4. Lack of budget
5. Lack of appropriate resources (such as skilled and experienced people in the area in question)
6. Lack of understanding what is important to measure (results don’t match key performance measures)
7. Evaluation techniques that don’t capture human performance, only the performance of the trainers and training materials
8. Lack of valid measurements, resulting in false data reports
9. Lack of data collection
10. Lack of data analysis and summary

All of these reasons fall under the two key problems of poor public relations and improperly designed and conducted evaluations. These two problems are so interrelated that it is difficult to separate them. However, identifying and discussing the points of these two problems might prove beneficial.

The Public Relations Problem

Evaluation is often an unwelcome component of a project (Scriven, 1974), and is often viewed as an adversarial process (Stevens et al., 1997). The misconception exists that evaluation is something done to a project (not for a project) after-the-fact by an outside group who is placing judgment upon the effectiveness of the program (Cronbach, 1975; Kinnaman, 1992; Stevens et al., 1997; Valdez, 2000b). Valdez points out that although this often happens in practice, it is not the only or even the correct way of conducting an evaluation. Kinnaman writes,

Despite its promise for progress, program evaluation still has a tarnished reputation among educators. In fact, for many of us, just hearing the word evaluation evokes a negative response. It conveys the risk of failure and creates an atmosphere of vulnerability. It tends to conjure up fears, warranted or not, that someone’s position or program may be in jeopardy (p. 2).

TenBrink (1974) asked pre-service and in-service teachers, “What comes to your mind when you hear the word evaluation?” Below are their responses, with the words occurring most often listed first.
Table 2.1. Pre-service and in-service teachers’ responses to hearing the word evaluation

<table>
<thead>
<tr>
<th>Pre-Service Teachers</th>
<th>In-Service Teachers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tests</td>
<td>Tests</td>
</tr>
<tr>
<td>Grades</td>
<td>Measurement</td>
</tr>
<tr>
<td>Achievement</td>
<td>Grades</td>
</tr>
<tr>
<td>Unfair</td>
<td>Accountability</td>
</tr>
<tr>
<td>Judgment</td>
<td>Invasion of Privacy</td>
</tr>
</tbody>
</table>

The two negative responses “unfair” and “invasion of privacy” speak to the distorted perceptions of evaluation. Scriven (1974) also speaks to the unhealthy perceptions of evaluation when discussing the fear management has of goal-free evaluation (GFE) because of their perceived loss of control at what might be “discovered”. He writes, “The reactions by management to GFE have really brought out the extent to which evaluation has become or has come to seem a controllable item, an unhealthy situation” (p. 39).

Kinnaman (1992) writes “…program evaluation is not about criticism and failure; it’s about improving the quality of educational programs” (p. 5). Without bad publicity, evaluation might find its way into the agendas of designers, developers, evaluators, and decision-makers, leading to evaluation plans and results that benefit both stakeholders and the program being evaluated.

Poorly-Planned and Poorly-Conducted Evaluations

Worthen and Sanders (1987) note that poorly-planned and poorly-conducted evaluations aren’t in and of themselves surprising, but the frequency and significance of them within the field of educational evaluation is the real problem.

One problem plaguing evaluation is the common misperception that evaluation is a lengthy, time-consuming, expensive, or complex process (Abernathy, 1999; Oetting & Cole, 1978; Valdez, 2000b). Either due to these perceptions or to lack of knowledge, time and money are often not budgeted at all into the design process (Dick et al., 2001). Therefore, if an evaluation is conducted at all, it is often poorly planned and poorly conducted, resulting in information that is less than optimal for program improvement, program validation, or decision-making.

Furthermore, even if a designer is open to the notion of evaluation, they are usually unclear about what evaluation can do and should try to do (Cronbach, 1975). The evaluation is often brought in too late, after the program has been completed, rather than during development. Cronbach says, “Evaluation is a fundamental part of curriculum development, not an appendage. Its job is to collect facts the course developer can and will use to do a better job, and facts from which a deeper understanding of the educational process will emerge” (p. 256).

A second problem (resulting from poorly conducted evaluations) is the lack of transfer from evaluation results to program improvement or decision-making (Harpel,
Kinnaman (1992) points out if any formal evaluation is conducted at all it rarely amounts to more than an end-of-the-year formality. While better than nothing, “end-of-the-year progress reports don’t ensure a link between program evaluation and program improvement” (p. 2). Valdez (2000b) writes, All evaluation is wasted unless the results are used to improve project operations or interventions. This essential step, however, is frequently overlooked… Evaluation activities should be fully incorporated into the project management process (design, implement, evaluate, redesign…). Too often, insufficient time and resources are available for the redesign stage. You should schedule project activities to allow time for reviewing evaluation results and modifying project design after results become available and before the next iteration of the intervention begins. (p. 9)

A third problem (resulting from poorly planned evaluations) appears when the wrong questions are asked and/or the wrong data is gathered. The evaluator must know the needs of the decision-maker(s) to have an impact on the decision (Harpel, 1978). A common “blunder” is to look only at attitudinal data and disregard the measurable data regarding value (Abernathy, 1999).

A poor evaluation design is the cause of many failed evaluations, thus perpetuating biases against evaluation. At one extreme are those evaluators who are very clinical in their research efforts, planning elaborate assessments, but who have turned up only research and have not chosen to use it. “As a result, they have run clever studies aimed at answering interesting theoretical questions that have little or no bearing on practical decisions” (Oetting & Cole, 1978, p. 35). At the other extreme are those evaluators who feel “research” is beyond their scope and thus design equally useless evaluations (Oetting & Cole, 1978). Good evaluation, of course, would avoid both of these extremes. Another problem that stems from poor evaluation design is planning a perfectly legitimate evaluation that cannot be conducted because the project’s time and money constraints were not taken into consideration when planning the evaluation (Tessmer, 1993). Brown (1978, pp. 61-66) has created a checklist of things to consider for an evaluation to be more useful and successful:

1. Know who the decision makers are.
2. Identify key issues.
3. Identify sources of resistance and conflict.
4. Make the content of the evaluation relevant.
5. Get the staff involved in planning the evaluation.
6. Focus the evaluation on the problem rather than the personnel.
7. Understand the administrator’s role.
8. Clarify expectations.

A fourth problem can be found with the personnel conducting the evaluation. Many experts agree that the evaluator can be part of the problem (Cronbach, 1975; Pace & Friedlander, 1978; Worthen & Sanders, 1987). Worthen and Sanders write that evaluators
often conduct an evaluation using their preferred method, ignoring other paradigms that may be more relevant to the evaluation at hand.

First, evaluation is a requirement frequently and increasingly attached to the granting of money for new programs and the continuation of old ones. Second, evaluation has become big business. One consequence of the first fact is a lot of hasty and often frenetic efforts to do something to appease the questioners. One consequence of the second fact is the emergence for hire of professional evaluators, evaluation agencies, and evaluation ‘experts’ committed to their own individual theories and techniques.” (Pace & Friedlander, 1978, p. 1)

Evaluation Types, Designs, Models, Approaches, and Tools

Although the words “type,” “design,” “model,” “approach,” and “tool” are quite often used interchangeably, they are distinctly different aspects of evaluation. Worthen and Sanders (1987) note the lack of agreed upon language in educational evaluation when they write, “The semantic undergrowth in the field of evaluation could hardly be termed univocal; some clearing of redundant verbiage is clearly called for” (p. 145). This author has categorized and defined the evaluation components as follows: Evaluation types are phases of evaluation; the type selected depends upon the purpose of the evaluation (Braden, 1992). Evaluation designs provide the method by which the evaluation will be conducted. Evaluation models provide a graphical or textual overview of the steps involved in evaluation. Evaluation approaches are neither models nor designs; approaches are overarching theoretical frameworks that give the evaluator some foundations on which to plan and implement an evaluation. Evaluation tools are instruments that aid in the planning or implementation of an evaluation without serious time investments in research and without serious potential for errors or reinventing the wheel. A review of each of these follows.

Evaluation Types

There are several basic types of evaluation: planning evaluation (Stevens et al., 1997) (similar to the discrepancy analysis (Braden, 1992)), formative evaluation (Braden, 1992; Stevens et al., 1997; Worthen & Sanders, 1987), summative evaluation (Braden, 1992; Stevens et al., 1997), and predictive evaluation (Braden, 1992). In addition to these, Worthen and Sanders also identify informal and formal evaluation, as well as internal and external evaluation (Worthen & Sanders, 1987).

Planning Evaluation

A Planning Evaluation helps in determining a project’s goals, objectives, strategies, and timelines. “The product of the Planning Evaluation is a rich, context-laden description of a project, including its major goals and objectives, activities, participants and other major stakeholders, resources, timelines, locale, and intended accomplishments” (Stevens et al., 1997, p. 4). A type of evaluation that is similar to (or the same as) the Planning Evaluation is Discrepancy Analysis Evaluation. “We use
Discrepancy Analysis Evaluation (including needs assessment) for establishing parameters during goal setting and some forms of decision making. Any decision that requires a choice will benefit from this kind of evaluation of the things that might be chosen” (Braden, 1992, p. 15).

**Formative Evaluation**

Formative evaluation is a systematic and empirical process, although rarely a scientific one. “The term formative evaluation was introduced in 1967 by Michael Scriven and originally referred to ‘outcome evaluation of an intermediate stage in the development of the teaching instrument’ (p. 51)” (Flagg, 1990, p. 5). The term has developed and expanded over the years. Scriven later reflected on the formative role of evaluation when he wrote that evaluation feedback stays within the development loop and serves to improve the product (Scriven, 1973). The current definition of formative evaluation is “a judgment of the strengths and weaknesses of instruction in its developing stages, for purposes of revising the instruction to improve its effectiveness and appeal” (Tessmer, 1993, p. 11). Formative evaluation is part of the instructional design process, and can be a “cost-saving measure to economically 'debug' instruction and increase client satisfaction (Tessmer, 1993, p. 13). Tessmer also cites Nathenson and Henderson’s study (1980) which showed that “over the last 30 years, a number of empirical studies have shown that formatively evaluating instructional materials has resulted in revised instruction that produces statistically significant increases in student performance over the original, unevaluated versions of the instruction” (p. 13). This is the case even when materials are reviewed only once. Although the primary goals of most formative evaluations are quality control (Braden, 1992) and to improve instructional effectiveness (Stevens et al., 1997), there can be additional reasons for conducting a formative evaluation. “One reason is political: by involving instructors, administrators, or learners in the evaluation process they obtain ownership in the product, and with their ownership comes a greater chance for their acceptance and use of the final product” (Tessmer, 1993, p. 20).

Formative evaluations are usually conducted by the instructional designer(s) (Tessmer, 1993). Tessmer notes that instructional designers are often the only ones who know how to conduct a formative evaluation. Instructional designers are intimately familiar with the content and design of the materials, and may best be able to translate comments and suggestions back into program improvement. The most difficult problem facing formative evaluators is how to translate comments and suggestions into appropriate solutions.

The formative evaluation process also gives the evaluator the opportunity to evaluate the evaluation instruments, as well as the instruction. Learners can pinpoint confusing questions and tasks, as well as point out problems on an attitude survey. “If any section of the instrument does not yield data useful for pinpointing strengths and weaknesses and for revising instruction, then it should also be revised or eliminated” (Dick et al., 2001, p. 311).
Formative evaluation can be divided into several phases: needs assessment (Flagg, 1990), pre-production formative evaluation (Flagg, 1990), production formative evaluation (Flagg, 1990) (also known as progress evaluation (Stevens et al., 1997), and implementation formative evaluation (Flagg, 1990; Stevens et al., 1997). The needs assessment phase is often not considered part of formative evaluation, but rather as part of the planning evaluation. A pre-production evaluation collects information that guides decisions during the design phase. Production or progress evaluations are conducted to determine whether a project is meeting its goals. Implementation evaluations are conducted to determine whether the project is being conducted as planned, sometimes called field testing.

Either the instructional process (process evaluation) or the instructional product (product evaluation) can be the focus of a formative evaluation (Braden, 1992). Although there are minor variations to these, the main stages of formative evaluation are: Expert Reviews, One-To-One Evaluations, Small Group Evaluations, and Field Trials (Dick et al., 2001; Tessmer, 1993).

**Expert Reviews**

The expert review is a phase conducted very early in the formative evaluation process, while the materials are in their roughest stages. Tessmer (1993) notes that the expert review is an intrinsic evaluation of the instruction, meaning that the focus is on the instruction’s content accuracy or technical quality, rather than on learner performance or overall effectiveness. “Of all the stages of formative evaluation, expert review most directly reveals the types of revisions that should be made. This is because experts not only tell us what is wrong, they often tell us how they would fix it!” (p. 65).

There are several types of expert reviewers: subject-matter (content) expert reviewers; teacher, parent, or instructor expert reviewers; technical (production) expert reviewers; instructional design or learning specialist expert reviewers, subject sophisticates, and others. Dick (1975) notes that it is beneficial to have at least one of the reviewers be experienced in program design. This experience enables the reviewer to offer comments that would be most useful to the designer(s).

**One-To-One Evaluations**

The purpose of the one-to-one evaluations is “to identify and remove the most obvious errors in the instruction, and to obtain initial performance indications and reactions to the content by learners” (Dick et al., 2001, p. 286). During this phase of the formative evaluation the designer or developer works individually with three or more learners selected from the target population to gather both intrinsic and learning effects information as early as possible. The one-to-one evaluations are an interactive process. Dick et al (2001) note that the effectiveness of the process is greatly reduced when an evaluator simply hands the instruction to a learner and asks for their feedback. The evaluator should go through the materials with the learner and use questions and discussion techniques to gather useful information.
Small Group Evaluations

Small group evaluations answer use and implementation questions as well as effectiveness questions (Tessmer, 1993). The small group evaluation contains between eight and 20 learners that are representative of the target population (Dick et al., 2001). The evaluator interacts with learners as little as possible in an environment similar to the one in which the instruction will be used. Quantitative and descriptive information is summarized and analyzed with the intention of further refining the instruction.

Field Trials

Field trials answer use and implementation questions as well as effectiveness questions (Tessmer, 1993). The field trial (also called ‘beta test’ or ‘field test’) serves to see if changes made after the small group evaluation were effective and whether or not the instruction can be used effectively in its intended context (a situated evaluation). The polished instruction is used with approximately 30 learners from the target population, with the evaluator acting only as an observer. “Field tests can furnish more information on the ease of use of instruction than any other type of formative evaluation” (Tessmer, 1993, p. 138).

Summative Evaluation

Summative evaluations are usually conducted by outside evaluators, after a project or program’s completion, for purposes of accountability (Flagg, 1990). Determining effectiveness, costs verses benefits, and decision-making are three common goals of summative evaluation (Braden, 1992). Some purposes of summative evaluation (Stevens et al., 1997) are:

- To determine overall project success.
- To determine whether or not specific goals and objectives were achieved.
- To determine if and how participants benefited from the program.
- To determine which components were most (or least) effective.
- To determine any unanticipated outcomes.
- To determine cost vs. benefits.
- To communicate evaluation findings to stakeholders (teachers, participants, program designers and developers, funding agency, and superiors.)

Harpel (1978) identifies two distinct types of summative evaluation, each with its own purpose: research-oriented and management-oriented. This identifies the need for evaluations that improve and validate programs (research-oriented) as well as for evaluations that determine whether or not those programs did what was needed at what cost (management-oriented).

Predictive Evaluation
“Instructional designers do a lot of predictive evaluating when they organize the management of projects. Budget estimation, time estimation, guesses about materials that will be used and the skills that will be required are all imprecise activities” (Braden, 1992, p. 16). Predictive evaluation is used to improve ones chances of being correct when making such guesses.

**Evaluation Approaches**

Guba and Lincoln (as cited in Worthen & Sanders, 1987) note that one of the few things these diverse approaches have in common is the idea that evaluation should determine the worth of the educational program. House (1980) groups evaluation approaches into two categories: **objectivism** and **subjectivism**. Objectivism requires “scientifically objective” information, using an empirical approach external to the evaluator. Subjectivism claims its validity on, “an appeal to experience rather than to scientific method” (Worthen & Sanders, 1987, p. 252). Worthen and Sanders write, “The validity of a subjectivist evaluation depends on the relevance of the evaluator’s background and qualifications and the keenness of his perceptions. In this sense, the evaluation procedures are ‘internalized’…” (p. 46). The major evaluation approaches are identified in a taxonomy created by House (1980, p. 23) as seen in Figure 2.3. House recognizes earlier classifications by Stake (1976), Popham (1975), and Worthen and Sanders (1973) when introducing his taxonomy. Each approach is addressed in more detail in the pages following the taxonomy.
### Figure 2.3. A Taxonomy of Major Evaluation Approaches

<table>
<thead>
<tr>
<th>Model</th>
<th>Major Audiences or Reference Groups</th>
<th>Assumes Consensus On</th>
<th>Methodology</th>
<th>Outcome</th>
<th>Typical Questions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Systems analysis</td>
<td>Economists, managers</td>
<td>Goals, known cause and effect, quantified variables</td>
<td>PPBS, linear programming, planned variation, cost-benefit analysis</td>
<td>Efficiency</td>
<td>Are the expected effects achieved? Can the effects be achieved more economically? What are the <em>most</em> efficient programs?</td>
</tr>
<tr>
<td>Behavioral objectives</td>
<td>Managers, psychologists</td>
<td>Prespecified objectives, quantified outcome variables</td>
<td>Behavioral objectives, achievement tests</td>
<td>Productivity, accountability</td>
<td>Is the program achieving the objectives? Is the program producing?</td>
</tr>
<tr>
<td>Decision-making</td>
<td>Decision-makers, especially administrators</td>
<td>General goals, criteria</td>
<td>Surveys, questionnaires, interviews, natural variation</td>
<td>Effectiveness, quality control</td>
<td>Is the program effective? What parts are effective?</td>
</tr>
<tr>
<td>Goal-free</td>
<td>Consumers</td>
<td>Consequences, criteria</td>
<td>Bias control, logical analysis, modus operandi</td>
<td>Consumer choice, social utility</td>
<td>What are <em>all</em> the effects?</td>
</tr>
<tr>
<td>Art criticism</td>
<td>Connoisseurs, consumers</td>
<td>Critics, standards</td>
<td>Critical review</td>
<td>Improved standards, heightened awareness</td>
<td>Would a critic approve this program? Is the audience's appreciation increased?</td>
</tr>
<tr>
<td>Professional review</td>
<td>Professionals, public</td>
<td>Criteria, panel, procedures</td>
<td>Review by panel, self study</td>
<td>Professional acceptance</td>
<td>How would professionals rate this program?</td>
</tr>
<tr>
<td>Quasi-legal</td>
<td>Jury</td>
<td>Procedures and judges</td>
<td>Quasi-legal procedures</td>
<td>Resolution</td>
<td>What are the arguments for and against the program?</td>
</tr>
<tr>
<td>Case study</td>
<td>Client, practitioners</td>
<td>Negotiations, activities</td>
<td>Case studies, interviews, observations</td>
<td>Understanding diversity</td>
<td>What does the program look like to different people?</td>
</tr>
</tbody>
</table>
Systems Analysis Approach

This approach has also been called the Research Model, the Consumer-Oriented Approach (Worthen & Sanders, 1987), or the Educational Science Model (Pace & Friedlander, 1978). Worthen and Sanders do not consider the Systems Analysis Approach an evaluation approach at all. They believe that the narrow research focus of systems studies on establishing causal links qualifies them more as research than evaluation. However, their Consumer-Oriented Approach seemed in many ways to fit with the Systems Analysis Approach in its comparison of programs and in its cost-effectiveness factor. Scriven and Komoski are two names closely associated with the Consumer-Oriented Approach, a predominantly summative evaluation approach (Worthen & Sanders, 1987). Scriven developed a product checklist of 13 standards and continues to be an advocate of this approach.

“In the educational science models an ideal evaluation has an experimental research design in which the variables are so manipulated and controlled that one can explain the causes of the results” (Pace & Friedlander, 1978, p. 9). The Systems Analysis Approach assumes that program goals exist, and that these goals have been stated such that they can be measured. Astin and Panos are associated with this approach that is “intended to explain the causes of whatever results are observed” (p. 3). Astin and Panos (as cited in Pace & Friedlander) developed an approach whose goal was “to provide information on how alternative educational programs are likely to affect students’ performance” (p. 9-10). In this evaluation approach, the program is compared with another. Three components are studied by the evaluator to find this information:

- Inputs (the incoming qualities of the students)
- Operations (the characteristics of the program or set of experiences that are intended to promote the students’ progress toward the desired aims)
- Outputs (the impact of the program on the participants’ attainments of its objectives)

Determining the connections among these three components is the purpose of the Systems Analysis Approach. “Knowing how well students performed…does not by itself provide a sufficient base for evaluation. If that is all one knows, one is left with the assumption, but not the proof, that the results are due to exposure to the program…. So, one needs systematic information about what the individuals were like before they were exposed to the program” (Pace & Friedlander, , p. 10).

Pace, who does identify it as a valid evaluation approach, does find fault with the approach and writes that the main weakness of the approach is that it deals with only one criterion measure at a time.

Although many criterion measures can be used, each is seen only singly. Educational programs are always embedded in environments that are larger than their own boundaries, and they always have a number of outcomes or consequences that overlap, interact, and sometimes contradict one another. Looking at outcomes one at a time can therefore obscure
these interactions and miss the larger significance that may be present. (p. 14)

**Behavioral Objectives Approach**

This approach has also been called the Goal-Oriented Approach, the Goal-Based Approach, the Objectives-Oriented Approach (Worthen & Sanders, 1987), the Tylerian approach (Alkin & Ellett, 1985), and the Educational Psychology Model (Pace & Friedlander, 1978).

Its assumptions about education and evaluation are fairly clear. First of all, it assumes that education is intended to change the behavior of students….

Thus, the first step in the process of evaluation is to identify the general aims or goals of the instruction or educational program. The second step is to specify more concretely what these changes in behavior are – that is, to transform the program goals into measurable and observable pieces of behavior… Another main assumption is that particular conditions or tasks may be set up so that students may experience, practice, or demonstrate the intended behavior… The fourth step is to devise or select methods that will enable the evaluator to determine whether and to what extent the intended changes in students’ behavior have in fact occurred. Finally, one related what has happened to what was intended, drawing conclusions about how successful the program has been in achieving its objectives. To the extent that results are not as positive as one had hoped, a need for improvement is implied (Pace & Friedlander, 1978, p. 4-5).

Due to its straightforward procedure of letting the achievement of objectives determine success, failure, necessary improvements, or termination, this approach has been the predominant educational evaluation approach since the 1930s (Worthen & Sanders, 1987). Several names have been linked with the Behavioral Objectives Approach, including Tyler, Provus, Metfessel and Michael, Hammond, Popham, Taba, Bloom, and Talmage (Worthen & Sanders, 1987). Worthen and Sanders identify several “approaches”– what this author has referred to as “models”- within the Behavioral Objectives approach, such as the Tylerian Evaluation Approach (see Tyler’s Model on page 25), Metfessel and Michael’s Evaluation Paradigm (see the Metfessel and Michael Model on page 25), Hammond’s Evaluation Approach (see Hammond’s Model on page 26), and Provus’s Discrepancy Evaluation Model (on page 26).

**Decision-Making Approach**

This approach has also been called the Decision-Oriented Approach, the Management-Oriented Approach (Worthen & Sanders, 1987), and the Educational Decision Model (Pace & Friedlander, 1978). This approach has been associated with Alkin, Provus, and Stufflebeam. Marvin Alkin (1973) defines evaluation as “the process of ascertaining the decision areas of concern, selecting appropriate information, and collecting and analyzing information in order to report summary data useful to decision
makers in selecting among alternatives” (p. 150). Pace and Friedlander (1978, p. 6) note that this definition reflects the following beliefs or assumptions:

1. The decision maker determines the nature of the domain to be examined.
2. The evaluator’s role is that of advisor or servant to the decision maker.
3. Evaluation is concerned primarily with gathering information.
4. The information gathered should be appropriate to the needs of the decision maker.
5. What information is relevant depends on the situation and the kinds of decisions to be made.

Daniel Stufflebeam, another name widely recognized with the decision-making approach, believes that evaluation is performed in the service of decision-making (Stufflebeam, 1973). He defines evaluation as “the process of delineating, obtaining, and providing useful information for judging decision alternatives” (p. 129). The role of the evaluator in this approach is that of an information-gatherer; the decision makers are the real driving force. For examples of models adhering to this evaluation approach, see Alkin’s Model on page 28 and Stufflebeam’s CIPP Model on page 28.

**Goal-Free Evaluation Approach**

Goal-free evaluation is “the evaluation of actual effects against (typically) a profile of demonstrated needs” (Scriven, 1974, p. 35) or the goals of the consumer or the funding agency, instead of evaluating against the goals set forth in the program. Scriven (1974) coined the phrase “goal-free evaluation” after noticing side effects that sometimes had a more positive (or negative) effect than did the intended goals. Scriven argued that, by using an outside evaluator who was “blind” to the intended goals, that evaluator would be more open to assessing what the real effects of a program were. In fact, goals, as written in program proposals, are often full of rhetoric and current jargon. Albeit based on good intentions, evaluators need to evaluate what is really happening, not what was intended to happen. Scriven wrote that evaluating goals is just evaluating the proposal, not the actual product. (Further, he noted that the proposal itself was often being substituted as evidence for success!) Pace (1978) writes,

To be sure, most writers on evaluation note that one should also look for “side effects” or “unintended results”; but if measurement is precisely matched to objectives, how does one discover side effects? What Michael Scriven (1972) means when he talks about “goal-free” evaluation is not that evaluation is devoid of values but that it should not be limited or restricted to a specific set of stated goals…. The notion of relevant criteria, rather than explicit objectives, at least tends to avoid a restrictive influence, for it opens one’s mind to thinking about a wider range of observations that might be pertinent in judging the worth of a program. (p. 12)

The value of GFE is not noticing what is expected, but noticing the unexpected.

GFE can be used both for formative and summative evaluations. Although the formative evaluator is usually part of the design team, Scriven suggests hiring an outside
evaluator for GFE, so that the evaluator does not have familiarity with the goals, which could interfere with an objective evaluation. This outside evaluator would not replace the inside evaluator; on the contrary, implementing the results found by the external evaluator depends in large part on the internal evaluator. Further, since “a crucial function of good formative evaluation is to give the producer a preview of the summative evaluation” (Scriven, 1974, p. 35), if goals alone are used for the formative evaluation, undesirable side effects may surface in the summative evaluation that were not anticipated. Scriven writes of GFE in summative evaluation, “When summative time comes around, the intended effects had better be large enough to be obvious to the unaided (but expert) eye or, in general, they are not worth very much” (p. 38).

Often real goals differ substantially from the alleged goals. Scriven (1974) writes, “Since almost all projects either fall short of their goals or overachieve them, why waste time rating the goals; which are usually not what is achieved?” (p. 37). A benefit of using GFE is the flexibility for shifting goals midway in a project; the evaluation has not been designed around those goals.

Scriven (1974) identifies arguments against GFE (both true and false). They are:

- The evaluator would substitute his or her own goals for those of the project (False)
- The evaluator would infer the goals of the project (False – Simply control the amount of data getting to the evaluator)
- The evaluator might miss (intended) results (True)
- Dismissing goals as an unimportant component to evaluation might result in no (or poor) goal formulation in the program planning process (False – These are necessarily unrelated)
- The evaluator can’t test for all possible effects (False – Both goal-based and goal-free evaluators should be looking for side effects.)
- GFE will be seen as a threat (True)

Scriven (1974) makes an analogous comparison between GFE and double blind medical studies. Keeping the evaluator unaware of the program’s goals keeps the evaluator objective and open to legitimate side effects, just as keeping the medical researcher unaware of who received the placebo pill keeps the medical researcher objective and open to legitimate side effects.

**Art Criticism Approach**

This approach has also been called the Expert Approach or the Expertise-Oriented Approach (Worthen & Sanders, 1987). Worthen and Sanders combine the Art Criticism Approach and the Professional Review Approach into one approach called the Expertise-Oriented Approach. Worthen and Sanders identify theater critics, art critics, and literary critics as educational connoisseurs and critics. These experts offer direct and efficient application of expertise to that which is being evaluated.
Professional Review Approach

This approach has also been called the Accreditation Approach, the Expertise-Oriented Approach (Worthen & Sanders, 1987), and the Professional Judgment Approach. As mentioned above, Worthen and Sanders combine the Art Criticism Approach and the Professional Review Approach into one approach called the Expertise-Oriented Approach. Worthen and Sanders identify formal and informal review systems, as well as ad hoc panel and ad hoc individual reviews. This approach focuses on the subjective professional expertise of experts for the purpose of passing judgment.

Quasi-Legal Approach

This approach has also been called the Adversary Approach, the Judicial Approach, or the Adversary-Oriented Approach (Worthen & Sanders, 1987). Several names associated with this approach are Rice, Wolf, Owens, Levine, and Kourilsky. The basis of this approach is to balance bias, rather than reduce it, by presenting both the positive and negative views in the evaluation (Worthen & Sanders, 1987).

Case Study Approach

This approach has also been called the Transaction Approach, the Naturalistic and Participant-Oriented Approach (Worthen & Sanders, 1987), or the Educational Change Model (Pace & Friedlander, 1978). Pace and Friedlander also refer to this approach as the “case study”, “self-study”, or “learning” model. By “case study” they mean “an individual self-evaluation in which the student is both the recorder and reflector of his or her experience” (p. 14). The evaluator is an internal, not an external, evaluator. This characteristic closely intertwines evaluation and learning (Pace & Friedlander, 1978). In this approach, evaluation has a strong possibility of being a learning activity.

Those who see evaluation results from someone else’s evaluation could be prone to dismissing those results. Results, however, from one’s own evaluation would not be second-guessed so easily. Believing in the validity of the results is a good step towards positive change. “Evaluation should be conducted in such a way that it builds confidence in the results and fosters readiness to change. Participation, making evaluation a genuine group enterprise, is one means of increasing the likelihood that the results will be put to good use” (Pace & Friedlander, 1978, p. 15).

Pace and Friedlander (1978, p. 16) also identified several ingredients for a successful self-study.

- The topic or focus of the study must be recognized as important.
- The people most directly concerned about it and most likely to be affected by the results must be active in carrying out the study.
- Maintaining active participation requires continuous and open communication, discussion, and debate.
- The interpretation of results must reflect consensus.
Plans must be made and followed for translating the conclusions into appropriate actions.

This approach “emphasizes the uses of evaluation for individual learning and of institutional self-study for institutional improvements. This school of thought requires the personal involvement of those people who will participate in the change: students and faculty members, as well as decision makers, rather than only researchers” (Pace & Friedlander, 1978, p. 3). Several names are associated with this approach, including Stake, Patton, Guba and Lincoln, Rippey, MacDonald, and Parlett and Hamilton (Worthen & Sanders, 1987). Worthen and Sanders identify several “approaches”- what this author has referred to as “models”- within the Case Study Approach: Stake’s Countenance Model, Transactional Evaluation, Illuminative Evaluation, Democratic Evaluation, Responsive Evaluation, and Naturalistic Evaluation. For further information, see Stake’s Countenance Model on page 27.

Evaluation Designs

Research designs have been identified as qualitative and quantitative (Valdez, 2000b; Worthen & Sanders, 1987). Worthen and Sanders see these two methods as complementary in educational evaluation and note an increased integration of these methods within educational evaluation. Valdez’s (2000b) organization of the designs follows:

Qualitative
- Qualitative interviews
- Focus groups
- Systematic observation

Quantitative
- Non-experimental (characterized by the absence of a control or comparison group – can make no generalizable claims about gains)
  - Posttest-only design
    - Pretest-posttest design
  - Quasi-experimental (subjects aren’t randomly assigned)
    - Time series design
    - Nonequivalent control group design
  - Experimental designs (random assignment of subjects)
    - Pretest-posttest control group design
    - Multiple intervention design

Hawkes (2000) identifies evaluation as having three main objectives: description, explanation, and validation. He outlines primarily qualitative methods for addressing description and explanation, and more quantitative methods for addressing validation.

Description and Explanation (Qualitative)
- Participant observation
- Interviews
- Focus groups
- Document review
• Participant logs or journals
• Portfolios
• Expert Review
• Case study

Validation (Quantitative)
• Surveys/questionnaires
• Standardized tests
• Experiments

Fitz-Gibbons and Morris (1987) write, "Without design, the information you present will in most cases be haunted by the possibility of reinterpretation. Information from a well designed study is hard to refute…" (p. 11). They identify seven primarily quantitative designs:
1. True control group, pretest-posttest design
2. True control group, posttest only design
3. Non-equivalent control group, pretest-posttest design
4. Single group time series design
5. Time series with non-equivalent control group design
6. Before and after design
7. Analysis of Variance (ANOVA)
Evaluation Models

Worthen and Sanders wrote in 1987 that during the previous two decades, more than 50 different evaluation models had been developed. M. C. Alkin (1985) wrote, Evaluation models either describe what evaluators do or prescribe what they should do. Generally, evaluators are concerned with determining the value or current status of objects or states of affairs. The term ‘model’ is used in two general ways. (a) A prescriptive model, the most common type, is a set of rules, prescriptions, prohibitions, and guiding frameworks which specify what a good or proper evaluation is and how evaluation should be carried out. Such models serve as exemplars. (b) A descriptive model is a set of statements and generalizations which describes, predicts, or explains evaluation activities. Such models are designed to offer an empirical theory.

The importance of studying evaluation models is shown in a number of ways. For example, understanding the various evaluation models provides insights and a framework for conducting evaluations in a more defensible way. Prescriptive models provide consistent frameworks and strategies for performing evaluations, and descriptive models present a range of validated possibilities for conducting evaluation.” (pp. 1760-1761)

Although the intent was to locate a comprehensive, prescriptive formative evaluation model, one will see that there is overlap between “formative” and “summative” models. Several models are so theoretical that they apply to either category. Consequently, the models that follow are not grouped by formative or summative procedures (although a few do fit one category better than another.) Additionally, no effort was made to group the models by their prescriptive or descriptive nature. Instead, what follows is a detailed list of the most prominent and relevant models in the field of evaluation.

Donald J. Kirkpatrick’s Level 4 Model

- Reaction
- Learning
- Behavior
- Results

Kirkpatrick first published a series of articles in 1959, describing this four-stage evaluation model. Abernathy (1999) refers to Kirkpatrick’s model as the “supermodel” of training evaluation. Abernathy quotes Paul Bernthal, manager of research at Development Dimensions International, “Kirkpatrick’s classic model has weathered well. But it has also limited our thinking regarding evaluation and possibly hindered our ability to conduct meaningful evaluations” (The Supermodel section, para. 3).
Too often, trainers jump feet first into using the model without taking the
time to assess their needs and resources, or to determine how they’ll apply
the results. When they regard the four-level approach as a universal
framework for all evaluations, they tend not to examine whether the
approach itself is shaping their questions and their results. The simplicity
and common sense of Kirkpatrick’s model imply that conducting an
evaluation is a standardized, prepackaged process (Abernathy, 1999).

John W. Humphrey, chairman of The Forum Corporation, is quoted as being
dissatisfied with the Kirkpatrick model. “Kirkpatrick’s four levels speak to a finite
intervention. For example, you can train salespeople and, at the end of the day, measure
whether or not it happened” (Abernathy, 1999, New Directions section, para. 6).
Humphrey points out that this does not take continuous learning into account.

**Tyler’s Model**

Worthen and Sanders (1987, p. 63) outline the seven steps of the “Tylerian
Evaluation Approach”:
1. Establish broad goals or objectives
2. Classify the goals or objectives
3. Define objectives in behavioral terms
4. Find situations in which achievement of objectives could be shown
5. Develop or select measurement techniques
6. Collect performance data
7. Compare performance data with behaviorally stated objectives

“Evaluation was viewed as a cycle that involved not only clarifying and measuring
objectives but adapting teaching methods and materials to make success more likely”
(Pace & Friedlander, 1978, p. 4).

Ralph W. Tyler has had an enormous impact on the field of educational
evaluation. He is often referred to as the Father of Evaluation. Many books are dedicated
to this man, and much of the work that has come after him has been based on his works.
His career spanned many decades, and his presence it still impacting the field of
evaluation.

**The Metfessel and Michael Model**

Worthen and Sanders (1987, p. 65-66) identify the eight steps in what is actually
called the “Metfessel and Michael Evaluation Paradigm”:
1. Involve the total school community as facilitators of program evaluation.
2. Formulate cohesive model of goals and specific objectives.
3. Translate specific objectives into a communicable form applicable to facilitating
   learning in the school environment.
4. Select or construct instruments to furnish measures allowing inferences about
   program effectiveness.
5. Carry out periodic observations using content-valid tests, scales, and other behavioral measures.
6. Analyze data using appropriate statistical methods.
7. Interpret the data using standards of desired levels of performance over all measures.
8. Develop recommendations for the further implementation, modification, and revision of broad goals and specific objectives.

This model was heavily influenced by the work of Tyler. Its major contribution was in expanding the possibilities regarding alternative instruments.

Hammond’s Model

Hammond (1973, p. 158) developed a three-dimensional cube (see Figure 2.4) that he called a “structure for evaluation”. The cube was designed to help the evaluator search for factors contributing to the success and failure of educational activities (Worthen & Sanders, 1987).

![Figure 2.4. The Hammond Model: A Structure for Evaluation](image)

Provus’s Discrepancy Model

As a program, product, or activity is being developed, Provus thought of it as going through four developmental stages, with a fifth optional step (Worthen & Sanders, 1987, p. 68).
1. Definition
2. Installation
3. Process (interim products)
4. Product
5. Cost-benefit (optional)

Provus (1973, p. 172) writes, “Evaluation is the process of (a) agreeing upon program standards, (b) determining whether a discrepancy exists between some aspect of the program and the standards governing that aspect of the program, and (c) using discrepancy information to identify the weaknesses of the program.”

**Stake’s Countenance Model**

Stake (1973, p. 113) identifies three elements in an evaluation program that can be found in his model: antecedents (the conditions that exist before intervention), transactions (the events or experiences that constitute the program), and outcomes (the effects of the program). Stake asserts that the two basics of evaluation are description and judgment.

Stake’s discussion has two special aspects… first, the distinction between intents and observation (what was supposed to be done and what was really done) and, second, the difference between standards and judgments about what effects occurred….Stake’s model can be comparative (Is A better than B?) or not (Does A do what it was meant to do?) (Pace & Friedlander, 1978, p. 6).

![Figure 2.5. Stake’s Countenance Model](image)

According to Worthen and Sanders (1987, p. 130), the evaluator would use this framework for the following steps:

1. Provide background, justification, and description of the program rationale (including its need)
2. List intended antecedents (inputs, resources, existing conditions), transactions (activities, processes), and outcomes
3. Record observed antecedents, transactions, and outcomes (including observations of unintended features of each)
4. Explicitly state the standards (criteria, expectations, performance of comparable programs) for judging program antecedents, transactions, and outcomes
5. Record judgments made about the antecedent conditions, transactions, and outcomes

Alkin’s Model

Alkin (1973) identifies five areas of evaluation that attempt to provide information to satisfy unique decision categories:
1. Systems assessment (Similar to context evaluation in the CIPP Model).
2. Program planning
3. Program implementation
4. Program improvement (Similar to process evaluation in the CIPP Model).
5. Program certification (Similar to product evaluation in the CIPP Model).

Pace and Friedlander (1978, p. 7) organize Alkin’s Model this way:

Table 2.2. Pace and Friedlander’s presentation of Alkin’s Evaluation Model

<table>
<thead>
<tr>
<th>Decision Area</th>
<th>Type of Evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Problem selection</td>
<td>Needs assessment</td>
</tr>
<tr>
<td>Program selection</td>
<td>Program planning</td>
</tr>
<tr>
<td>Program operationalization</td>
<td>Implementation evaluation</td>
</tr>
<tr>
<td>Program improvement</td>
<td>Progress evaluation</td>
</tr>
<tr>
<td>Program certification</td>
<td>Outcome evaluation</td>
</tr>
</tbody>
</table>

Stufflebeam’s CIPP Model

The CIPP Model (Context, Input, Process, and Product) can be used for both formative and summative evaluation. “Perhaps the most significant characteristic of CIPP is that it makes provision for holistic evaluation. Its elements are systems oriented, structured to accommodate universal evaluation needs” (Braden, 1992, p. 14). Braden also notes the rarity of an evaluation model that offers process evaluation, as this one does.

- **Context evaluation**, to serve planning decisions (Worthen & Sanders, 1987) – “is intended to describe the status or context or setting so as to identify the unmet needs, potential opportunities, problems, or program objectives that will be evaluated” (Pace & Friedlander, 1978, p. 8).
- **Input evaluation**, to serve structuring decisions (Worthen & Sanders, 1987) – “the evaluator provides information to help the decision maker select procedures and resources for designing or choosing appropriate methods and materials” (Pace & Friedlander, 1978, p. 8).
• **Process evaluation**, to serve *implementing decisions* (Worthen & Sanders, 1987) – “making sure that the program is going as intended, identifying defects or strengths in the procedures” (Pace & Friedlander, 1978, p. 9).

• **Product evaluation**, to serve *recycling decisions* (Worthen & Sanders, 1987) – “a combination of Alkin’s progress and outcome evaluation stages” (Pace & Friedlander, 1978, p. 9) that serves to judge program attainments.

The CIPP model deals with products or outcomes not only at the conclusion of the program but also at various points during the program. Outcomes are then related to objectives; differences are noted between expected and actual results; and the decision maker decides to continue, terminate, or modify the program (Pace & Friedlander, 1978, p. 9).

**Management Model of Evaluation**

![Figure 2.6. Management Model of Evaluation](image)

The Management Model of Evaluation (Harpel, 1978, p. 21) is cyclical in nature. One can enter the process at any point, although each stage builds on another. Harpel notes that the “iterative nature of the model is self-correcting” (p. 21), because eventually all steps will be addressed.

A manager’s evaluation takes into account not only the effectiveness of a program, but also the efficiency of a program (Harpel, 1978). Managers must take into account the resources available and whether or not those resources were used to their best advantage while also keeping in mind the goals of the parent institution or organization. Results of a management evaluation are primarily reviewed by external decision-makers rather than internal program staff. This is due to the heavy emphasis on resource allocation, since external decision-makers (administration, budget boards, etc.) will use these findings for future funding.
Walker’s Evaluation Design Model

Neff Walker (as cited in Valdez, 2000b) created an evaluation design for the SUCCEED program that can serve as a model for developing an evaluation design and process. Valdez articulates Walker’s evaluation model:

Step 1: Define the Purpose of the Evaluation
Action steps for investigators:
• Decide on the purpose(s) of the project evaluation.
• Determine what primary questions need to be answered.
• Determine in what order these questions should be answered.
• Decide who will be the primary audience(s) for the evaluation results.
• Determine how the results will be used.

Step 2: Clarify Project Objectives
Action steps for investigators:
• Review and, if necessary, revise existing project objectives.
• Ensure that appropriate impact, outcome, and process objectives have been specified.

Step 3: Create a Model of Change
Action steps for investigators:
• Develop a model of change for the project, making it as specific and complete as possible.
• Review each of the assumptions (lines) in the model. Using the criteria presented above, identify the priority assumptions to be addressed through the project evaluation.

Step 4: Select Criteria and Indicators
Step 5: Identify Data Sources and Define How Often Indicators Will Be Measured
Step 6: Design Evaluation Research
Action steps for investigators:
• Design evaluation research studies for key questions.
• Carry out studies and report the results.
• Where appropriate, use the results to improve project interventions or operations.

Step 7: Monitor and Evaluate
Action steps for investigators:
• Implement your plan for project evaluation. Ensure that you have a working management information system for the project as well as sufficient money, time, and people to carry out planned evaluation activities, and that regular meetings are held to review and use the evaluation results.

Step 8: Use and Report Evaluation Results
Action steps for investigators:
• Ensure that your project timeline includes adequate time for the interim analysis and review of evaluation results, and for modification of project interventions.
• Publish and present your evaluation findings in appropriate journals and at relevant conferences.
This evaluation design model is referred to as a “primer” and is an attempt to clearly outline the actions evaluators should take. Walker had a great sense for reducing the evaluation process down to its most basic components. The model is useful as a basis. However, it does lack, as most models do, the specifics of how to operationalize many of the steps, even if specific “action steps for investigators” are outlined.

**Tips for Conducting an Evaluation**

Stevens et al. (1997) overview the process of conducting an evaluation in the *User-Friendly Handbook for Project Evaluation*. The steps or “tips” are summarized into five main points (p. 29).

1. Develop Evaluation Questions
   - Clarify goals and objectives of the evaluation.
   - Identify and involve key stakeholders and audiences.
   - Describe the intervention to be evaluated.
   - Formulate potential evaluation questions of interest to all stakeholders and audiences.
   - Determine resources available.
   - Prioritize and eliminate questions.

2. Match Questions with Appropriate Information-Gathering Techniques
   - Select a general methodological approach.
   - Determine what sources of data would provide the information needed.
   - Select data collection techniques that would gather the desired information from the identified sources.

3. Collect Data
   - Obtain the necessary clearances and permission.
   - Consider the needs and sensitivities of the respondents.
   - Make sure data collectors are adequately trained and will operate in an objective, unbiased manner.
   - Cause as little disruption as possible to the ongoing effort.

4. Analyze Data
   - Check raw data and prepare data for analysis.
   - Conduct initial analysis based on the evaluation plan.
   - Conduct additional analysis based on the initial results.
   - Integrate and synthesize findings.

5. Provide Information to Interested Audiences
   - Provide information to the targeted audiences.
   - Deliver reports and other presentation in time to be useful.
   - Customize reports and other presentations.
Sherman Model for Selecting Evaluation Procedures

Figure 2.7. The Sherman Model for Selecting Evaluation Procedures

The Sherman Model for Selecting Evaluation Procedures (class notes, Spring, 2000) was developed by Dr. Greg Sherman to aid in selecting the type of evaluation one needs (i.e. program vs. product, formative vs. summative) and consequently what procedures one would need to engage in to meet those needs.

Tessmer's Formative Evaluation Planning Process

Tessmer (1993, pp. 38-43) outlines 11 steps that are necessary when planning an evaluation. Although not graphical, the steps are prescriptive in nature and this model constitutes one of only a few formative evaluation models.

1. Determine the goals of the instruction
2. Determine the resources for and constraints upon the evaluation
3. Obtain a complete task analysis of the instruction
4. Describe the learning environment
5. Determine the media characteristics of the instruction
6. Outline the information sought from the evaluation
7. Choose parts of the instruction for evaluation
8. Select the stages and subjects for the evaluation
9. Select the data gathering methods and tools for each stage
10. Plan the report(s) of the evaluation
11. Implement the evaluation

In addition to the steps outlined above, Tessmer (1993, pp. 26-37) also identifies questions that one must ask when planning a formative evaluation:

- What are my evaluation goals?
  - To improve the effectiveness of the instruction
    - Identify deficiencies in learning effectiveness
    - Locate ease of use problems
    - Evaluate the efficiency of the instruction
    - Analyze instructional strengths
  - To gain acceptance of the instruction
  - To demonstrate the utility of formative evaluation
- Who will be making evaluation decisions?
  - Who will see the evaluation data?
  - Who will make decisions on what is to be revised?
  - Who is responsible for revising it?
  - What is the best way to present the information to those who make the decisions?
- What will be learned?
  - Identify the terminal objective(s), sub-objectives, and entry skills (task analysis)
- What is the setting for the learning?
  - What are the user requirements?
  - In what learning environment(s) will they learn?
  - What are the physical characteristics of these learning environments?
  - What is the pattern of use of instruction?
  - What kind of support environment is required?
- What media formats will be used?
- What are the evaluation resources and constraints?
  - When is the project due to be finished?
  - What money and time can be encumbered for the evaluation?
  - What personnel and resources are available for the evaluation?
  - What is the level of expertise of the evaluators?
  - How revisable is the instruction?
  - When must the evaluation data be submitted in order for changes to be feasible?
- What should be assessed, learning or work environment?

Tessmer (1993, p. 20) identifies eight circumstances when a thorough formative evaluation effort is particularly warranted:

- When the designer is relatively new to the practice of instructional design
- When the content is new to the designer or team
- When the technology is new to the design team
- When the learner population is new to the design team
- When unfamiliar or experimental instructional strategies are part of the instruction
When accurate task performance from the instruction is critically important
When materials will be disseminated in large quantity
When the chances for revisions or newer versions of the instruction after its release is slim

Braden Formative Evaluation Model

Braden (1992) feels that a major weakness of nearly all instructional design models is that they dedicate only one step to formative evaluation at the end of the process, if they include an evaluation step at all. Braden argues that formative evaluation needs to be conducted at each step of the process, thereby evaluating both the process and the product. “Formative evaluation of the deliverable of any given ISD process step is product evaluation initially. However, as soon as a flaw in the product is determined, the procedure shifts to an evaluation of the process which permitted the flaw to occur” (p. 14). He calls this front-to-back evaluation, which enables a system to be truly self-correcting. Braden notes a movement towards a more expanded role of formative evaluation in instructional design, and predicts an increased acceptance of front-to-back formative evaluation. While Braden acknowledges that formative evaluation is often conducted throughout the instructional design process in practice, he felt a model would make the process more systematic.

The model proposed by Braden (1992, p. 77) is an elaboration of the Dick and Carey Instructional Design Model (Dick et al., 2001, p. 2-3). This model (see Figure 2.8) was selected for its recognition value. Embedded in each step is a circle containing the letters F. E., which stands for formative evaluation, and denotes that formative evaluation is the final “sub-step” within each step of the model. The key to success of a front-to-back evaluation model is that proceeding to the next step is contingent upon having formatively evaluated the deliverables at the previous step (and having done any necessary revisions).

Figure 2.8. The Braden Formative Evaluation Model

The Braden Model differs from the Dick and Carey Model in three ways: the addition of the F. E. “nodes”, the addition of an “organize management” box, and the
re-naming of the formative evaluation step to “Conduct Prototype Test.” Braden points out that one could add formative evaluation to each step of any instructional design model, not just the Dick and Carey model. He writes that most instructional design models tend to be very similar to each other, consisting of the same 10 or 11 steps. The following two points were identified as requirements for applying the front-to-back formative evaluation component to an existing instructional design model (Braden, 1992, p. 12):

1. The model should be linear.
2. The model should be a prescriptive model for, not a descriptive model of.

The Braden Model works by asking questions of consistency and adequacy at each step of the process. However, Braden notes that as with all evaluations, the exact methods depend on what information one wishes to collect.

*Four Phases of Formative Evaluation*

Flagg (1990, p. 4) introduced the four phases of formative evaluation (see Table 2.3). Although this is not a model per se, it did contribute greatly to a broader definition of formative evaluation. “Taken together these phases represent a front-to-back implementation of formative evaluation methods” (Braden, 1992, p. 4).

*Table 2.3. Flagg’s Four Phases of Formative Evaluation*

<table>
<thead>
<tr>
<th>Phases of Program Development</th>
<th>Phases of Evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phase 1: Planning</td>
<td>Needs assessment</td>
</tr>
<tr>
<td>Phase 2: Design</td>
<td>Pre-production formative evaluation</td>
</tr>
<tr>
<td>Phase 3: Production</td>
<td>Production formative evaluation</td>
</tr>
<tr>
<td>Phase 4: Implementation</td>
<td>Implementation formative evaluation</td>
</tr>
<tr>
<td></td>
<td>Summative evaluation</td>
</tr>
</tbody>
</table>

*Evaluation Overview: Steps of a “Typical” Evaluation*

Quinones, Kirshstein, & Loy (1998, p. 6) note that although each and every evaluation is slightly different, there are steps that are generally followed in any evaluation. The following is an overview of the steps in a “typical” evaluation (see Figure 2.9).
**Step 1**
Get an Overview of the Program

**Step 2**
Determine Why You Are Evaluating

**Step 3**
Determine What You Need to Know and Formulate Research Questions

**Step 4**
Figure Out What Information You Need to Answer Questions

**Step 5**
Design the Evaluation

**Step 6**
Collect Information/Data

**Step 7**
Analyze Information

**Step 8**
Formulate Conclusions

**Step 9**
Communicate Results

**Step 10**
Use Results to Modify Program

*Figure 2.9. Evaluation Overview: Steps of a “Typical” Evaluation*
TenBrink (1974) writes that the value of having a model of the evaluation process is that it provides an ideal, even though one might deviate from the specific steps. This 10-step model (see Figure 2.10) is broken into three stages. The first stage is Preparation, and includes the first five steps of the model. The second stage is Data Collection, and includes steps six and seven. The third stage is Evaluation, and includes steps eight, nine, and 10. This model (TenBrink, 1974, p. 20) was specifically designed for the classroom teacher.

Figure 2.10. The Evaluation Process: A Model for Classroom Teachers
Evaluation Tools

Evaluation tools are instruments that help with planning or implementing an evaluation without serious time investments in research and without serious potential for errors or reinventing the wheel. Their purpose is to expedite the process, or at the very least make the process consistent and thorough. Below are several evaluation tools, each with a unique format, medium, or depth.

The Flashlight Program

The Flashlight Program aims to help institutions evaluate the Web’s influence on educational outcomes through its growing set of investigative tools and support services (Ehrmann, 2000). The Program has several components, including Flashlight Online, a Web-based service used by instructors and administrators to gather information from currently enrolled students; Flashlight Current Student Inventory, a database of almost 500 survey and interview questions; and the Flashlight Cost Analysis Handbook. The Flashlight Current Student Inventory has many questions that specifically focus on the uses of technology that increase access or that implement the “Seven Principles of Good Practice in Undergraduate Education”, compiled by Arthur Chickering and Zelda Gamson in 1987.

Evaluation Guide

This guide (see Figure 2.11) was developed to aid in accomplishing an evaluative task (Lucas, Miles, & Weber, 1973, p. 34; Weber, Worner, & Harris, 2000). Presented in matrix format, the guide depicts a systems approach to evaluation. The evaluation guide has two main components: the decision components (shown as column headings), and the program components (shown as row headings). The decision components are those elements that illustrate the format for evaluating program components. The program components are those elements of the program that are to be evaluated.
The authors note the difficulty of determining one set of components appropriate to all program evaluations, but they feel this guide is suggestive of the things evaluators may want to include when developing their own, more specific guides.

**Program Evaluation Standards, 2\textsuperscript{nd} Edition**

Valdez (2000a) writes, “Most evaluators would agree that the most comprehensive and ideal framework and definitions for how to assess evaluation of education programs may be found in *Program Evaluation Standards, 2\textsuperscript{nd} Edition*, developed by the Joint Committee on Standards for Educational Evaluation.” The standards identified four attributes contained in a good program evaluation:

- Utility
- Feasibility
- Propriety
- Accuracy

The first edition, entitled *Standards for Evaluations of Educational Programs, Projects, and Materials* (The Joint Committee On Standards for Educational Evaluation, 1981) established 30 standards that were a “working philosophy of evaluation” (p. 16). The standards outlined did not provide specific prescriptions of how to conduct an evaluation, but rather addressed principles that need to be considered when conducting an evaluation. The Joint Committee recognized and supported a need for the continuing review, revision, and development of evaluation standards.

---

<table>
<thead>
<tr>
<th>Program Components</th>
<th>DECISION</th>
<th>COMPONENTS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Program Goals</td>
<td>Operationalized Outcomes</td>
</tr>
<tr>
<td>Students</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adminstration</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Material</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Economics</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Personnel</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Figure 2.11. The Evaluation Guide*
An Educator’s Guide to Evaluating the Use of Technology in Schools and Classrooms

This handbook, produced by the U.S. Department of Education, has the goal of providing educators who have little or no formal training in evaluation a resource with which to “jump into the evaluation process, learning as they go” (Quinones et al., 1998, p. v). The guide understands the pressures being put on schools to incorporate technology into the curriculum, and understands the strings attached to the grants that enable this to happen. Quite often a teacher or administrator is given the task of analyzing and reporting on the implementation and effectiveness of technology programs. This guide provides an overview of the basics of evaluation, and includes worksheets and surveys that can be reproduced when conducting the evaluation. The guide is organized by typical questions one might ask while going through the evaluation process:

- Why am I evaluating? *
- What is evaluation anyway?
- Where do I start? *
- What questions should I ask? *
- What information do I need to collect? *
- What’s the best way to collect my information?
- What are my conclusions? *
- How do I communicate my results? *
- Where do I go from here? *

The asterisked questions contain a worksheet to help the evaluator answer the question. The guide also contains district and administrative surveys, teacher surveys, and student surveys.

Although the guide may be helpful, it misses many of the basic steps of evaluation. It also misses a thorough explanation of each step that it does include. The guide’s strongest point is its overview and example of translating goals into indicators, benchmarks, and measures. The format of the guide is very good, starting with the overview (44 pages) and ending with approximately 68 pages of reproducible worksheets and surveys. This is a strong idea and might be at the appropriate level for educators that are novices in evaluation, but would not be comprehensive enough for those conducting more extensive evaluations.

Designer’s Edge

Designer’s Edge (Allen Communication & (Tyler Tribe, 1999) is software that walks an instructional designer step-by-step through the instructional design process. The 12-phase process, which is based on a combined and refined approach of Gagne, Mager, and other instructional design theorists, was created for training design and planning (http://www.mentergy.com). The 12 phases are:

1. Analyze Needs
2. Mission Statement
3. Create Audience Profile
4. Write Objectives
5. Analyze & Outline Content
6. Layout Course Map
7. Define Treatment
8. Select Learner Activities
9. Create Detailed Plan
10. Produce Media
11. Author Course
12. Evaluate Course

Phase 12, the “Evaluate Course” phase, consists of the following steps:
• Collect evaluation data
• Conduct alpha test
• Revise course as necessary
• Conduct beta test
• Revise course as necessary
• Write evaluation report

The following forms were created to aid in achieving those steps:
• Alpha Test Checklist
• Beta Test Checklist
• Course Opinion Survey
• Objective Summary

The “Final Course Evaluation Report” can be automatically generated using an 11-step report wizard. The first step requests identification information (i.e. name) and the remaining 10 steps request information obtained from the “Course Opinion Survey.”

Three other phases in Designer’s Edge also contain components of evaluation. Analyze needs, the first phase, contains the following forms: Focus Group, Observation Form, Knowledge Assessment Form, One-on-one Interview, and a Questionnaire. These forms are for gathering information for a needs analysis, however, not for a formative or summative evaluation.

Layout Course Map, the sixth phase, has as one of its steps to “Write Course Evaluation Plan.” This also has a report wizard that can generate a report using eight steps that request the following information:
1. Identification information
2. Certification or accreditation?
3. List of those who will review course data
4. Purpose of collecting data
5. What kind of data is needed to fulfill the purpose
6. Checklist of data types
7. Course sections for data collection
8. Criteria for judging course success

Author Course, the eleventh phase, contains two forms that help to test all courseware before sending it our for field testing:
• Defect Report
• Quality Assurance Checklist
Although this software includes evaluation as a component, it does so only in the context of getting feedback on the very training being designed. It does not discuss how to evaluate, when to evaluate, or what to evaluate. It would not be possible to use this with an outside project or program.
CHAPTER 3: TOWARDS A FORMATIVE EVALUATION TOOL

Arguments For the Need of a New Formative Evaluation Model

Short of evaluators or investigators reading books about the evaluation process, or researching previous evaluations, each researcher would in essence have to “recreate the wheel” as far as evaluation planning, instrument creation, scheduling, and evaluation reporting. Worthen and Sanders (1987) write that evaluation literature is “badly fragmented and is often aimed more at fellow evaluation theorists than practitioners” (p. 44). Practitioners do not have the time necessary to consolidate and interpret the literature. Worthen and Sanders also write that evaluators are too busy to set off on an “intellectual expedition” to discover new methods and techniques that are possibly more relevant to the evaluation than their preferred method.

Worthen and Sanders reference Antonoplos (1977) and Gephart (1977), who have criticized the abundance of evaluation models and have suggested that existing models be synthesized. Antonoplos argues that the saturation of evaluation models confuses the person who wants an evaluation conducted, causes problems when planning and implementing evaluations, fragments and reduces evaluation theory-building efforts, and makes the learner’s task horrendous. “He calls for empirical research on the effectiveness of existing models and proposes ‘synthesizing them in a manner that is parsimonious, that prevents premature closure, and that reduces burdensome verbiage (Antonoplos, 1977, p. 9)” (Worthen & Sanders, 1987, p. 147). Gephart feels that it is logical to assume that there is one ideal model. He notes that the immaturity of the field inhibits us from yet seeing that all evaluation models could fit together, let alone how they fit together.

Pace (1978, p. 1) writes, “Identifying what is good or bad presumes an accepted theology, or something like it. In evaluation, we don’t have it.” TenBrink (1974) writes, “The process of evaluation involves many concepts. Even the experts are not in complete agreement as to exactly what activities should properly be included in the process of evaluation” p. 5).

Worthen and Sanders (Worthen & Sanders, 1987) note that the most relevant definition of a “model” is “something eminently worthy of imitation, an exemplar, and ideal,” but they write, “without clearer operational guidelines and procedures, most evaluation ‘models’ are so vague and general as to elude emulation in any strict sense” (p. 145). They note that most evaluation “models” aren’t models at all. They also write that evaluation “theories” aren’t actually theories, but simply conceptions and descriptions.

Alkin and Ellett (1985) discuss the difference between prescriptive and descriptive models.
A model is prescriptive because it gives advice, recommendations, warnings, and guidance about doing evaluations. The recommendations appear as recipes, flow diagrams, maxims, priority rules, strategies, and general guidelines. Often, a prescriptive model points out problems,
pitfalls, demands, and restrictions in doing an evaluation. In developing an evaluation model, a theorist is implicitly (or explicitly) specifying which standards, criteria, principles, and guidelines are the proper or appropriate ones to regulate the activities of an evaluator" (p. 1763). Alkin and Ellett then write, “A descriptive model is a ‘theory.’ An evaluation theory is a general set of empirical statements containing generalizations (or laws) for describing, predicting, or explaining evaluation activities” (p. 1765).

Reigeluth (1987) also delineates between descriptive theories or models and prescriptive theories or models when discussing instruction. A descriptive theory of instruction describes the effects of a whole model of instruction (integrated set of strategy components), instead of just the effects of a single strategy component. A prescriptive theory of instruction prescribes when a given model or set of models should be used. It identifies the instructional model that should be used for a given desired outcome and condition(s). The more comprehensive an instructional theory is, the more models it prescribes for different kinds of desired outcomes and conditions (p. 2).

Reigeluth argues that what the field of instruction needs are more integrated and comprehensive prescriptions. Although Reigeluth’s argument is for prescriptions for instructional theory, one can also see the benefits of more detailed prescriptions for components of instructional theory, i.e. for formative evaluation within the process of instructional design.

Alkin and Ellett note that since the 1970s there has been no single study aimed at developing a comprehensive theory of evaluation. Instead, most research has been related to particular aspects of evaluation (Alkin & Ellett, 1985). They emphasize the importance of developing descriptive evaluation models, not only because of the usefulness of understanding evaluation activities, but also because “descriptive evaluation models should undergird prescriptive models. Descriptive models provide the limitations and possibilities for the prescriptive models” (p.1766).

Braden (1992) discussed the need for formative evaluation checklists, routines, or questions to help minimize time investments when conducting formative evaluations using front-to-back formative evaluation models. Braden writes, “there is a high level of transfer” (p. 13) of these tools from project to project. Braden also feels that there is carryover from evaluation plan to evaluation plan, another integral step in his front-to-back formative evaluation model.

Braden also notes that the vast majority of instructional design models lack process evaluation. In fact, only three models other than the one he proposed (The Braden Formative Evaluation Model) provide for process evaluation: the CIPP Model (Stufflebeam, 1973), the VanPatten Model, and the Flagg Phases (Flagg, 1990) (all non-graphic). Braden quotes Reeves (1989, p. 164), who recognized the necessity for process evaluation when he wrote, “Obviously, prototype products will be a major focus of formative evaluation during an instructional design project, but there should be an equal
emphasis on collecting formative evaluation to improve the overall instructional design process…”

“The ways in which educators think about evaluation and how they go about it are necessarily related to what they are assessing and why. How evaluation should be conducted is thus related to the purpose, the program, and the personnel. Some standard recipes exist, but they are not interchangeable. Their differences arise largely from differing assumptions about why evaluation is conducted and what uses it serves” (Pace & Friedlander, 1978, p. 2).

“When evaluation is carried out in the service of course improvement, the chief aim is to ascertain what effects the course has – that is, what changes it produces in pupils” (Cronbach, 1975, p. 246). How the program produces its effects and the parameters influencing its effectiveness are of paramount importance. This means looking at all aspects of the instruction, more than just whether or not it is effective or ineffective. Cronbach notes that lumping information into a single score is a mistake because it masks failures and successes. He goes on to say, “The greatest service evaluation can perform is to identify aspects of the course where revision is desirable” (p. 247). Cronbach praises the virtues of formative evaluation by stating that “evaluation, used to improve the course while it is still fluid, contributes more to improvement of education than evaluation used to appraise a product already placed on the market…” (p. 247).

Cronbach (1975) notes that no specific set of guidelines exists for evaluation, and that no one set would be appropriate for all situations. However, there once was such heavy emphasis on achievement testing that testing became the methodology for evaluation. Testing, however, is not the only way and may not even be the best way, for program improvement. Cronbach also notes that both systematic (i.e. testing) and unsystematic data (i.e. observations and comments) are of great value. Systematic data becomes more profitable after obvious problems with early versions of the instruction have been fixed.

“Studies have shown that thousands of the instructional products sold in the United States each year have not been evaluated with learners and revised prior to distribution. Other studies have demonstrated that simply trying out materials with a single-learner and revising the materials on the basis of that data can make a significant difference in the effectiveness of materials” (Dick et al., 2001, p. 284).

Tessmer (1993) notes that the most effective evaluations begin with an evaluation plan. The process of evaluation planning needs simplification. Dozens if not hundreds of resources exist that contain valuable information on evaluation, but all lack an intuitive step-by-step user-friendly design. Many resources speak of evaluation from the perspective of one who will evaluate something to be used in his or her classroom, and not from the perspective of a designer, developer, or evaluator who needs to assess the program or product in question.
Valdez (2000b) identified three types of summative questions that should be addressed for any intervention: intervention efficacy, intervention effectiveness, and intervention costs. Intervention cost is a strong argument for an automated evaluation process. Although the time to implement an intervention may not be able to be reduced, the time invested in planning can be greatly reduced. One could also argue that intervention implementation time might also be reduced via proper and thorough evaluation planning. Evaluation planning is often overlooked which leads to evaluations that aren’t thorough, are asking the wrong questions, don’t answer the right questions, or just aren’t conducted at all.

Valdez (2000b) writes, “The key to a good evaluation plan is the design of the study or studies to answer the evaluation questions” (p. 5). Perhaps this seems obvious. This requires two things: prepared evaluation questions, and an evaluation plan. Too often neither of these components are prepared. Instruments from other investigations are used (which don’t answer questions directly related to the project at hand), or questions are created and instruments are developed, but the evaluation takes place at the end of the implementation, rendering no formative evaluation. Furthermore, without an evaluation plan in place which dictates what will take place when, as well as what questions to have answered and how to implement the results, the results are often NOT incorporated at all, rendering the evaluation altogether useless!

One possible reason for this is lack of a systematic formative evaluation process. Speaking of the Management Model of Evaluation, Harpel (1978, p. 20) says, “There is value in formalizing the process, making it explicit and systematic. Like a road map, it tells you where you are, where you want to go, and what routes will take you there. It is not intended to be a set of rules that must be adhered to rigidly: As on any journey, you may take some shortcuts. But it reminds you what routes you have missed in doing so, in case sooner or later you need to return.”

Worthen and Sanders (1987) note that the two of the greatest contributors to inadequate evaluation are the lack of performance information regarding educational products, practices, and programs; and the lack of an established system for acquiring such information. Worthen and Sanders (pp. 6-7) identify six arguments that support the use and improvement of evaluation:
1. A need for systematic school improvements
2. A need for cost-benefit analysis of programs that require large sums of money
3. A need to test theories about the effects of education on student development
4. A need for educators to appraise the quality of their school programs
5. A need to reduce uncertainty about educational practices when experience is limited
6. A need to satisfy external agencies’ demands for reports, to legitimize decisions, or to improve public relations through credible, data-based decision making.

“Unless evaluation practices improve significantly in the years ahead, its potential for improving education may never be realized” (p. 9).

Arguments Against the Need of a New Formative Evaluation Model
Not all believe that evaluation needs to be blueprinted or streamlined. Lucas et al. (1973) write,

It is not critical for all evaluation plans to conform to set procedures, and especially not important that program components being evaluated be the same for all programs…. The evaluation plan should be responsive to the needs of the system and meaningful and useful to those who are in charge of the program. (p. 35)

Worthen and Sanders (1987) feel that synthesizing models could be dangerous. They note that some evaluation approaches are so different in their philosophical assumptions that it would be inappropriate if not impossible to combine them. They also fear that “discipleship” would increase even more if models were synthesized. They also wrote that the field is too young to be restricted by uniform evaluation methods. “Moving toward one omnibus model at this time could bring premature closure to expansion and refinement within the field” (p. 148). They feel it would be more beneficial to tolerate the contradictory and confusing masses of models and theories than to force them into a unified but “impoverished” conception of evaluation. They further question whether one or two models could be generalized and relevant to all possible evaluation situations.

The movement towards a more prescriptive, comprehensive formative evaluation model does not have to be at the expense of the field of educational evaluation. In fact, more prescriptive, comprehensive models should further the field. Developing a model for formative evaluations is not an attempt to merge all previous models and approaches, but rather an attempt to address needs within the field.

The Development of a New Formative Evaluation Model

Dick et al. (2001) distinguish between formatively evaluating ones own instructional materials and formatively evaluating externally selected materials. Many evaluation articles, strategies, and tools are geared towards the end-user (instructor, teacher, etc.) to help to determine the effectiveness or appropriateness of a piece of instruction or instructional material once selected. “Although most evaluation theorists have said that their models are intended to influence and assist audiences, for the most part the nature and the role of these audiences has been given little or no special emphasis” (Alkin & Ellett, 1985, p. 1763). The author’s interest focused on the formative evaluation of instruction by the designer or developer within the context of the instructional design process. My interest is in helping designers and developers incorporate the step of formative evaluation into their design process for the purpose of improvement, not leaving it to the end-user to determine appropriate fit.

Tessmer (1993, p. 21) notes that all formative evaluations will follow these general steps:
1. Plan the evaluation
2. Introduce the evaluation to participants
3. Conduct the evaluation
4. Gather and organize data
5. Make data-based revisions
6. Evaluate revised version (if possible)

TenBrink (1974, p. 5) identifies three things that a model of the process of evaluation should do:
1. Clearly specify each step in the process of evaluation
2. Be applicable to all types of evaluation problems
3. Be easy to understand and easy to use

In fact, most formative evaluations follow the same procedures. However, what was found when reviewing the literature to find the most comprehensive formative evaluation on which to base a tool was that no one model contained all steps that the author felt were necessary. Although it must be said that often a book addressed each step, the goal of such a model was to avoid having a designer or developer need to read an entire book just to perform one step in the instructional design process, the step of formative evaluation. In lieu of one comprehensive prescriptive model, one was developed, pulling from the author’s own experience and backed by the repetition of many of the same steps in other formative evaluation models.

The following is the proposed prescriptive formative evaluation model, using the systems approach. The model incorporates the steps identified by Tessmer as being general to all formative evaluations, and honors the three guidelines proposed by TenBrink for a good evaluation model. Many of these steps can also be found in other models. (The referenced models with similar steps have been listed below each step.) These 14 steps condense the formative evaluation process into one that is comprehensive and manageable.

1. Clarify product/program goals and objectives
   [Tyler Model (steps 1 and 2), Metfessel and Michael Model (step 2), Provus’ Discrepancy Model (termed “standards”), Stake’s Countenance Model (Step 1 – termed “program rationale”), Stufflebeam’s CIPP Model, Walker’s Evaluation Design Model (step 2), Tessmer’s Formative Evaluation Planning Process (step 1), Evaluation Overview (step 1)]

2. Define purpose of evaluation
   a. Clarify evaluation type
   b. Clarify evaluation goals and objectives
   [Walker’s Evaluation Design Model (step 1), Tips for Conducting an Evaluation (step 1), Evaluation Overview (step 2), The Evaluation Process (step 1)]

3. Determine evaluation questions

4. Plan the evaluation
   a. Determine evaluator’s role(s)
   b. Identify evaluation design
1. Clarify product/program goals and objectives

Valdez (2000b, p. 2-3) writes, “A prerequisite for evaluation is the development of a project plan with measurable objectives that are logically related to one another and to the goals and interventions defined in the project proposal.” Stating goals can be done deductively or inductively. Deductively stating goals results from translating the needs assessment into the mission of the institution or organization’s programs and...
activities. Inductively stating goals results from assessing the institution’s or organization’s current programs and activities and determining the goals from those activities (Harpel, 1978). Harpel writes, “Whereas the focus of the deductive approach is on the relationship between goal statements and students’ needs, the inductive mode stresses the relation between goals and activities” (p. 26). Mager (1997) also warns against confusing objectives with goals. Whereas goals are broad, objectives should be more specific and measurable.

Due to the importance of goal statements, everyone involved should be included in the formation and review of goals statements. An evaluation will best serve its clients when the goals and objectives have been negotiated with input from all relevant stakeholders (Valdez, 2000a). Unless using a goal-free approach, an evaluator should learn as much as he or she can about the background, goals, and objectives of the program being evaluated (Quinones et al., 1998).

Harpel (1978, p. 26) identifies questions that test for the adequacy of goals:

- Does the goal address the needs of those who are served by the program?
- Does it relate to the activities of the program?
- Does it clearly identify the ideal result of the activities?
- Does the goal recognize the constraints of the environment (for instance, is it consistent with the purposes of the institution)?

2. Define purpose of evaluation

Clarify evaluation type

Work closely with all relevant stakeholders to determine what they hope to get out of the evaluation. Determining whether your evaluation is a formative or summative evaluation, or a process or outcome evaluation will help you determine the purpose(s) of your evaluation.

Clarify evaluation goals and objectives

Worthen and Sanders (1987, p. 24) distinguish between what evaluation is, “the act of rendering judgments to determine value”, and what goals evaluation has (i.e. decision-making). The goals and objectives are the stated purposes of the evaluation, and include decision-making, problem identification, impact determination, and continued program improvement (Quinones et al., 1998).

3. Determine evaluation questions

The evaluation questions determine what data needs to be gathered, when, and from whom. Having stakeholders give input into the process of determining evaluation questions helps to minimize the possibility of a fruitless or disappointing evaluation.

4. Plan the evaluation
Several authors have stressed the importance of evaluation planning (Dick, 1975; Kinnaman, 1992; Lucas et al., 1973; Valdez, 2000b). Valdez writes, A good evaluation plan should be developed before a project is implemented, should be designed in conjunction with the project investigators, and should ensure that the evaluation serves two broad purposes. First, evaluation activities should provide information to guide the redesign and improvement of the intervention. Second, evaluation activities should provide information that can be used by both the project investigators and other interested parties to decide whether or not they should implement the intervention on a wider scale. (p. 1)

Lucas (1973) writes, “Rejection of an evaluation can be avoided or minimized if the evaluation plan is well-conceived and well-developed” (p. 45).

**Determine evaluator’s role(s)**

A formative evaluator should be recognized as part of the program design team whose job is to acquire information that will help the design team improve the program (Dick, 1975). Dick writes, “A clear distinction should be made between the role of the evaluator during the formative stage of the program and during the summative stage” (p. 260).

“A formative evaluator should be knowledgeable in the areas of measurement and data analysis, but should also have a background in learning theory and experience in the construction of learning materials” (Dick, 1975, p. 260). Dick sees the evaluator’s role as more than just the acquirer of useful information. He feels the evaluator should explain to the design team the procedures involved in the evaluation, including data analysis, and should also help interpret the data from the evaluation. Dick sees the evaluator’s role as an asset not only in evaluation planning and data collection, but also in program revision.

**Identify evaluation design**

Determine what data acquisition strategies will be used (Valdez, 2000a). Many evaluations fail due to inadequate research designs and poorly selected or constructed outcome measures (Oetting & Cole, 1978). The evaluation question(s) will help determine the most appropriate design. “Understanding the strengths and weaknesses of various research designs will help evaluators match the research question with the best possible approach” (Hanson, 1978, p. 98). “The purpose of the evaluation should influence the design of the evaluation” (Quinones et al., 1998, p. 3).

5. **Select criteria, indicators, and data sources**

Before obtaining information, one must first know what judgments and/or decisions are to be made. Knowing what judgments and decisions are to be made not only aids in obtaining the right information, but also how and when to obtain the information (TenBrink, 1974). Determine what data will be necessary to obtain in order to make informed decisions (Valdez, 2000a).
Dick (1975) noted in his study of programmed instruction that the ability of the student mattered to the designer when considering the comments that student made. Determine if this might be a factor to you or the designers and administrators making decisions, and determine indicators and create instruments accordingly.

Determine what data is useful for the evaluation that already exists, and determine what new data will need to be collected (Valdez, 2000a). “Failures in evaluative research are frequently due to a poor link between whatever was assessed and the goals of the program. Thus, the first step in deciding on outcome measures must be to determine carefully what the program is meant to accomplish and exactly what changes would finally occur if it met its goals” (Oetting & Cole, 1978, p. 47).

6. Develop instruments

Get input from relevant parties during development

It is extremely important when developing instruments to get input from all relevant parties. The instruments need to gather data that is relevant and productive for those who designed and developed the product, and for those who will make decisions regarding the product. Dick et al. (2001) speak to the design of formative evaluation instruments from an instructional design perspective, “Using the instructional strategy as the frame of reference for developing evaluation instruments and procedures should help you avoid designing a formative evaluation that is either too narrowly focused or too broad” (p. 298).

Dick et al (p. 310) write, “Good formative evaluation includes assessing both achievement and attitudes.” Cronbach (1975) also notes the importance that designers place on attitudes as an outcome. Cronbach writes that interviews and questionnaires are valuable if not trusted blindly. Attitude questionnaires have been criticized because of the distortion that results when students answer less than honestly (too positively or too negatively).

When measuring proficiency, standardized tests are useful. However, Cronbach (1975) notes that giving different questions to different students can be a good technique when the evaluator wants to acquire more information (i.e. for the assessment of an entire course.) Essay tests and open-ended questions, although possibly inappropriate for large groups, can be profitable and beneficial in some situations. Cronbach also speaks to test development when he says, “If you wish only to know how well a curriculum is achieving its objectives, you fit the test to the curriculum; but if you wish to know how well the curriculum is serving the national interest, you measure all outcomes that might be worth striving for” (p. 253).

“Outcome measures fall into one of three groups: direct measures of exactly what you are trying to change, “off-the-shelf” measures that have been pretested in other settings, and “build-your-own” measures designed along with the study. The first two are
obviously preferable, since negative results can then be attributed to the program” (Oetting & Cole, 1978, p. 46). The key is to determine whether the instrument is valid for your purposes. “Choose your measures so that the information will generate answers to the questions you are asking” (Oetting & Cole, 1978, p. 47).

7. **Submit evaluation proposal**

**Determine time schedule**

The time needed for an evaluation depends on many factors, including (but not limited to): the number of steps to be incorporated, the type of steps to be incorporated, the number of participants selected, the participants’ schedules, the evaluation budget, the time required to incorporate changes, and the overall product or project timeframe.

**Determine budget**

Determining the budget of an evaluation is a difficult task. A rough estimate could vary greatly depending on the type of project, the purpose of the evaluation and the depth of the evaluation. More specifically, determining the exact costs associated with an evaluation is an elusive task. There are the costs of adoption of a product/program, costs for human resources, and costs for material resources. There are differences between projected costs and actual costs, initial costs and ongoing costs (Anderson, Ball, Murphy, & Associates, 1975).

Determining budgetary considerations (possible human and material resources) may be the first step in budgetary planning for an evaluation. Having a tool that could help to simplify the process of determining resources and costs might prove very useful. Exact cost estimates may never be able to be achieved, but is worthy of further research.

Another area of budgetary concern is cost efficiency/effectiveness (also known as cost-benefit analysis or systems analysis). Specifically, is what is gained worth the costs involved? Can the same rewards be reaped at a lesser cost? Can greater benefits be reaped at the same cost? (Anderson et al., 1975). This is an area of cost analysis not currently incorporated into the tool in its current design, although it is certainly a worthy undertaking. The tool’s current goal of budget determination centers on the determination of factors directly related to implementing a formative evaluation of the program/product being developed: the time involved (human resources), the materials required (materials resources), and other miscellaneous or overhead expenses.

**Submit evaluation plan, time schedule, budget, and instruments for feedback.**

Feedback from stakeholders (whether management, the funding agency, or others involved less directly) as to the thoroughness, timeliness, and fiscal responsibility of the planned evaluation is much more acceptable at this stage than at a later stage. At this stage, comments or changes can be incorporated relatively easily. What one wants to avoid is conducting an evaluation that doesn’t meet the purpose outlined due to the
gathering of the wrong information, or the information was gathered at the wrong time, or a plan was accepted that could not be implemented due to financial restrictions.

8. **Make revisions to the evaluation proposal (if necessary)**

   Feedback from all stakeholders (and especially those directly involved in the evaluation) may prompt a revision of the evaluation plan, time schedule, budget, instruments, or all of these. This step provides for revisions now, when such changes can be made relatively easily.

9. **Schedule appointments for evaluation**

   Identifying who will be involved, how, and when is an integral step in the process of the evaluation.

10. **Conduct evaluation**

    **Revise evaluation instruments (if necessary)**

   After conducting a portion of the evaluation, problems in question wording or instrument presentation may be realized. Identifying such problems and addressing them can make the information gathered during subsequent stages much more useful.

11. **Organize and analyze results**

    Organizing and analyzing the data can be done using several possible strategies. These may include describing, counting, coding, clustering, comparing, finding covariation, and telling a story (Weiss, 1998, p. 285).

12. **Report results**

    Reporting evaluation results to relevant parties is a crucial step in the evaluation process. Designers, developers, and decision-makers will use the report to make improvements, revisions, and modifications to the instruction or materials. Administration, management, and funding agencies will use the report as documentation of the evaluation efforts and for funding considerations. Evaluation participants will view the report to see how their input affected the results. Valdez (2000a) writes of the importance in recognizing the report audience and their informational needs.

    The best evaluation is totally useless if the reports are not read or are not understood by those responsible for making decisions or expected to make the changes. Very often reliance on a single report will prove inadequate because different populations have different needs for the degree of technical sophistication, comprehensiveness, and detail of information. One should consider when developing appropriate reports for targeted audiences identifying the audience's special information needs. The most appropriate communication strategies for each audience should be
identified as soon as possible. A timeline needs to be developed that will
guide when specific information will be shared with targeted audiences
especially as the reports relate to formative evaluation. (pp. 2-3)
“Getting the right people to use the findings for making decisions is the most difficult
task associated with evaluation” (Hanson, 1978, p. 98). Hanson believes the days when
evaluators simply collected and reported information are over. He believes a greater
involvement is necessary; an evaluator needs to be a “consultant”.

The Joint Committee recognizes Report Clarity as one of its 30 standards, and
says, “The evaluation report should describe the object being evaluated and its context,
and the purposes, procedures, and findings of the evaluation, so that the audiences will
readily understand what was done, why it was done, what information was obtained, what
conclusions were drawn, and what recommendations were made” (The Joint Committee
On Standards for Educational Evaluation, 1981, p. 37). Reports must be appealing and
compelling to their audience. Different reports may have to be provided for different
audiences (Stevens et al., 1997).

Valdez (2000b) also accentuates the importance of conveying in the report how
the evaluation results were used. “All reports should include not only evaluation results
and reports of progress, but detailed explanations of how those results were used to
reinforce, refine, or modify project activities” (p. 9).

13. Revise instruction accordingly

Determine if problems were due to a faulty program, to a faulty instrument, or

14. Publish results

Publishing or publicizing your results is an important step in the evaluation
process. These results may be made available to outside, interested parties via journals,
newsletters, web sites or other electronic venues, or other publications.

In addition to the 14 steps outlined above, the following suggestions regarding
formative evaluation are valuable to keep in mind. Valdez (2000b, p. 9) lists four
strategies that help to ensure that evaluation activities are an integral part of any project.
- Establish a routine information system for the project, including inputs (time,
  resources), outputs (activities completed, student contact hours), and outcomes
  (student course grades, interim results of evaluation activities). Once established,
a member of the project staff should be held responsible for keeping the
  information system up to date.
- Include evaluation activities in the project budget.
- Hold regularly scheduled monitoring and evaluation meetings for project staff.
  All those who work on a project should be familiar with the project objectives and
  how they will be evaluated.
Encourage review and revision of the evaluation plan… Do not hesitate to revise aspects of the evaluation plan – to strengthen the research designs, select alternative indicators if the original ones are not sufficiently sensitive to project achievements, or incorporate the results of formative research.

Pace and Friedlander (1978, p.12) identify questions an evaluator may want to keep in mind:

- How are the objectives or criteria determined? By whom? In what detail?
- What is the program or set of experiences to be evaluated? Is it clear?
- Who are the evaluators and what role do they play? Is the evaluator an independent scientist? One whose activity is largely guided by the needs of a client? A participant in program planning and improvement? A neutral observer? An information collector? A judge?
- Why is the evaluation being made? Who wants it? For what purpose?
- Who (what person or group) makes a judgment about the merit or worth of the program? And what happens next?

Arguments For the Need of a Formative Evaluation Tool

The development of a tool of any sort requires at least one thing: justification for the tool. Charles Reigeluth (1987) writes, “The purpose of the field of instruction is to provide educators and trainers with prescriptions for making their instruction more effective and appealing” (p. 1). Speaking on instructional theory as a whole, he notes that methods of instruction have been analyzed to identify the fairly elementary components that comprise them, but that the prescriptions developed for the use of those components is mostly “piecemeal”. It stands to reason that if instructional theories are in need of further prescriptive development, then so too are evaluation theories. Another way to look at it is that evaluation is a step within most instructional design models. If instructional theory needs further prescriptive development (perhaps via more prescriptive models), then so too should each step within the prescriptive models, including evaluation.

The Standards for Evaluations of Educational Programs, Projects, and Materials (The Joint Committee On Standards for Educational Evaluation, 1981) was developed by the Joint Committee On Standards for Educational Evaluation. The Committee, chaired by Daniel Stufflebeam, had most top names in evaluation engaged as committee members, consultants, officers, writers, participants, reviewers, and design consultants. The Committee recognized a need to standardize the skills required of one who commissions, conducts, or employs results from evaluations designed to improve education. The Committee believed that “a set of professional standards could play a vital role in upgrading the practice of educational evaluation. Finally, the Committee members agreed at the outset of this project that no adequate standards for educational evaluation existed” (p. 5). The book identifies 30 separate standards, for which it established “guiding principles” (p. 9) for each. Each standard contains practical guidelines and warnings of pitfalls and caveats. Standards range from evaluator credibility to report writing. The book was, no doubt, the answer of the day to a poorly structured field. The
Committee openly supported ongoing reviews and revisions of the standards as they themselves perceived the book to be “a necessary, imperfect beginning” (p. 12). As Reigeluth (1987) implies, it is our responsibility to continue to develop the field, and to continue to develop prescriptions to make the job of developing instruction (and evaluating instruction) easier.

The Joint Committee foresaw several benefits from the development of sound standards: a common language to facilitate communication and collaboration in evaluation; a set of general rules for dealing with a variety of specific evaluation problems; a conceptual framework by which to study the often-confusing world of evaluation; a set of working definitions to guide research and development on the evaluation process; a public statement of the state of the art in educational evaluation; a basis for self-regulation and accountability by professional evaluators; and an aid to developing public credibility for the educational evaluation field” (The Joint Committee On Standards for Educational Evaluation, 1981, p. 5).

The reasons listed above are still valid today as the field continues to be developed. They attest to what could be gained by further developments in the field. The objective of the Committee was “to ensure that evaluations would be conducted effectively, fairly, and efficiently” (p. 6). That too is the objective of the formative evaluation tool being proposed.

The development of a formative evaluation tool builds on the many evaluation models and approaches proposed over the years. Specifically, the formative evaluation tool builds on the usefulness of the formative evaluation model described earlier. The tool would provide step-by-step procedural knowledge of how to conduct each step of planning and conducting a formative evaluation by providing guidelines, examples, and worksheets to the user. After an extensive literature review that focused on the models, approaches, and tools of evaluation, it is my opinion that no such tool currently exists. Although many books contain this much needed information, the secondary goal of such a tool was to expedite the process, so that no user would have to read books on the topic, just to carry out something that should be inherent in the instructional design process.
CHAPTER 4: METHODOLOGY

A thorough literature review was conducted (see Chapter 2: Literature Review) to determine whether or not a formative evaluation tool existed, and whether such a tool (if found) was prescriptive and comprehensive in nature. No such tool was found. The literature review also focused on evaluation models with the intention of finding a model comprehensive of the formative evaluation process on which an evaluation tool could be based. In lieu of finding one comprehensive model, a synthesized model was created (see Chapter 3: Towards A Formative Evaluation Tool). This model would serve as the basis for the future development of the formative evaluation tool.

Evaluation Planning

The next phase was to plan the design, development, and evaluation of such a tool. The model itself was utilized to demonstrate its usefulness when planning an evaluation. The following is the formative evaluation model, with planning comments below each step. The information contained in this example has been left unchanged since it was initially drafted during the planning process. It is meant to demonstrate and document the process.

1. Clarify product/program goals and objectives
   To aid instructional designers and developers in the process of formatively evaluating their instruction and materials. To expedite the planning and procedures inherent in formative evaluation by introducing them in a step-by-step manner and including within each step guidelines, printouts, and prompts for an efficient and thorough formative evaluation.

2. Define purpose of evaluation
   a. Clarify evaluation type
      Formative evaluation for product improvement.
   b. Clarify evaluation goals and objectives
      The purpose of the evaluation is to improve the formative evaluation tool through feedback regarding design, usability and usefulness.

3. Determine evaluation questions
   a. What is the current role or job of the expert reviewer? How do they intend to use this tool?
   b. Is the purpose of this tool apparent?
   c. Is it clear who the intended user of this tool is?
   d. Is the tool appropriate for the intended user?
   e. Is the tool appropriate for the intended environment?

   The following evaluation questions are categorized into inputs, processes, and outcomes.

   INPUTS:
   a. Is the purpose of this tool apparent?
   b. Is it clear who the intended user of this tool is?
   c. Is the tool appropriate for the intended user?
   d. Is the tool appropriate for the intended environment?

   PROCESSES:
   a. Is this tool well designed?
Is it visually pleasing to the eye?
Is the navigation of this tool intuitive?
Is the tool accessible?
Did you have difficulty opening it on your computer? If so, what type of computer are you using?
Does the tool cover all aspects of planning and implementing a formative evaluation?
Is the tool well designed for collaboration? If not, how could it be improved?
Is the tool efficient?
Does each step in this tool seem necessary for your purposes? If not, explain.
What would you add or remove to improve the tool?

OUTCOMES:
- Does each step contain the components necessary to make the formative evaluation process thorough and efficient (i.e. worksheets, definitions, guidelines)?
- Would utilizing this tool save you time when conducting a formative evaluation?
- Could this tool be used as a summative evaluation tool?

Overall comments and suggestions will also be solicited.

4. Plan the evaluation
   a. Determine evaluator’s role(s)
      My role is as the internal formative evaluation planner and facilitator. Those engaged in the expert reviews will be external formative evaluators and experts in the field of evaluation or instructional design.
   b. Identify evaluation design
      Qualitative in nature: expert reviews utilizing questionnaires.

5. Select criteria, indicators, and data sources
   Expert Reviews: Comments and feedback will be elicited once from expert reviewers. They will be given materials and documentation, as well as a questionnaire, and will have two weeks to answer questions regarding the evaluation tool.

6. Develop instruments
   See Appendices.

7. Submit evaluation proposal
   In this case, rather than submitting an evaluation proposal to the design team or stakeholders or management, the evaluation plan will be submitted to the committee members for approval.
   a. Determine time schedule
      Rapid prototyping of the tool will begin almost immediately and will continue through March 2002.
b. **Determine budget**

N/A

c. **Submit evaluation plan, time schedule, budget, and instruments for feedback**

The instruments will be evaluated on several levels. First, feedback will be obtained from committee members before the instruments are utilized. Second, instruments will be revised as necessary when specific questions are determined for the formative evaluation tool (Step 3). Finally the instruments will again be revised if necessary as they are implemented in the expert reviews (Step 10a).

8. **Make revisions to the evaluation proposal (if necessary)**

Revise evaluation plan and/or instruments accordingly.

9. **Schedule appointments for evaluation**

Three expert reviewers will review the materials and documentation from March 4 – March 15.

10. **Conduct evaluation**

a. **Revise evaluation instruments (if necessary)**

11. **Organize and analyze results**

Results will be organized around the questions posed in the questionnaire and around the comments elicited with each step.

12. **Report results**

Results will be reported in Chapter 5: Results and Conclusions, rather than in an independent report.

13. **Revise instruction accordingly**

In this case, the evaluation tool would be revised accordingly and would be run through another round of evaluations.
14. Publish results

In this case, the results would be published as part of the thesis and dissertation process.

Tool Development

After utilizing the model to aid in planning the design, development, and evaluation of an evaluation tool, the next phase was to actually develop the prototype of the tool. The 14-step formative evaluation model outlined previously was the basis for the formative evaluation tool. In other words, the tool embodied the 14 steps of the model. The purpose behind researching existing models (and subsequently developing a synthesized model when an “ideal” was not found) was to start with the most comprehensive and prescriptive set of steps in formative evaluation on which to base a formative evaluation tool. The model was the foundation and guiding framework for the tool’s development. Within that framework will be the opportunity to operationalize, automate, and functionalize the 14 steps of the model.

A formative evaluation tool based on the prescriptive model outlined previously could have taken many forms. The two that seemed most logical were paper-based (such as a handbook) or electronic-based (such as software). The software option would have allowed the user to input information and save it, and would also have allowed the option for printing worksheets or reports so the user could have a paper-based copy if so desired. The medium selected for this tool was electronic-based, understanding that the steps and related handouts and “worksheets” are printable for those who like to work from paper. Specific software package possibilities for design and development included Microsoft PowerPoint, which would have produced a fully designed but inoperable prototype, Adobe Acrobat which would have allowed for input but only minimal operationalization, and Macromedia Director, which would have produced a fully designed and fully operational prototype. The latter option would also have allowed for easy distribution. Macromedia Director can create a projector, which is a standalone application, or it can use Aftershock to create a Shockwave movie that would display the tool on the web. Time constraints and lack of familiarity with software packages dictated the software selected for developing the tool.

A prototype for the sake of illustration was created using Microsoft PowerPoint (see Appendix A). The tool was temporarily named Evaluation Express. The prototype illustrates the framework that will guide future development of the tool. The 14 steps of the formative evaluation model are consistently displayed on the left-hand side of the screen, with the step currently being employed highlighted. The rest of the window (to the right) contains information, guidelines, and space to input information for each step. The users have the ability to save information input into the tool (for purposes such as collaboration or for future report generation) as well as the ability to print. Often informal evaluations are done in the formative stages of instruction, but are not well documented. The goal of this tool was to simplify and streamline the formative evaluation process, including documentation.
Tool Evaluation

When planning this round of tool evaluation, it was determined that three experts would be used during the expert review process. Tessmer (1993) identifies the expert review as the review of the instruction by experts, with or without the evaluator present. He notes that experts can be content experts, technical experts, designers or instructors. Tessmer’s book, *Planning and Conducting Formative Evaluations* was itself formatively evaluated using an expert review and a field test evaluation. The decision was made to utilize three experts: two evaluation (content) experts and one instructional design expert (who happened to be very well educated on evaluation as well). Two expert reviewers were committee members who were not exposed to the tool prototype prior to their expert review, and the third expert reviewer was an external expert reviewer. The expert reviewers evaluated the tool from the perspectives of usability: ease of use, intuitive flow, content, etc.

The expert reviewers were given the tool itself (see Appendix A on page 85) as well as the tool’s flowchart (see Appendix B on page 110). An introduction to the tool was also provided (see Appendix C on page 112), as well as the capabilities of each step within the tool, since this was not a completely functioning prototype (see Appendix D on page 115). Finally, a questionnaire was included that could be filled out from a distance. This questionnaire can be seen in Appendix E on page 121. The instrument Tessmer used when Walter Dick formatively evaluated his book was requested for comparison’s sake, but was unavailable. That evaluation was conducted in the same context as this evaluation was – at a distance. The timeline for the expert reviews was about two weeks. An additional two weeks were needed to synthesize the data obtained from the expert reviews.

Contrasting the experts who reviewed the tool, the tool would ideally be used by instructional designers and developers that are not very familiar with the evaluation process and therefore need guidance when formatively evaluating their instruction or instructional materials. Therefore, the users would be training developers, instructors, and instructional designers who are not evaluation experts, but are engaged in the field and can benefit most from the finite process prescribed as well as the end results of such an evaluation. Future evaluation steps (i.e. the one-to-one evaluation) would provide insights from the target user. Future evaluation steps are discussed in Chapter 5: Results and Conclusions. The methodology outlined for this dissertation (three expert reviews) is provided as foundational support for the concept, (the concept and justification of the concept being the main thrust behind this developmental dissertation.) A fully functioning and fully evaluated tool was not the intended result of this developmental dissertation. However, this may be the next step.
Table 4.4. Overview of methodology

<table>
<thead>
<tr>
<th>To be evaluated</th>
<th>Methodology</th>
<th># of evaluators</th>
<th>Instrument(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tool</td>
<td>Expert Review</td>
<td>3</td>
<td>Appendix A: Prototype</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Appendix B: Flowchart</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Appendix C: Introduction</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Appendix D: Software Capabilities</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Appendix E: Formative Evaluation Tool Expert Review</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Instrument</td>
</tr>
</tbody>
</table>

Quantitative or Qualitative Methodology?

In Jared Danielson’s dissertation (1999, p. 86), Dr. Vic Bunderson of Brigham Young University is given credit for proposing that most instructional software requires at least four cycles of implementation, evaluation and revision to achieve its full effectiveness. This is obviously the first round of evaluation that the tool is undergoing. The extent of this round of evaluation is an expert review. Quantitative measures would not be very helpful at this point because of the small number of participants. Instead, extensive qualitative information will be obtained that will give insight into the future design changes to the tool. Later iterations of the evaluation could focus on gathering quantitative information, or could at least share the focus with qualitative information.
CHAPTER 5: RESULTS AND CONCLUSIONS

Executive Summary

As was outlined in Chapter 4: Methodology, three experts reviewed the formative evaluation tool (which can be seen in Appendix A): two evaluation experts and one instructional design expert. (Their roles are being stated here to clarify who was involved; they addressed themselves only as expert reviewers when responding to the question posed on the questionnaire.) To differentiate their comments as well as to identify their roles as experts, the three expert reviews will be identified as Instructional Design Expert (IDE), Evaluation Expert #1 (EE#1), and Evaluation Expert #2 (EE#2).

Overall, the concept of an automated evaluation tool was accepted as having its place. Many suggestions were given that if incorporated would benefit the prototype immensely. Comments were also given that lent support to certain ideas while other comments identified areas of confusion. All comments given are listed below. Most comments are direct quotes; changes were only made to comments that either revealed an expert’s identity or affiliation, or that were confusing and needed clarification.

Each step and questionnaire question has the corresponding comments and suggestions summarized below it with possible solutions added. However, the following points represent the recurring comments and the strongest consistencies among the feedback given.

- **Documentation was strongly recommended.** Documentation was provided in the form of brief explanations of capabilities to the expert reviewers (see Appendix D on page 115), and would need more expansion for the end user. Simply adding documentation would clarify many of the confusions from the expert reviewers, and would address many of the other concerns they voiced regarding the end user’s ability to fully utilize and understand the tool.

- **A tutorial or sample evaluation was recommended.** Adding an example evaluation that utilized each step or creating a tutorial that demonstrated each step was suggested as a means of clarifying the tool’s utility.

- **Adding information within the tool was recommended.** There were several places where simply providing a definition, an example, or even a handout would have improved the usability greatly.

- **Adding collaborative techniques was recommended.** Providing the ability to rank order evaluation questions or prioritize objectives were two examples of collaboration enhancements that would benefit the tool.

- **The tool was not seen as only a formative evaluation tool.** Rather, the tool was seen as an evaluation tool that could be utilized for both summative and formative purposes. In fact, the question posed to the expert reviewers “Could this tool be used as a summative tool?” was almost unanimously met with surprise that the tool hadn’t been designed with summative evaluations in mind.
Expert Reviewer’s Step-By-Step Comments

The results of the expert reviews are detailed below. Comments and suggestions for the tool’s improvement are listed below with their corresponding step (as seen in Appendix D).

Opening screen

(EE#1) It’s not clear what the purpose of the program is or who it is for. You may want to put buttons or rollovers on the opening screen.

Summary of comments, suggestions, and possible solutions:
- Clarify program intentions and clarify intended audience.
- Put rollovers on opening screen that present this information.

1. Clarify product/program goals and objectives

(IDE) I strongly agree within step one that involving multiple other parties in the evaluation is a good idea, having stakeholders clarify and focus the evaluation's intent, primary goals and objectives, up-front. I'm unsure of your development possibilities, but it would be nice if the tool could be "packaged" online so that someone could create an evaluation "project" within the tool structure, and then the multiple stakeholders could be given access to that project's files (via user names/passwords). For stage one, then, each stakeholder could list their top goals/objectives, and it would be great if you had a feature where the stakeholders could then prioritize/rank the goals/objectives listed by others (maybe just a pull-down list next to each recommended goal--high priority, medium priority, low priority). A “consensus-building” feature... “six of eight stakeholders rated this goal as having a ‘high priority’ for investigation.”

(EE#1) For those who may not know what “goals” or “objectives” are, there should probably be a hot link from these terms to define for users what these terms mean and to give contrasting examples.

(EE#2) A statement (such as the one provided in Appendix D: Software Capabilities – maybe briefer) would help the user understand the process a bit more. It might help to provide an example of a goal and objective so the user would know what is expected.

Summary of Step 1 comments, suggestions, and possible solutions:
- Make the tool available online.
- Enhance collaboration features to include a “consensus-building” feature that ranks or prioritizes goals and objectives.
- Make rollovers or hot links for the terms “goals” and “objectives” to provide definitions and examples of these terms to the user.
- Clarify that when determining “goals” and “objectives”, this is of the product/program, not of the evaluation.
2. Define purpose of evaluation

(IDE) The relationship is not clear between the three dichotomies at top... formative/summative, process/outcome, overt/covert, and the check-box overt reasons for conducting an evaluation listed below the dichotomies. It appears the “decision-making” check-box items are related to summative evaluation, and the ‘program/product improvement’ check-box items are related to formative evaluation. Maybe these items are better clumped under the formative/summative dichotomy? Also, the dichotomies may be setting up an expectation... that formative evaluations look at processes, and summative evaluations look at outcomes (which logically they do), but then you have the third dichotomy overt/covert that doesn't really fit the pattern... formative/process/overt, summative/outcomes/covert? You might want to think about the organization here. I think it might confuse some folks. Maybe list the check-box items first, asking, “Do your evaluation goals focus on the following program/product improvement issues?” If yes, your evaluation is “formative” to improve a program/product.

(EE#1) I’m not sure (the term) stakeholders will be clear to everyone without some definition and, of course, they will need to be representives rather than the entire stakeholder group sometimes, won’t they? … I kind of feel like “purpose” may need to go first or at least earlier. Doesn’t purpose affect stuff like stakeholders and goals?

(EE#2) I don’t understand the reason for “Highlighting Goals”.

Summary of Step 2 comments, suggestions, and possible solutions:
- Better explain differences between formative/summative evaluation, process/outcome evaluation, and overt/covert evaluation.
- Separate descriptions of the purposes and the check-box items into two separate slides.
- Make rollovers or hot links for the term “stakeholders” to provide the definition of this term to the user.
- Stakeholders may in fact determine the purpose. On the other hand, the purpose may determine the stakeholders. Move listing of stakeholders from Step 1 to Step 2.

3. Determine evaluation questions

(IDE) Really like the features you have planned here... the group can create a list of questions for investigation. Again, it might be useful to have a ranking feature whereby the most important questions could be pinpointed more easily by stakeholders. Also, it might be useful to carry the goals or objectives forward from step one, and allow questions here to be written or shuffled/placed within that initial structure (i.e., six questions related to goal one, four questions related to goal two, etc.). The sample questions you have provided are well constructed. I noticed you had inputs, processes, and outcomes or products, which mirrors Stufflebeam's CIPP (context/input/process/product) structure. You might want to include sample questions in the "context” category as well. That one set seems to be missing and is important to
include. Also, under the category "outcomes," you might want to consult the Flashlight program's 14 sub-scales which list common outcome categories such as motivation, critical thinking, active learning, time on task, faculty-student interaction, student-student interaction, and other types of outcomes. It might give you some expansion of the sample questions you have already listed there.

(EE#1) I like the “click here for evaluation questions” approach.

(EE#2) OK.

Summary of Step 3 comments, suggestions, and possible solutions:
- Enhance collaboration features to include a “consensus-building” feature that ranks or prioritizes evaluation questions.
- Create an organizational feature that would allow evaluation questions to be organized under goal/objective subheadings.
- Add “context” as a category for evaluation questions.

4. Plan the evaluation

(IDE) I'm not sure I understand why focus group, expert review, and field test, are listed on this page and not on the step five criteria/indicators/data sources page? Some of the step four items you have seem like mechanisms to collect data just like an observation or questionnaire in step five. Maybe collapsing steps four and five would be less confusing?

Then when someone selects a data source, whatever that may be, your idea for placing it on the time schedule (step 7) and the appointments sheet (step 9) would still be valid. I think your idea of evaluation ‘procedures’ and/or coming up with an evaluation plan is already built-into your automatic, electronic time schedule (step 7). You may not need this extra step for that if you let people pick their data sources and instruments (step five) and then use the schedule to plan and organize their flow.

(EE#1) Looks just fine to me.

(EE#2) Again, a statement (such as the one provided in Appendix D: Software Capabilities) would help the user know what to do.

Summary of Step 4 comments, suggestions, and possible solutions:
- Collapse step 5 into step 4.
- Provide documentation to guide the user through the process.

5. Select criteria, indicators, and data sources

(IDE) It seems like the objectives from step one should be carried over to this step, so that the user can define criteria and/or indicators for each. On a related note, the evaluation questions that were selected in step three should be carried over and shown here so the user can match specific data sources with specific questions (e.g., “To answer evaluation question one, we will use interviews and focus groups...”). You'll love this
one... it would be nice to have short, one-page job aids that described each data source and instrument (quick description, why they are used, when they are best used, variations, and samples).

(EE#1) Seems very clear to me.

(EE#2) In some (many?) formative evaluations much of the data are for information purposes only and criteria do not need to be established!

Summary of Step 5 comments, suggestions, and possible solutions:
- Create an organizational feature that would allow criteria and indicators to be organized under goal/objective subheadings.
- Create an organizational feature that would allow data sources to be organized under evaluation question subheadings.
- Add pop-up windows that describe/define data sources or instruments.

6. Develop instruments

(IDE) The sample instruments are a good idea.

(EE#1) Good. It would probably help to have something like “tips for interviews and observations.” Will the Test and Survey buttons stay grayed out? If not, they will probably need tips also.

(EE#2) Some questions listed under inputs may be processes - some questions under processes may be inputs - some questions under outcomes don’t specifically address an outcome.

Summary of Step 6 comments, suggestions, and possible solutions:
- Create “Interview and Observation Tips,” “Test Tips,” and “Survey Tips” handouts. Link these from the Evaluation Instruments page.
- Create sample test and survey?
- Reorganize and reword questions to appropriately and clearly reflect their intention.

7. Submit evaluation proposal

(IDE) No comments at this time.

(EE#1) Very good. I’m assuming “by” means “person responsible” – if so, you may want to say it that way.

(EE#2) Does “by” reflect a date or a person? Otherwise, seems clear.

Summary of Step 7 comments, suggestions, and possible solutions:
- Rephrase “By” to “Date for submission”.
8. **Make revisions (if necessary)**

(IDE) No comments at this time.

(EE#1) Works fine except I’m not sure what “component” is – one of the 14 steps? A piece of a step like objectives or a piece of a piece like single objective, or, since it references back to step 4, is it an evaluation procedure?

(EE#2) The word “component” bothers me. I don’t know what it refers to. “Deliverables” too. It sounds like jargon.

*Summary of Step 8 comments, suggestions, and possible solutions:*
- Clarify the meaning of “component” or reword entirely. (Here, “component” refers to any part of the evaluation as it was presented in the proposal - the instruments selected, the timeline, etc.)
- Clarify the meaning of “deliverables” or reword entirely. (Here, “deliverables” refers to the resulting revisions on the “component” of the evaluation - the instruments selected, the timeline, etc.)

9. **Schedule appointments for evaluation**

(IDE) Only thing to consider is what I said before in step four. It seems that scheduled events would also be necessary for surveying, conducting observations, conducting interviews, etc. The procedures you have listed in step four are not the only things that would require scheduled appointments. I guess in step five you have deadlines for instrument *creation* that get placed on the calendar, but what about the actual dates of instrument *utilization*. Can those dates be passed onto the calendar as well from some prior step?

(EE#1) Good.

(EE#2) OK.

*Summary of Step 9 comments, suggestions, and possible solutions:*
- Clarify the steps and events that utilize the schedule/calendar feature. Identify these within the tool as well as in the documentation.
- Make the tool customizable so that any step or event can be added to the schedule/calendar when needed (a pull-down menu feature?).

10. **Conduct evaluation**

(IDE) No comments.

(EE#1) Clear. Might be nice if at the end of all the management stuff you could summarize and print a chart which could be stuck on the wall. Could also be that several
of these pieces that are separated could be redesigned to print separately, but tape together as a whole. Some of the report stuff printed at this point may be the answer.

(EE#2) Instead of “Conducting the Evaluation”, as the title suggests, the material in the section appears to emphasize revisions to the evaluation plan. It’s confusing.

Summary of Step 10 comments, suggestions, and possible solutions:
- Make individual steps printable.
- Emphasize (in documentation?) that although the step seems intuitive and self-sufficient, changes may have to be made (thus the possible revisions listed.)

11. Organize and analyze results

(IDE) It would be nice to have examples or samples of data that show when each analysis strategy is useful. For what types of data do you count, and for what types of data do you code? Is there some chart that could be developed here to help people pick appropriate analysis strategies? For instance, if someone has chosen to conduct interviews, does that method typically utilize a specific analysis strategy? Coding?

(EE#1) May need a brief descriptor and/or example in a button or hot link.

(EE#2) This section seems more like “Conducting an Evaluation” than the previous section.

Summary of Step 11 comments, suggestions, and possible solutions:
- Create a link to a chart that describes analysis strategy and appropriate uses of analysis strategies.
- Utilize documentation to clarify the purpose and differences of each step.

12. Report results

(IDE) I like the idea of an auto-report that lays out the common sections needed, and automatically drops-in bulleted lists of items like evaluation questions. I'm not sure the first part of step twelve is as important, generating letters to the stakeholders that a report is available. Assuming stakeholders have participated in creating an evaluation project via your tool, a formal mechanism for sending them a report would not really be necessary. They would already be in the loop.

(EE#1) Should probably default to the stakeholders from Step 1 in the real implementation. Also maybe allow for some stakeholders that were directly contacted as part of the process.

(EE#2) Shouldn’t “Report Results” have a finding section?

Summary of Step 12 comments, suggestions, and possible solutions:
• Clarify that the report (not a letter) can also be generated for stakeholders not directly involved in the evaluation process (i.e., evaluation participants).
• Utilize documentation to explain that a “Results” section would be added automatically. The check-boxed items are not automatic, you may choose to include them or not.

13. Revise instruction accordingly

(IDE) This step should be more generic or expanded to represent summative evaluations as well as formative evaluations. Revision of instruction assumes formative, but another checklist could be generated asking the evaluators to make judgments about the initiative they are evaluating (adopt the practice and state primary reasons why, reject the practice and state primary reasons why, adopt with reservation and state modifications needed, etc.)

(EE#1) Again, the term “component” is not clear to me.

(EE#2) Again, the term “component” bothers me! This step is a bit detached from your evaluative procedure and you need to explain why you need to include it. Maybe the section ought to be called “Recommendations and Conclusions”.

Summary of Step 13 comments, suggestions, and possible solutions:
• Expand tool’s capabilities to include summative evaluation.
• Incorporate the information selected in Step 2 (Define purpose of evaluation) to prompt stakeholders to pass summative judgments relative to the initial reasons for conducting the evaluation.
• Clarify the meaning of “component”. (Here, “component” refers to any part of the instruction as it was presented in the program/product.)
• Utilize documentation to clarify the purpose of each step.
• Rename Step 13 to “Recommendations and Conclusions”?

14. Publish results

(IDE) This is O.K. as a step, but if you try to list journals you will likely have broken links inside of six months. Could be hard to revise and keep up to date. Also, with so many fields represented, it could be hard to list every potential publishing outlet. Most fields have education or teaching-related journals. I would recommend renaming this step to “publicize results,” and add a collaborative feature perhaps that allows stakeholders to decide who to present their results to... “I recommend Journal X...” “I recommend a Web site...” The group could rank order the recommendations and assign various parties to work on publicizing.

(EE#1) A good list of possibilities. May imply that you can go straight from the report to the publication though.

(EE#2) See journals added and consider them. (Three journal titles added to list.)
Summary of Step 14 comments, suggestions, and possible solutions:

- Rename step to “Publicize results”.
- Remove list of journals and links.
- Add collaborative feature that allows stakeholders to input venues for publicizing and allows the venues to be rank ordered.
- Add collaborative feature that assigns various publicizing responsibilities and inputs those responsibilities and deadlines on the calendar.

Expert Reviewer’s Questionnaire Responses

Answers to the questions posed to the expert reviewers (including additional comments and suggestions) are listed below with the corresponding question from the Formative Evaluation Tool Expert Review Instrument (as seen in Appendix E).

Describe your current role as an intended user and/or expert reviewer of this tool.

(IDE) As an instructional designer and evaluator, I would consider using this tool to plan large scale evaluations as part of my departmental duties. I would also recommend the tool to faculty planning to write grants with evaluation components. (Instructional design expert)

(EE#1) Expert reviewer. (Evaluation expert)

(EE#2) Committee member. (Evaluation expert)

INPUTS:

Is the purpose of this tool apparent?

(IDE) For the most part, yes.

(EE#1) No. It is not clear for the intended audience.

(EE#2) Not really. You need some more “general” introductory information to explain what the tool will do.

Summary of comments, suggestions, and possible solutions:

- Create an introductory slide that presents the purpose and utility of the tool.
- Provide documentation to accompany the tool.

Is it clear who the intended user of this tool is?

(IDE) Not formally stated that I remember, so an up-front introduction would be useful.

(EE#1) No. It is not clear for whom this is intended.

(EE#2) Not really. You need more introductory information to explain who the tool is for.
Summary of comments, suggestions, and possible solutions:
- Create an introductory slide that presents the purpose and utility of the tool (for the intended audience.)
- Provide documentation to accompany the tool.

Is the tool appropriate for the intended user? The intended environment?
(IDE) Faculty would most likely not have the time to fully utilize such a tool, unless they are mandated to do so for a grant or large-scale teaching initiative. For most faculty, quicker evaluation tools like Angelo and Cross’ classroom assessment techniques would be more realistic. Evaluation planning tools like this one are useful for those who have to design and carry out large-scale evaluation projects, grant-related evaluations, etc.

(EE#1) Absolutely, but there needs to be more help or support or examples.

(EE#2) Generally yes – but I think some suggestions I made earlier might be incorporated.

Summary of comments, suggestions, and possible solutions:
- Provide documentation to accompany the tool.
- Provide an example(s) in tutorial and/or documentation.

Processes:

Is this tool well designed?
(IDE) Conceptually, yes, but I think additional features would be nice to have and would increase the likelihood that the tool would be used. In general, the features that allow an evaluation planner to carry forward their goals/objectives, evaluation questions, and other items to next steps are very useful for keeping the plan focused and organized; very useful for matching instruments with questions, etc. Some additional features would be nice and might increase the likelihood that the tool will be used.

(EE#1) The interface design is good as are the “Steps.” More support needs to be given for the intended audience.

Summary of comments, suggestions, and possible solutions:
- Provide documentation to accompany the tool.
- Provide information and definitions within the body of the tool to increase the usability.
- Add features such as increased collaboration (i.e. a ranking and prioritization feature for evaluation questions), examples and samples, and optimize features that already exist (such as carrying the objectives through so that criteria and indicators can be established for each.)

Is it visually pleasing to the eye?
(IDE) Not bad, improvements could be made by a graphic designer’s touch, depending on developmental budget.

(EE#1) Yes.

(EE#2) Oh yes.

Summary of comments, suggestions, and possible solutions:
- Conduct expert review with a graphic designer to find ways of strengthening the interface design.

Is the navigation of this tool intuitive?
(IDE) I think it would be given the actual tool, although difficult to tell from a prototype. The navigation guide along the left-hand column is helpful to know where one is located in the overall flow.

(EE#1) Clear, simple, straightforward.

(EE#2) In some parts – but I feel clarification is needed for someone who will see this tool for the first time.

Summary of comments, suggestions, and possible solutions:
- Create tutorial that demonstrates navigation through the tool.
- Provide documentation to accompany the tool.

Is the tool accessible?
This question was interpreted in two ways: disabled/handicapped accessible and technology accessible.

(IDE) No problem accessing the prototype. Unsure of accessibility for the disabled.

(EE#1) Not sure what the question implies here. It is usable from the standpoint of the interface and the steps IF more support is available.

Summary of comments, suggestions, and possible solutions:
- Provide documentation to accompany the tool.
- The question was intended to address technical usability issues – the question needs to be reworded for future evaluation cycles.

Did you have difficulty opening it on your computer? If so, what type of computer are you using?
(IDE) No.

(EE#1) Not really applicable. Static prototype ran smoothly on G3 (OS X) and G4 (OS 9.2.2) Macs.
Does the tool cover all aspects of planning and implementing a formative evaluation?
(IDE) Yes.

(EE#1) Yes.

(EE#2) No.

I had a little difficulty in seeing how it differed from a summative evaluation tool.

Summary of comments, suggestions, and possible solutions:
- Redefine the purpose of the tool – from a formative evaluation tool to an evaluation tool.

Is the tool well designed for collaboration? If not, how could it be improved?
(IDE) I think additional features could help.

(EE#1) To be honest, it is hard to see how this will work collaboratively. It could be as simple as printing out pieces and taping to the wall – could be run from a server – collaboration is not clear.

(EE#2) I don’t understand what “collaboration” means!

Summary of comments, suggestions, and possible solutions:
- Provide documentation to accompany the tool.
- Create an introductory slide that presents the purpose and utility of the tool (specifically address collaboration.)
- Add features such as increased collaboration (i.e. a ranking and prioritization feature for evaluation questions), examples and samples, and optimize features that already exist (such as carrying the objectives through so that criteria and indicators can be established for each.)

Is the tool efficient? (More efficient or less efficient than not having this tool?)
(IDE) Yes, it could aid in planning.

(EE#1) Yes, I would think so. Could save a lot of definitional lookup to see what they are supposed to do. Could keep track of who is doing what so it could be accessed easily and definitely could speed up reporting. Producing consistent reports should, over time, make reading reports more efficient for everyone involved.

(EE#2) More efficient (than not having this tool.)

Does each step in this tool seem necessary for your purposes? If not, explain.
(IDE) I would probably lump steps four and five together. Otherwise, yes, this is similar to evaluations I’ve had to conduct.
(EE#1) I think so but it may be worth thinking of whether it could be customized to let someone skip a step and have that step pulled from the report. I can’t think of which step that might be though to be honest.

(EE#2) Some more than others.

Summary of comments, suggestions, and possible solutions:
- Lump steps four and five together.
- Make tool customizable to enable specific step selection or allow steps to be bypassed.

What would you add or remove to improve the tool?
(IDE) Additional features could bolster the tool.

(EE#1) I would add more support for the intended audience.

(EE#2) Examples – or maybe a ‘full blown’ example as an illustration for the user to see.

Summary of comments, suggestions, and possible solutions:
- Provide documentation to accompany the tool.
- Provide an example(s) in tutorial and/or documentation.
- Add features such as increased collaboration (i.e. a ranking and prioritization feature for evaluation questions), examples and samples, and optimize features that already exist (such as carrying the objectives through so that criteria and indicators can be established for each.)

OUTCOMES:

Does each step contain the components necessary to make the formative evaluation process thorough and efficient (i.e. worksheets, definitions, guidelines)? If not, explain.

(EE#1) Given the additional support I have suggested, yes.

(EE#2) I think more explanatory material needs to be provided.

Summary of comments, suggestions, and possible solutions:
- Provide documentation to accompany the tool.
- Add features such as increased collaboration (i.e. a ranking and prioritization feature for evaluation questions), examples and samples, and optimize features that already exist (such as carrying the objectives through so that criteria and indicators can be established for each.)

Would utilizing this tool save you time when conducting a formative evaluation?
Could this tool be used as a summative evaluation tool? Explain.

(IDE) Yes, I thought it was generic for both formative and summative.

(EE#1) Yes, it frankly ‘feels’ summative.

(EE#2) Probably yes – I have some problems distinguishing it from a summative tool.

Summary of comments, suggestions, and possible solutions:
- Redefine the purpose of the tool – from a formative evaluation tool to an evaluation tool.

OVERALL COMMENTS AND SUGGESTIONS:

(IDE) The title “Evaluation Express” might need to be changed; nothing about evaluation is “express”, rapid, quick… tedious maybe.

(EE#1) I think the prototype suggests an excellent final product if revisions are made.

(EE#2) You need to provide more explanatory introductory materials, examples (generic) that are easy to understand and directions. Each of the 14 sections should have a ‘brief’ explanation of purpose or clarification.

Summary of comments, suggestions, and possible solutions:
- Rename tool.
- Provide documentation to accompany the tool.
- Provide an example(s) in tutorial and/or documentation.
- Introduce each step’s purpose and procedures within the tool.

Meta-Evaluation

When analyzing the comments and suggestions given by the expert reviewers, two things were realized. First, there were some things that were confusing, needed improvement, or clarification regarding the evaluation tool. Second, there were some things that were confusing, needed improvement, or clarification regarding the evaluation itself. Below are some items I would change for the next round of the tool’s evaluation.
The Formative Evaluation Tool Expert Review Instrument (Appendix E) has three questions that need revising. The first question in need of revision, “Describe your current role as an intended user and/or expert reviewer of this tool” was intended to elicit from the expert reviewer his or her role as a potential user (i.e. instructional designer, technology expert, evaluation expert). The question was taken too literally due to poor wording and the responses conveyed the role the expert reviewer had now (i.e. expert reviewer or committee member.) Understanding each expert reviewer’s potential use of the tool (i.e. as an instructional designer) would provide depth to the answers provided in future iterations of evaluation. (In this particular case, these expert reviewers were selected for their strengths in evaluation, instructional design, and programming. This enables me to convey to you the responses of each expert without them having declared their role via their response to that question.) The second question in need of revision, “Is the tool accessible?” was interpreted in two ways, and rightly so. The question was intended to elicit responses regarding user-friendliness. Instead, it led the expert reviewers to conclude I wanted information regarding handicap-accessibility. This question warrants clarification. The third question in need of revision, “Did you have difficulty opening it on your computer? If so, what type of computer are you using?” was intended to catch technological accessibility issues, which wasn’t really necessary at this point in time with only a PowerPoint prototype.

It was also pointed out by one expert reviewer that Stufflebeam’s CIPP model step of “Context” (C) was missing as a category of questions under Step 3: Determine Evaluation Questions. This is also true of the Formative Evaluation Tool Expert Review Instrument (Appendix E). Both of these included questions regarding inputs (I), processes (P), and outcomes (or products - P) Adding questions regarding the context of the tool may prove beneficial.

Future Steps

The comments received from the three expert reviewers need to be weighed and those deemed beneficial will be incorporated into the tool. Most comments and suggestions were not only extremely insightful, but were also quite practical. Below is a synthesized list of the comments and suggestions given that will be incorporated during the next phase of revisions. Many of the specific comments elicited by the reviewers fall under one or more of the bulleted items listed below.

- The tool is more comprehensive than just formative. This needs to be rectified not only in the representation of the tool, but also in the development of the tool.
- Documentation will be provided in two ways:
  - Information will be provided within the framework of the tool utilizing links, rollovers, or pop-up windows, to provide examples, introductions, definitions, samples, worksheets, or a tutorial.
  - Information will accompany the tool in the form of a handbook that explains the purpose and capabilities of each step and the tool as a whole.
- Improve collaboration features
- Add a ranking and prioritization feature for evaluation questions, goals, and objectives.
- Add a feature that assigns various publicizing responsibilities and inputs those responsibilities and deadlines on the calendar.
- Improve organizational features
  - Enable criteria and indicators to be organized under goal/objective subheadings.
  - Enable data sources to be organized under evaluation question subheadings.
  - Enable evaluation questions to be organized under goal/objective subheadings.
- Make the tool more customizable
  - Enable specific step selection or allow steps to be bypassed.
  - Make individual steps printable.
- Create an introductory slide that presents the purpose, audience, and capabilities of the tool
- Add “context” as a category for evaluation questions.
- Rename step 10 from “Conduct Evaluation” to “Collect Data”. This is a much more useful description of what the step intends to accomplish.
- Rename step 13 from “Revise Instruction Accordingly” to “Revise Program Accordingly” to better represent the potential usefulness of the tool.
- Rename step 14 from “Publish Results” to “Publicize results” and the journals and links will be removed in favor of a collaborative feature that will allow stakeholders to input their own venues for publicizing.
- Make the tool available online.

After revising the tool utilizing the feedback received from evaluation experts and an instructional design expert, the tool could next undergo another round of expert reviews, specifically a technical (programming) expert review and a graphic design expert review. Comments would again be weighed and incorporated.

The tool could then go through a round of one-to-one evaluations. These one-to-ones would be conducted with the intended users, practicing instructional designers and developers and perhaps faculty working on a grant (as was suggested by one expert reviewer.) This round of evaluation would capture comments regarding the usability (with new programming and database capabilities in place.) In addition, the previous suggestions that were incorporated would be assessed for their success or failure. Comments would again be weighed and incorporated.

The final round of evaluation would be a field trial. Again, the tool’s intended users, instructional designers and developers and select others, would be utilized for this step. This step would take place in the “real” environment intended for this tool; the users would utilize the tool for their purposes in their spaces. The purpose of this evaluation would be to see that all previous incorporated suggestions are working well and that the tool is appropriate in the user’s environment.
As was mentioned earlier, Dr. Vic Bunderson addressed the development of instructional software in Jared Danielson’s dissertation (1999, p. 86). He proposed that most instructional software requires at least four cycles of implementation, evaluation and revision to achieve its full effectiveness. This tool has already undergone an expert review process that utilized evaluation experts and an instructional design expert. Future evaluation endeavors of a secondary round of expert reviews, one-to-one evaluations, and field trials would provide the minimum of four cycles of which Dr. Bunderson speaks.

Conclusion

The need for a prescriptive, step-by-step evaluation modus operandi prompted the idea behind this developmental dissertation. The actual focus of the dissertation, however, was on the justification behind such a need – the proof that such an accessible tool did not already exist in the complete and procedural manner that the author envisioned. (Granted, numerous books speak volumes about the process of evaluation. The goal, however, was to find or create a tool that would help those active in the field to avoid spending so much time reading books and more time doing the evaluation utilizing the envisioned tool.) Unable to locate such a tool, the first phase was to find or develop a model that was comprehensive of the formative evaluation process. The second phase was to develop the prototype of a tool around that model. The third phase was to have the feasibility, usability, and necessity of such a tool be evaluated by those who know best – evaluation and instructional design experts.

The expert review of the prototype was extremely successful – not because the tool was without fault but because the review process truly brought out what was in need of improvement, what would work well, and what could be done to enhance the tool even more. The findings of this expert review process will allow for the future development of the evaluation tool. The end result to this process is the development of a prototype of an evaluation tool that has its place in the world of evaluation. Although developed for formative evaluation, it is agreed by all experts to be applicable to summative evaluation as well. Although the tool might not be useful to all target users, other potential users were identified in the expert review process.

The ultimate success or failure of such a tool will depend upon the thorough formative evaluation of the evaluation tool. This process was outlined previously and could be done with the confidence that the 14 steps the tool is based upon are selected from a thorough review of the literature. The expert review of the evaluation tool provides confidence in the content and structure of the prototype. The process by which the prototype was researched, developed, and evaluated is the issue at hand. This author feels the process was not only successful, but also rewarding.
REFERENCES


Hanson, G. R. (1978). Conclusions and further resources. *New Directions for Student Services*(1), 97-100.


APPENDIX A: FORMATIVE EVALUATION TOOL PROTOTYPE

1. Clarify product/program goals and objectives
2. Define purpose of evaluation
3. Determine evaluation questions
4. Plan the evaluation
5. Select criteria, indicators, and data sources
6. Develop instruments
7. Submit evaluation proposal
8. Make revisions (if necessary)
9. Schedule appointments for evaluation
10. Conduct evaluation
11. Organize and analyze results
12. Report results
13. Revise instruction accordingly
14. Publish results

Evaluation Express

A Collaborative Formative Evaluation Tool for Instructional Designers, Developers, and Evaluators

BEGIN
Step 1: Clarify Product/Program Goals and Objectives

Goals:

Objectives:

[Blank lines for input]
Step 1: Clarify Product/Program Goals and Objectives

List all stakeholders (including, but not limited to: the funding agency, management, instructional designers, developers, evaluation participants, and any other persons involved.)

Name: __________________________
Company/Title: __________________________
Contact Information: __________________________

Name: __________________________
Company/Title: __________________________
Contact Information: __________________________

Name: __________________________
Company/Title: __________________________
Contact Information: __________________________

Name: __________________________
Company/Title: __________________________
Contact Information: __________________________

Have all relevant stakeholders give input into clarifying the goals and objectives. An evaluation will reap the most rewards and be most accepted when its focus is clarified by all involved.
Step 2: Define Purpose of Evaluation

Formative vs. Summative evaluation: Relates to the evaluator's intentions – do you wish to improve a program/product (formative) or to judge a program's/product's effectiveness (summative)?

Process vs. Outcome evaluation: Relates to the part of the program/product being studied – do you wish to examine what happens while the program/product is in progress (process) or to determine what participants gain/lose as a result of the program/product (outcome)?

Overt vs. Covert reasons for evaluation: Below are overt reasons for conducting an evaluation. There are often many covert reasons for conducting an evaluation - good public relations, fulfilling a requirement, postponing a decision, or other reasons. Identify all reasons for this evaluation.

☐ Decision-making
  ☐ Asset allocation
  ☐ Continuation/Discontinuation
  ☐ Corrections
  ☐ Cutting back/Expanding
  ☐ Choosing the best option
  ☐ Other____________________

☐ Program/Product improvement
  ☐ Outcomes and results
  ☐ Objectives attainment
  ☐ Other____________________

☐ Policy improvement
  ☐ Other____________________

☐ Organizational Learning
  ☐ Program documentation
  ☐ Feedback to practitioners
  ☐ Highlighting goals
  ☐ Other____________________
Step 3: Determine Evaluation Questions

What do you hope to determine from this evaluation?
(Note: Have all relevant stakeholders give input into clarifying the evaluation questions.)

________________________________________________________________________
________________________________________________________________________
________________________________________________________________________
________________________________________________________________________
________________________________________________________________________
________________________________________________________________________
________________________________________________________________________
________________________________________________________________________

Click here for examples of evaluation questions.
Step 3: Determine Evaluation Questions

Database: Sample evaluation questions

(Click to add a question to your list)

**Inputs**
- Is there a market for this (text, tool, other)?
- Have the user’s incoming capabilities been properly identified?
- Was the (text, tool, other) a proper match for the audience’s (attention span, age, other)?

**Processes**
- What is (liked/disliked) about the (text, tool, other)?
- Are the procedures outlined appropriate for the intended outcome?
- Is the (text, tool, other) (intuitive, readable, usable, other)?
- How (when, where) would users utilize this (text, tool, other)?
- Does the (text, tool, other) address all the relevant information in enough detail?

**Outcomes**
- Is ______ more cost-effective than ________?
- Is ______ better than ________?
- Were the stated objectives met?
- Did user’s (attitudes, test scores, other) improve as a result of the intervention?
- How could the (text, tool, other) be improved?
Step 4: Plan the Evaluation

Evaluation Procedure

☐ **Pilot test**: A trial run of a product/program with a few participants for immediate feedback.

☐ **Rapid prototype**: An abbreviated version of a product/program is implemented directly into the field, where intended users provide feedback and input for revisions.

☐ **Expert review**: An evaluation of the product (instruction, program) from an expert’s perspective (content, technical, instructional design, etc.)

☐ **One-to-one**: An evaluation with one intended user, but not necessarily in the intended environment. This step allows for major problems to be identified.

☐ **Focus group**: Similar to the one-to-one, but several intended users are brought together to discuss the product/program.

☐ **Small group**: An evaluation with a group of intended users in an environment similar to the intended environment. Revisions to problems identified earlier are tested and other problems are identified.

☐ **Field trial (field test)**: An evaluation with a group of intended users in the intended environment. The evaluator has no role in this procedure except as observer. Revisions to problems identified earlier are tested and final revisions are identified.

☐ **Observation**

☐ **Other**

---

[Links and buttons for next page]
Step 5: Select Criteria, Indicators, and Data Sources

Criteria: The standard of judgment (Example: Passing a test)

Indicators: An index, gauge, or value that determines whether or not the criteria have been met (Example: 65% and above equals passing)

Data Sources: The documents or primary reference works that supply information.

Examples:

☐ Administrative records
☐ Existing data
☐ Formal interviews
☐ Informal interviews
☐ Observation Logs
☐ Program records
☐ Questionnaires
☐ Tests (informational, skill, interpretational, attitudinal, projective, other)
☐ Simulation games
☐ Diaries
☐ Clinical exams
☐ Document review (financial records, media stories, photos, videos, other)
☐ Other

☐ Other
Appendix A. Formative Evaluation Tool Prototype

Step 5: Select Criteria, Indicators, and Data Sources

Data Sources selected:

**Formal interviews**
Person(s) responsible
Instrument Needed? Y □ N □ Deadline for instrument creation __________
Deadline for completion ________________

**Observations**
Person(s) responsible
Instrument Needed? Y □ N □ Deadline for instrument creation __________
Deadline for completion ________________

**Program records**
Person(s) responsible
Instrument Needed? Y □ N □ Deadline for instrument creation __________
Deadline for completion ________________

**Other**
Person(s) responsible
Instrument Needed? Y □ N □ Deadline for instrument creation __________
Deadline for completion ________________

---

1. Clarify product/program goals and objectives
2. Define purpose of evaluation
3. Determine evaluation questions
4. Plan the evaluation
5. Select criteria, indicators, and data sources
6. Develop instruments
7. Submit evaluation proposal
8. Make revisions (if necessary)
9. Schedule appointments for evaluation
10. Conduct evaluation
11. Organize and analyze results
12. Report results
13. Revise instruction accordingly
14. Publish results
Step 6: Develop Instruments

Evaluation Instruments
(Click on highlighted instruments for a sample.)

- Interview Questionnaire
- Observation Log
- Survey
- Field Notes Log/Journal
- Tests

Click here for sample instrument questions
# Observation Log

<table>
<thead>
<tr>
<th>Learner (Teacher, Trainer, Participant, Other):</th>
<th>Time: ____________________________</th>
</tr>
</thead>
<tbody>
<tr>
<td>Comments/Notes:</td>
<td>Location: ________________________</td>
</tr>
<tr>
<td></td>
<td>Who is being observed:</td>
</tr>
<tr>
<td></td>
<td>What is being observed:</td>
</tr>
<tr>
<td></td>
<td>What to look for:</td>
</tr>
<tr>
<td>Behavior:</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Field Notes

Time: ___________________  Location: ___________________

Who is being observed: __________________________________________

What is being observed:
________________________________________________________________
________________________________________________________________
________________________________________________________________

What to look for:
________________________________________________________________
________________________________________________________________
________________________________________________________________

Notes:
________________________________________________________________
________________________________________________________________
________________________________________________________________
________________________________________________________________
________________________________________________________________
________________________________________________________________

Back
Step 6: Develop Instruments

Database: Sample instrument questions

(Click to add a question to your list)

Inputs
- Is the (text, software, other) a proper match for your (attention span, age, environment)?
- Did your computer(s) encounter technical problems when running the software?

Processes
- Is the (text, software, other) (intuitive, readable, usable, other)?
- How (when, where) would you utilize this (text, software, other)?
- Does the (text, software, other) address all the relevant information in enough detail?

Outcomes
- Do you feel your (attitudes, test scores, other) improved as a result of the intervention?
- How can the (text, software, other) be improved?
- Was the (text, software, other) readily accepted by (you, your students, your school system)?
Step 7: Submit Evaluation Proposal

Determine time schedule
Determine budget

Submit for feedback

- evaluation plan
- time schedule
- budget
- instruments

By: ________________________________
Deadline for feedback: _______________

Submit to: (see List of Stakeholders)

______________________________
______________________________
______________________________
______________________________
______________________________
______________________________
______________________________
______________________________

[Print]
Step 7: Submit Evaluation Proposal

Time Schedule

The time needed for your evaluation varies depending on several factors: the number of steps you intend to incorporate, the type of step you incorporate, the numbers of participants you select, your participants’ schedules, your budget, the turn-around time for incorporating changes between each step, and your overall timeframe. Below is a timetable that may help you get an idea of the time involved.

<table>
<thead>
<tr>
<th>Formative evaluation step</th>
<th>Approximate # of weeks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conduct expert reviews</td>
<td>2</td>
</tr>
<tr>
<td>Analyze expert reviews</td>
<td>1</td>
</tr>
<tr>
<td>Report on expert reviews</td>
<td>1</td>
</tr>
<tr>
<td>Incorporate changes</td>
<td>2</td>
</tr>
<tr>
<td>Conduct one-to-one interviews</td>
<td>1</td>
</tr>
<tr>
<td>Analyze one-to-one interviews</td>
<td>1</td>
</tr>
<tr>
<td>Report on one-to-one interviews</td>
<td>1</td>
</tr>
<tr>
<td>Incorporate changes</td>
<td>2</td>
</tr>
<tr>
<td>Conduct small group(s)</td>
<td>1</td>
</tr>
<tr>
<td>Analyze small group(s)</td>
<td>1</td>
</tr>
<tr>
<td>Report on small group(s)</td>
<td>1</td>
</tr>
<tr>
<td>Incorporate changes</td>
<td>2</td>
</tr>
<tr>
<td>Conduct field trial(s)</td>
<td>1</td>
</tr>
<tr>
<td>Analyze field trial(s)</td>
<td>1</td>
</tr>
<tr>
<td>Report on field trial(s)</td>
<td>1</td>
</tr>
<tr>
<td>*Incorporate changes</td>
<td>2</td>
</tr>
<tr>
<td>Final evaluation report</td>
<td>1</td>
</tr>
</tbody>
</table>

* Denotes activities that occur in conjunction with other ongoing activities.
Step 7: Submit Evaluation Proposal

**Budget**

The evaluation budget is typically no less than 2% and no more than 10% of the total project budget.  

TOTAL ESTIMATED BUDGET  

There are many things to be considered when determining a budget.  

(No specific cost data is currently available for the categories listed below.)

**WAGES**

Salaries  Hourly

Training

**OVERHEAD**

Rent  Utilities

**EQUIPMENT**

Office supplies  FAXs

Copier  Copies

Computers  Software

**MISCELLANEOUS**

Travel expenses

**OTHER**

TOTAL

---

PRINT  BACK  NEXT
Step 8: Make Revisions (if necessary)

Revisions to evaluation proposal

Component to be revised: __________________________
Person(s) Responsible: __________________________
Deliverables: ____________________________________
Due date: __________________________

Return to Step 4: Plan the evaluation? Y □ □ N □

Component to be revised: __________________________
Person(s) Responsible: __________________________
Deliverables: ____________________________________
Due date: __________________________

Return to Step 4: Plan the evaluation? Y □ □ N □

Component to be revised: __________________________
Person(s) Responsible: __________________________
Deliverables: ____________________________________
Due date: __________________________

Return to Step 4: Plan the evaluation? Y □ □ N □
Step 9: Schedule Appointments for Evaluation

<table>
<thead>
<tr>
<th>Procedure</th>
<th>Who (Evaluator and…)</th>
<th>Where</th>
<th>Instrument(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expert Review</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>One-to-one</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Small Group</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Field Test</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1. Clarify product/program goals and objectives
2. Define purpose of evaluation
3. Determine evaluation questions
4. Plan the evaluation
5. Select criteria, indicators, and data sources
6. Develop instruments
7. Submit evaluation proposal
8. Make revisions (if necessary)
9. Schedule appointments for evaluation
10. Conduct evaluation
11. Organize and analyze results
12. Report results
13. Revise instruction accordingly
14. Publish results
### Step 9: Schedule Appointments for Evaluation

**Database: Calendar**

<table>
<thead>
<tr>
<th>Sunday</th>
<th>Monday</th>
<th>Tuesday</th>
<th>Wednesday</th>
<th>Thursday</th>
<th>Friday</th>
<th>Saturday</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>11</td>
<td>12</td>
<td>13</td>
</tr>
<tr>
<td>14</td>
<td>15</td>
<td>16</td>
<td>17</td>
<td>18</td>
<td>19</td>
<td>20</td>
</tr>
<tr>
<td>21</td>
<td>22</td>
<td>23</td>
<td>24</td>
<td>25</td>
<td>26</td>
<td>27</td>
</tr>
<tr>
<td>28</td>
<td>29</td>
<td>30</td>
<td>31</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Step 10: Conduct Evaluation

Although this step seems very straightforward, there are things to keep in mind.

- Instrument revisions – When utilizing instruments you may find questions are confusing or unnecessary, or you may feel the need to add a question.
- Schedule changes – Times can change, people involved can change. Update the evaluation plan, time schedule, and calendar as necessary.

Revision made: 
By: ___________________________ Date: ___________________________

Revision made: 
By: ___________________________ Date: ___________________________

Revision made: 
By: ___________________________ Date: ___________________________

Revision made: 
By: ___________________________ Date: ___________________________

Revision made: 
By: ___________________________ Date: ___________________________

Revision made: 
By: ___________________________ Date: ___________________________

Revision made: 
By: ___________________________ Date: ___________________________
Step 11: Organize and analyze results

Information should be organized around the questions that the evaluation sought to answer (see Step 3).

Analysis strategies:
- Describing
- Counting
- Coding
- Clustering
- Comparing
- Finding covariation
- Telling the story
Step 12: Report Results

Stakeholders to receive copy of report:

☐ Name: ______________________
  Company/Title: ______________________
  Contact Information: ______________________

☐ Name: ______________________
  Company/Title: ______________________
  Contact Information: ______________________

☐ Name: ______________________
  Company/Title: ______________________
  Contact Information: ______________________

☐ Name: ______________________
  Company/Title: ______________________
  Contact Information: ______________________

☐ Name: ______________________
  Company/Title: ______________________
  Contact Information: ______________________

☐ Name: ______________________
  Company/Title: ______________________
  Contact Information: ______________________

NEXT
Step 12: Report Results

Include:

- Program/product goals and objectives
- Purpose of evaluation
- Evaluation questions
- Evaluation proposal (including evaluation plan, time schedule, budget, and instruments)
Step 13: Revise instruction accordingly

Component to be revised:________________________________________________________________________

Person(s) Responsible:________________________________________________________________________

Deliverables:__________________________________________________________________________________

Due date:________________________________________

Return to Step 9: Schedule appointments for evaluation? Y □  N □

Component to be revised:________________________________________________________________________

Person(s) Responsible:________________________________________________________________________

Deliverables:__________________________________________________________________________________

Due date:________________________________________

Return to Step 9: Schedule appointments for evaluation? Y □  N □

Component to be revised:________________________________________________________________________

Person(s) Responsible:________________________________________________________________________

Deliverables:__________________________________________________________________________________

Due date:________________________________________

Return to Step 9: Schedule appointments for evaluation? Y □  N □

Component to be revised:________________________________________________________________________

Person(s) Responsible:________________________________________________________________________

Deliverables:__________________________________________________________________________________

Due date:________________________________________

Return to Step 9: Schedule appointments for evaluation? Y □  N □
Step 14: Publish results

Database: Journals

Topical Publications

Trade/Industry Publications

Educational/Instructional Design Publications

THE Journal Online
Journal of Computing in Higher Education
Educational Researcher (ER)
American Educational Research Journal (AERJ)
Review of Educational Research

Evaluation Publications

Practical Assessment, Research, and Evaluation
Evaluation Review
Evaluation
Educational Research and Evaluation
Assessment and Evaluation in Higher Education
American Journal of Evaluation
APPENDIX B: EVALUATION TOOL FLOWCHART

1) Clarify product/program goals and objectives (Slide 2)

2) Define purpose of evaluation (Slide 4)

3) Determine evaluation questions (Slide 5)

4) Plan the evaluation (Slide 7)

5) Select criteria, indicators, and data sources (Slide 8)

6) Develop instruments (Slide 10)

7) Submit evaluation proposal (Slide 14)

List relevant stakeholders (Slide 3)

Examples of evaluation questions (Slide 6)

Data sources selected (Slide 9)

Observation Sheet Example (Slide 11)

Sample questions (Slide 13)

Field Notes Example (Slide 12)

Time schedule (Slide 15)

Budget (Slide 16)
1) Make revisions (if necessary) (Slide 17)

2) Schedule appointments for evaluation (Slide 18)

3) Conduct evaluation (Slide 20)

4) Organize and analyze results (Slide 21)

5) Report results (Slide 22)

6) Publish results (Slide 25)

7) Components to include in report / Generate Report (Slide 23)

8) Make revisions (if necessary) (Slide 17)

9) Schedule appointments for evaluation (Slide 18)

10) Conduct evaluation (Slide 20)

11) Organize and analyze results (Slide 21)

12) Report results (Slide 22)

13) Revise instruction accordingly (Slide 24)

14) Publish results (Slide 25)

Key:
- Primary level of step
- Secondary level of step
- Link to outside database
APPENDIX C: INTRODUCTION TO THE EXPERT REVIEWER

Evaluation is an integral part of the instructional design process. Most instructional design models include evaluation as a step, whether formative or summative, and whether it is at the conclusion of the process or an ongoing loop. Despite this, evaluation has long been held in poor esteem. It is seen as something done to a product or program, rather than being seen as something done for the product or program. The very nature of summative evaluation may have led to this notion, as it is conducted after the fact and often by an outside entity. Employing formative evaluation is one means of determining problems internally and prior to mass distribution. Formative evaluation is often conducted informally, if it is conducted at all. The lack of know-how, lack of documentation, and lack of agreed upon procedures for how and when to implement findings often renders formative evaluation empty of value.

While working with NASA to formatively evaluate their “Why?” Files program, it occurred to me that there is very little evaluation know-how, even among those very experienced in education. Furthermore, there was an abundance of literature regarding evaluation, and specifically formative evaluation, but no one source seemed to convey the essentials quickly or concisely. To clarify, among the many wonderful books out there (and there are many), those engaged in the field of instructional design rarely have the time to stop and read an entire book on evaluation (or any other topic), let alone develop and incorporate what they have learned.

This fueled my desire to find (or develop) something more comprehensive, more prescriptive, and less time consuming than reading a book (or books). My literature review was an extensive review of evaluation models, evaluation approaches, evaluation designs, and evaluation tools. I found that (besides books) there were no tools that could walk a user through the evaluation process. Having proof (through the literature review) that no such tool existed, I now had an argument for developing such a tool. The first step towards developing an evaluation tool was determining what model or set of steps would be the basis for the tool. Again, the literature review pointed to many existing models. Many models had many similar steps; almost all models covered the same basics. However, it was my opinion that no one model was comprehensive enough (especially not for a novice to evaluation) and no one model was prescriptive enough to translate into
a tool that guided the user step-by-step. In lieu of finding such a model, one was created. The model is backed by the many models from which its steps are similar, but pulls all steps from many models to be more comprehensive.

The next step in tool development was to decide what form such a tool would take. A software-based tool (versus paper-based) was what I had envisioned. This would enable collaboration. The next questions, and the questions most pivotal in doing this developmental dissertation, were 1) what would I use to develop this tool, and 2) to what extent did this tool need to be developed? Both of these questions would determine the amount of time required in terms of development. Developing an evaluation tool in Director would enable a lot more capabilities than a PowerPoint prototype, but would also require a greater amount of learning on my end. Developing the tool fully (creating all possible databases that it could utilize) would take a lot more time than creating an example to give a general idea of potential capabilities. Thus, the prototype you are being asked to evaluate has been developed in PowerPoint, thus allowing minimal functionality but enabling one to see possible capabilities (i.e. links work). Further, only examples have been created of potential databases. For example, a database of hundreds of potential instrument questions is exemplified by one slide of such questions.

The purpose of your expert review is to determine whether or not a tool such as this is needed, whether it has been fully developed, whether or not it meets the needs of those who would use it, and to get comments and suggestions for the tool’s improvement. Vic Bunderson commented that instructional software will go through a minimum of four cycles of implementation, evaluation and revision before achieving its full effectiveness. This is only the first round – an expert review of a prototype. Each of you has been selected for your prominence and/or active participation in the field of evaluation or instructional design. Your comments and suggestions are welcomed.

In addition to receiving this introduction/explanation of the history behind the development of this tool, you have also received a flowchart of the tool, a description of the capabilities at each step (since many capabilities aren’t activated in this more static prototype) with room for comments, a questionnaire, and finally, the references from the literature review that have shaped the development of this tool. Of course, also included is a CD-ROM that contains the tool itself. My request is that all comments and
suggestions, including the questionnaire, be returned to me by Friday, March 15. My contact information is listed below in case there are problems or questions. I sincerely thank you for your participation as an expert reviewer in my developmental dissertation.
APPENDIX D: SOFTWARE CAPABILITIES

The benefit of a software-based tool (verses paper-based) is the ability to collaborate. The steps in this tool are accessible to any person granted password access: the instructional designer(s), developer(s), evaluator(s), or stakeholders. The calendar (seen in Step 9) is the organizational center for the tool. All deadlines and appointments (entered throughout the tool) are automatically placed on the calendar. This calendar (accessible by all with password access) is modifiable and printable. As you work your way through the tool, keep in mind that the tool is not only meant to guide the user through the evaluation process, but also to be a collaborative and organizational tool in that process.

1. **Clarify product/program goals and objectives**

   Determining the goals and objectives of the product/program being evaluated helps to discern what the product/program was designed to accomplish. This helps when formulating the evaluation questions (Step 3) as well as the instrument questions (Step 6). Identifying all stakeholders involved in the product/program and the evaluation (secondary slide) helps to ensure that the necessary people have input. This list of stakeholders is also utilized later when determining who receives an evaluation proposal (Step 7) and when determining who receives a copy of the final report (Step 12).

   Comments:______________________________________________________________
   __________________________________________________________________________
   __________________________________________________________________________
   __________________________________________________________________________
   __________________________________________________________________________

2. **Define purpose of evaluation**

   Determining why the evaluation is being conducted not only sheds light on what stakeholders hope to gain from the evaluation, but also helps in formulating evaluation questions (Step 3). Typical reasons for conducting an evaluation are listed with check boxes so you can identify your evaluation’s purpose(s).

   Comments:________________________________________________________________________
   ______________________________________________________________________________
   ______________________________________________________________________________
   ______________________________________________________________________________
   ______________________________________________________________________________

3. **Determine evaluation questions**
This step enables all relevant stakeholders to see the list of evaluation questions, and to have input into the creation of the list. A link at the bottom of the list of evaluation questions takes you to a database of sample evaluation questions. Clicking a check box automatically adds that question to the list, where it can be amended if necessary.

Comments:______________________________________________________________
_______________________________________________________________________
_______________________________________________________________________
_______________________________________________________________________
_______________________________________________________________________

4. Plan the evaluation
The first step in planning the evaluation is to determine which procedures you will implement. Listed are common evaluation procedures and their definitions. A check box is provided next to each for easy selection. The procedures selected will automatically show up in Step 7 (Time schedule) and Step 9 (Appointments sheet).

Comments:______________________________________________________________
_______________________________________________________________________
_______________________________________________________________________
_______________________________________________________________________
_______________________________________________________________________

5. Select criteria, indicators, and data sources
Criteria, indicators and data sources are first defined. A list of possible data sources is provided, along with a check box next to each so that you can flag possible sources of information for your evaluation. A secondary slide comes up that lists only the data sources that you have checked. Here you can specify what instruments you plan to use, make, or retrieve, from where they would be retrieved, who is responsible for retrieving or making them, and by what deadline. The deadline (and person responsible) is automatically added to the calendar.

Comments:______________________________________________________________
_______________________________________________________________________
_______________________________________________________________________
_______________________________________________________________________
_______________________________________________________________________

6. Develop instruments
This step lists evaluation instruments and allows the user to print standard instruments (i.e. observation logs and field notes), rather than recreating the wheel. A database of sample instrument questions is accessible from this step that could be used or modified for instrument creation. The questions would be categorized by input, processes, and outcomes.

Comments:
________________________________________________________________________
________________________________________________________________________
________________________________________________________________________
________________________________________________________________________

7. Submit evaluation proposal
This step addresses the time schedule, the budget, and the submission of the evaluation plan, time schedule, budget, and instruments for feedback. The time schedule (secondary slide) will vary widely depending on the pieces of evaluation that are selected, the number of participants, and individual schedules. A rough guideline of time requirements will be outlined. Only the evaluation steps selected in Step 4 will be included. (When looking at the information provided, assume the user selected the expert review, one-to-one, small group and field trials procedures.) The budget (secondary slide) will also vary widely depending on the same factors, and many others. A rough guideline of budgetary factors and considerations will be outlined. These of course, could be vastly improved upon and expanded in the future. Submitting the evaluation plan, time schedule, budget, and instruments for feedback is a crucial step to making sure that all stakeholders are satisfied with the path the evaluation will take, the time it will take, and the amount of money it will take. A list of persons to receive this information would be made at this step. The deadline(s) for feedback would automatically be placed on the calendar.

Comments:
________________________________________________________________________
________________________________________________________________________
________________________________________________________________________
________________________________________________________________________

8. Make revisions (if necessary)
What needs revising, who is responsible for the revisions, and the deadline for the revisions are all determined in this step. The deadline is automatically placed on the calendar. This step prompts the user to return to Step 4 to rethink the evaluation plan. If that option is selected, the tool would skip back to Step 4.
9. Schedule appointments for evaluation
The table of scheduled appointments is a quick and easy way to see what is being done with whom, when, and what instruments are needed. Only the evaluation steps selected in Step 4 will be included under “Procedure”. The information in this table will be automatically entered into the calendar. The calendar is perhaps the most important organizational component within this evaluation tool. It is available to all who have password access. Each deadline and each appointment is placed on the calendar automatically (if entered in at the appropriate step) and any relevant information, deadline, or appointment can be entered directly by going to the calendar. The calendar is printable, and every month is only a click away.

Comments:
________________________________________________________________________
________________________________________________________________________
________________________________________________________________________

10. Conduct evaluation
As straightforward as this step seems, there are always adjustments that may be necessary. The instruments may need adjustment if you find a question is confusing, or a question may need to be added. Time schedules may need adjustment depending on the overall time schedule of the project, or depending on individuals’ time schedules which are bound to change. This step allows for the specification of these changes and the party responsible for the changes. This information is automatically added to the calendar.

Comments:
________________________________________________________________________
________________________________________________________________________
________________________________________________________________________

11. Organize and analyze results
This step requires the information gathered from the data sources selected to be organized around the questions that the evaluation sought to answer. Several common analysis techniques are listed in this step.

Comments:______________________________________________________________
________________________________________________________________________
________________________________________________________________________
________________________________________________________________________
________________________________________________________________________

12. Report results
The list of stakeholders who were identified in Step 1 resurfaces here as a checklist. The report would automatically be generated with the correct letterhead to the persons whose names are checked on this list. The report can be automatically generated (secondary slide) to include many of the earlier steps including program/product goals and objectives (Step 1), evaluation purpose(s) (Step 2), evaluation questions (Step 3), and the evaluation proposal (Step 7) which includes the evaluation plan, time schedule, budget, and instruments. The report is complete when a “Results” section is added. Any part of the report can be revised manually after being automatically generated.

Comments:______________________________________________________________
________________________________________________________________________
________________________________________________________________________
________________________________________________________________________
________________________________________________________________________

13. Revise instruction accordingly
The components of instruction needing revision should be listed in this step. Responsible persons or departments would be identified, as well as the deliverables. This step in the evaluation tool would also allow the revision and due date to be placed directly on the calendar, which is accessible by all who are coordinating on the project. This step also prompts the user to return to Step 9 for further evaluation. If that option is selected, the tool would skip back to Step 9 so revisions could be made to the original evaluation plan (i.e. an abbreviated secondary evaluation).

Comments:______________________________________________________________
________________________________________________________________________
14. Publish results
Publishing the results of an evaluation is an important and often overlooked step. This step provides access to a database of journals in which you might publish your findings. The database of journals provides (at a minimum) the journals’ name (categorized by topic) and an on-line address, if available. The database of journals might also include the journals’ mission and their publishing guidelines. The database could be made searchable by topic or by journal name.

Comments:________________________________________________________________________
________________________________________________________________________
________________________________________________________________________
________________________________________________________________________
________________________________________________________________________
APPENDIX E: FORMATIVE EVALUATION TOOL EXPERT REVIEW INSTRUMENT

Describe your current role as an intended user and/or expert reviewer of this tool.
________________________________________________________________________
________________________________________________________________________
________________________________________________________________________

INPUTS:

Is the purpose of this tool apparent?
________________________________________________________________________
________________________________________________________________________
________________________________________________________________________

Is it clear who the intended user of this tool is?
________________________________________________________________________
________________________________________________________________________
________________________________________________________________________

Is the tool appropriate for the intended user? The intended environment?
________________________________________________________________________
________________________________________________________________________
________________________________________________________________________

PROCESSES:

Is this tool well designed?
________________________________________________________________________
________________________________________________________________________
________________________________________________________________________

Is it visually pleasing to the eye?
________________________________________________________________________
________________________________________________________________________
________________________________________________________________________

Is the navigation of this tool intuitive?
Is the tool accessible?

Did you have difficulty opening it on your computer? If so, what type of computer are you using?

Does the tool cover all aspects of planning and implementing a formative evaluation?

Is the tool well designed for collaboration? If not, how could it be improved?

Is the tool efficient? (More efficient or less efficient than not having this tool?)

Does each step in this tool seem necessary for your purposes? If not, explain.

What would you add or remove to improve the tool?
OUTCOMES:

Does each step contain the components necessary to make the formative evaluation process thorough and efficient (i.e. worksheets, definitions, guidelines)? If not, explain.

________________________________________________________________________
________________________________________________________________________
________________________________________________________________________

Would utilizing this tool save you time when conducting a formative evaluation?
________________________________________________________________________
________________________________________________________________________
________________________________________________________________________

Could this tool be used as a summative evaluation tool? Explain.
________________________________________________________________________
________________________________________________________________________
________________________________________________________________________

OVERALL COMMENTS AND SUGGESTIONS:
________________________________________________________________________
________________________________________________________________________
________________________________________________________________________
________________________________________________________________________
________________________________________________________________________
________________________________________________________________________