

This section includes shorter (e.g., 10–15 double-spaced manuscript pages or less) papers describing methods and techniques that can improve evaluation practice. Method notes may include reports of new evaluation tools, products, or services that are useful for practicing evaluators. Alternatively, they may describe new uses of existing tools. Also appropriate for this section are user-friendly guidelines for the proper use of conventional tools and methods, particularly for those that are commonly misused in practice.

Evaluation Checklists: Practical Tools for Guiding and Judging Evaluations

DANIEL L. STUFFLEBEAM

ABSTRACT

This article describes a project designed to provide evaluators, their clients, and other stakeholders with checklists for guiding and assessing formative and summative evaluations. The checklists pertain to program, personnel, and product evaluations, and reflect different conceptualizations of evaluation. They are constructed for use in planning, contracting, conducting, reporting, and judging evaluations. The checklists, along with papers on the logic and methodology of checklists and guidelines for developing checklists, are conveniently available on the Western Michigan University Evaluation Center's Web site www.wmich.edu/evalctr/checklists/.

INTRODUCTION

A checklist is a list for convenient checking and reference. An evaluation checklist is a list for guiding an enterprise to success (formative orientation) and/or judging its merit and worth (summative orientation).

Sound checklists can have profound evaluative applications. Familiar examples are evaluations of behavior against moral codes like the Ten Commandments; evaluations of legal matters against the U.S. Bill of Rights and other constitutional provisions; evaluations

Daniel L. Stufflebeam • Harold and Beulah McKee Professor of Education and Director of The Evaluation Center, Western Michigan University, Kalamazoo, MI 49008; Tel: (616) 387-5895; E-mail: daniel.stufflebeam@wmich.edu.

American Journal of Evaluation, Vol. 22, No. 1, 2001, pp. 71–79. All rights of reproduction in any form reserved.
ISSN: 1098-2140 Copyright © 2001 by American Evaluation Association.

of residents' actions against lists of community covenants; evaluations of construction projects against electrical, structural, plumbing, and other codes; evaluations of hospitals and colleges against accreditation criteria; and evaluations of evaluations against professional standards. Such checklists can provide valuable assistance to evaluators, their clients, and other stakeholders as they plan, conduct, and judge evaluations.

This article's purpose is to report on a project designed to provide evaluators with a growing set of useful evaluation checklists. The project is housed in the Western Michigan University Evaluation Center and is funded through Arlen Gullickson's National Science Foundation-funded Materials Development, Training, and Support Services (MTS) project. My colleagues in the checklist development project are Michael Scriven and Lori Wingate.

We are collecting, developing, and disseminating checklists designed to help evaluators improve all important aspects of their evaluation work. Individually and in combination, these checklists provide guidance for planning and contracting for evaluations; collecting, organizing, analyzing, synthesizing, and reporting information; managing evaluation operations; and arriving at judgments of merit and worth. They may be used independently or in combination, and some are keyed to professional standards for evaluations.

In summarizing this project, I will divide my remarks into two parts. First, I will provide some background by relating how and why I got involved in developing and using evaluation checklists. Second, I will characterize the checklists so far included in our project within a conceptual framework. In closing, I will consider what additional checklists are needed.

SOME PERSONAL HISTORY CONCERNING THE DEVELOPMENT OF EVALUATION CHECKLISTS

I first became sensitized to the value of evaluation checklists early in my evaluation career. Some personal history may help the reader understand how and why I came to develop, employ, and teach others to use evaluation checklists.

In the early 1960s, I was directing Ohio State University's Test Development Center and conducting research on measurement. I thought that was how I would spend my career. When the War on Poverty began in 1965 with its associated evaluation requirements, suddenly I was assigned to lead Ohio State's efforts to help schools meet federal evaluation requirements. It took me only a short time to realize that almost everything I had learned about experimental design, measurement, and statistics—from Ben Winer, Norbit Downey, Bill Owens, Julian Stanley, and other renowned methodologists—was largely irrelevant to evaluating new, heavily funded but ill-defined projects within the turmoil of poverty-stricken schools.

Gradually, I began to evolve an approach to evaluation that seemed to work in the dynamic, confusing world of school-based innovations. This approach was grounded in interactions with school personnel and other stakeholders. It was directed toward the design of evaluations that would address stakeholders' evaluative questions; provide them a flow of timely, relevant information; help them make informed decisions; and provide an information base for reports to sponsors. The ultimate aims were to help schools, through creditable evaluations, to both improve projects and meet accountability requirements.

School personnel and their federal sponsors welcomed and supported this approach. The U.S. Office of Education even negotiated with my university to have me become the lead advisor on the federal-level evaluation of the multibillion-dollar Elementary and Secondary

Education Act program. During 1965 and 1966, I spent two days a week in Washington on the federal government assignment and three days a week in Columbus directing the Ohio State evaluation work. This provided both large-scale and close-up views of the issues in evaluating the War on Poverty federal education programs.

At Ohio State, I engaged a cadre of very able graduate students to assist in carrying through my evaluation assignments. These students included future evaluation luminaries, such as Tom Owens and Blaine Worthen. The students began pressing me to explain what exactly I was doing in designing project evaluations. Answering their questions proved difficult. In retrospect, I was developing and exercising a kind of personal art of evaluation design rather than selecting and following any particular systematic approach. When the students persisted in pressuring me to give them an evaluation planning protocol, I decided to try to respond.

Thus, I developed my first evaluation checklist. It included all of the questions I thought I had been trying to answer in laying out and negotiating evaluation plans. I had been keeping these questions in mind while I met with evaluation clients and other stakeholders, reviewed pertinent materials, and retreated to write the evaluation plan. I doubt that I ever tried to pose these questions to clients and other stakeholders in any particular sequence. Moreover, I didn't ask clients to answer particular questions if I had gotten the information simply by listening and reading materials. Nevertheless, to serve our students, I listed all the questions in a conceptually ordered set. This became my first evaluation checklist.

As I recall, this first, primitive checklist had six categories of questions: Focusing the Evaluation, Collecting the Needed Information, Organizing the Information, Analyzing the Information, Reporting the Findings, and Administering the Evaluation. Each category included about six particular questions. I emphasized to the students the importance of addressing and answering all questions in the course of planning an evaluation.

However, I also observed that the checklist probably wouldn't work if they treated it as a rigid, linear protocol for interviewing clients. Instead, I suggested that they internalize the checklist's contents and use it as their mental guide. I also advised them to apply the checklist after the fact to assure they had obtained, through various means, the intelligence needed to set up a sound evaluation plan. In the intervening years, I believe that the students for whom I prepared this checklist, and quite a few other evaluators, have found the checklist useful. Its structure has been evident in several evaluation publications, including the Worthen–Sanders textbooks on evaluation. Over the years, I have evolved and improved this first evaluation checklist. The current version is titled the Evaluation Plans and Operations Checklist and is available through our project Web site.

I have also developed or helped to develop a number of other checklists. These include lists of professional standards for personnel and program evaluation, specific checklists keyed to these sets of standards, an evaluation contracting checklist, and, most recently, a checklist for developing checklists. Fortunately, I have found a valuable colleague in the checklists area. Michael Scriven (1991) has argued persuasively about the value of checklists and has developed some important and widely used ones. These include the Key Evaluation Checklist, the Product Evaluation Checklist, and the Duties of the Teacher Checklist.

Recently, when Arlen Gullickson made a funding opportunity available, Dr. Scriven and I agreed to collaborate in conducting the Evaluation Checklist Project. We were fortunate to engage Ms. Lori Wingate as our project manager. The project's objectives, as we have evolved them, are to provide: (1) a paper on the logic and methodology of checklists; (2) updated versions of the various checklists we had developed; (3) new checklists, by Dr.

Scriven and me, for evaluation areas not currently served by checklists; (4) additional checklists by other writers; (5) a methodology for developing checklists; and (6) a WMU web-based repository of evaluation checklists and related information. More or less, we have achieved all six objectives during this past year. You may make your own judgment about our progress by accessing the project's web site.

OUR CURRENT REPOSITORY OF EVALUATION CHECKLISTS

The checklists currently available from our project are identified in Table 1. Each is characterized by its source, intended area of application, intended users and beneficiaries, purpose, criterial emphases, and level of guidance or extent of detail provided.

Source

The first three checklists were authored by Dr. Scriven, the next six by me, and the last one by Ernest House and Kenneth Howe.

Areas of Application

Collectively, these checklists have a wide range of applications. These include guidance for conducting program, personnel, and product evaluations, plus criteria and guidance for assessing evaluations and evaluation systems. They also address particular aspects of evaluation, including evaluation contracts, teacher competency and performance, and steps in the evaluation process. The Key Evaluation Checklist is designed for application to virtually any evaluation, but probably is more useful in summative rather than formative evaluations. The others are focused somewhat more narrowly, with some having special utility in formative evaluations.

Certain evaluation checklists may be used in combination with another checklist. For example, the Evaluation Contracts Checklist could be used along with most of the other checklists. Also, the Duties of the Teacher Checklist could be used in combination with the Personnel Evaluation Systems Metaevaluation Checklist.

Intended Users/Beneficiaries

Most of the checklists are aimed both at evaluators and persons served by evaluators. In my experience, when both groups use an evaluation checklist they can reach clearer understandings of evaluation needs and planned processes, better agreements on evaluation matters, a useful means of recording such agreements, and an agenda for periodically reviewing an evaluation's progress. Only the Checklist Development Checklist has a narrow, particular audience. Mainly, checklist developers would be interested in this narrowly focused checklist. Such checklist developers include mainstream evaluators because they often need to develop new checklists to address the requirements of particular evaluations.

Purposes

The ten checklists vary in their purposes. The Key Evaluation Checklist is aimed at comprehensiveness in pulling together the information needed to reach firm conclusions about the merit and worth of any of a broad range of entities. The Product Evaluation Checklist has a similar orientation, but is focused on educational products, such as curricular materials. Scriven's Duties of the Teacher Checklist provides school administrators a set of philosophically based criteria by which to evaluate teachers' competence and performance.

Three of the checklists I developed are grounded in the national Joint Committee standards for sound evaluations, and are keyed to guiding and judging evaluations and evaluation systems. My Evaluation Plans and Operations Checklist provides guidance and criteria pertaining to the main steps of the evaluation process. This checklist has broad applicability across disciplines. It is the latest generation of the first checklist I developed back in the 1960s. My Evaluation Contracts Checklist actually is a considerably more detailed version of the contracting checkpoints in my other evaluation checklists. Finally, my Checklists Development Checklist was developed at the request of Dr. Gullickson. He judged that evaluation checklists are valuable and that more checklists are needed. Consequently, he asked our project to develop a checklist that could help individuals design good checklists for their own purposes. This new checklist needs testing, and I will welcome feedback from persons who put it to the test.

The final checklist in our set of ten was developed by Ernest House and Kenneth Howe. It reflects their new Deliberative Democratic Evaluation Model. I was especially pleased that Lori Wingate invited Drs. House and Howe to develop this checklist for our project. We now have checklists reflecting three different theoretical perspectives: the formative/summative, consumer-oriented perspective of Michael Scriven; House and Howe's Deliberative Democratic perspective; and the Joint Committee professional standards orientation. I think it will be advantageous to seek checklists that represent additional, creditable theoretical perspectives on evaluation.

Criterial Emphasis

A rich feature of the checklists is their explication of evaluative criteria. Dr. Scriven's checklists spell out in considerable detail the criteria associated with the generic concepts of merit and worth, the generic duties of teachers, and the dimensions of educational product soundness.

My checklists are keyed to professional standards and help to operationalize the Joint Committee's program and personnel evaluation standards. These checklists list the main requirements of each standard that the Committee provided to define the concepts of Utility, Feasibility, Propriety, and Accuracy. Specificity is also provided in my other checklists regarding the dimensions of checklists and the clarity and comprehensiveness of evaluation contracts. I think House and Howe's checklist will be welcomed for its utility in operationalizing and applying their concepts of inclusion, dialogue, and deliberation.

Level of Guidance Offered

In general, most of the checklists provide sufficient operational guidance to qualify as stand-alone evaluation tools. Moreover, they tend to be backed up by relevant detailed

TABLE 1.
Evaluation Checklists

<i>Checklists</i>	<i>Descriptors</i>					
	<i>Source</i>	<i>Area of application</i>	<i>Intended users/ beneficiaries</i>	<i>Purpose</i>	<i>Criterial emphasis</i>	<i>Level of guidance offered</i>
1. Key Evaluation Checklist	Scriven	Generic (e.g., products, personnel, programs, evaluations)	Evaluators and consumers	Foster comprehensiveness in evaluations	Dimensions of merit and worth	Intermediate, backed up by the <i>Evaluation Thesaurus</i>
2. Product Evaluation Checklist	Scriven	Educational products	Evaluators and consumers	Foster comprehensiveness in product evaluations	Dimensions of product soundness	Operational
3. Duties of the Teacher Checklists	Scriven	Teacher competency and performance	Teacher educators and supervisors, teachers, and students and their parents	Foster teacher evaluations that are grounded in teaching responsibilities	Duties of the teacher as defined by Scriven	Intermediate, backed up by detailed materials
4. Checklist Development Checklist	Stufflebeam	Checklist development	Checklist developers, including evaluators	Provide direction for developing sound, useful checklists	Dimensions of sound checklists	Operational
5. Evaluation Plans and Operations Checklist	Stufflebeam	Generic: planning and implementing evaluations	Evaluators and clients	Provide guidance and criteria for planning and assessing evaluations	Components of a sound evaluation process	Operational
6. Program Evaluations Metaevaluation Checklist	Stufflebeam/Joint Committee Standards	Program evaluations: judging evaluation plans and reports	Client, evaluators, consumers, and metaevaluators	Foster professionally defensible program evaluations	A program evaluation's utility, feasibility, propriety, and accuracy	Operational, backed up by detailed professional standards

TABLE 1.
Continued

<i>Checklists</i>	<i>Descriptors</i>					
	<i>Source</i>	<i>Area of application</i>	<i>Intended users/ beneficiaries</i>	<i>Purpose</i>	<i>Criterial emphasis</i>	<i>Level of guidance offered</i>
7. Personnel Evaluations Metaevaluation Checklist	Stufflebeam/Joint Committee Standards	Personnel evaluations: judging evaluation plans and reports	Evaluators, service providers, consumers, and metaevaluators	Foster professionally sound personnel evaluations	A personnel evaluation's propriety, utility, feasibility, and accuracy	Operational, backed up by detailed professional standards
8. Personnel Evaluation Systems Metaevaluation Checklist	Stufflebeam	Personnel evaluation systems: judging the components of evaluation systems	Evaluators, administrators, clients, consumers, and metaevaluators	Provide a tool for assessing and strengthening personnel evaluation systems	An evaluation system's propriety, utility, feasibility, and accuracy	Operational, backed up by detailed professional standards
9. Evaluation Contracts Checklist	Stufflebeam	Negotiating and judging contracts for evaluations	Evaluators, clients, and metaevaluators	Foster evaluations that are grounded in sound contracts	An evaluation contract's clarity and comprehensiveness	Operational
10. Deliberative Democratic Evaluation Checklist	House and Howe	Program evaluations: judging evaluation plans and reports	Evaluators and stakeholders	Foster fairness and validity in evaluations	An evaluation's adherence to principles of inclusion, dialogue, and deliberation	Operational

documents. Dr. Scriven's Key Evaluation Checklist and Product Evaluation Checklist are supported by his well-known and detailed *Evaluation thesaurus* (Scriven, 1991). My checklists are supported by the Joint Committee (1988, 1994) personnel and program evaluation standards, by writings on metaevaluation (Stufflebeam, 2000), and by my recent *AJE* article titled "Lessons in contracting for evaluations" (Stufflebeam, 2001). Finally, Drs. House and Howe (2000a, 2000b) have been issuing important publications to explain their new Deliberative Democratic evaluation model.

CONCLUSION

As seen in Table 1, in less than a year, our project has put together a fairly extensive set of evaluation checklists. As I reflect on the ten checklists included so far, I think there is room for valuable additions. Particularly, I would like to see us add checklists that cover more of the evaluation models. For example, it would be especially useful to include checklists by Bob Stake on case studies and responsive evaluation, Yvonna Lincoln and Egon Guba on constructivist evaluation, Michael Patton on utilization-focused evaluation, and perhaps me on the CIPP Model. I think we also need to work on a detailed checklist concerned with disseminating evaluation findings. Such a checklist should address the difficult matters of prerelease reviews,¹ right-to-know audiences, modes of effective presentation, and dealing with the public media. In addition, Dr. Scriven is developing a checklist for evaluating computer software. With more thought and planning, I think a case can be made for continuing the Evaluation Checklist Development project.

Of course, the proof of the worth of this project must rest in actions and assessments by intended users. We need to assess the extent to which the intended users learn about the checklists, obtain them, apply them, and find them useful in improving their evaluations. Ms. Wingate, Dr. Scriven, and I invite interested evaluators and clients of evaluations to review and apply the checklists on our web site: www.wmich.edu/evalctr/checklists/. We will gratefully receive and thoughtfully consider your critical feedback. We will also be glad to entertain proposals for adding checklists to the site. We hope that this web site and its contents will grow and improve over time, and that it will provide evaluators and their clients with valuable tools for strengthening their evaluations.

NOTES

1. Those persons who have used my Evaluation Contracts Checklist may wish to access a February 2001 updated version on the Evaluation Center's web site that includes a new section on prerelease reviews of reports.

References

- House, E. R., & Howe, K. R. (2000a). Evaluation as a democratic process: Promoting deliberation, dialogue, and inclusion. In K. E. Ryan & L. DeStefano (Eds.), *Deliberative democratic evaluation*. New Directions for Evaluation, 85 (pp. 3–12). San Francisco: Jossey-Bass.

- House, E. R., & Howe, K. R. (2000b). Deliberative democratic evaluation in practice. In D. L. Stufflebeam, G. F. Madaus, & T. Kellaghan (Eds.), *Evaluation models* (pp. 409–421). Boston: Kluwer.
- Joint Committee on Standards for Educational Evaluation. (1988). *The personnel evaluation standards: How to assess systems for evaluating educators*. Newbury Park, CA: Sage.
- Joint Committee on Standards for Educational Evaluation. (1994). *The program evaluation standards: How to assess evaluations of educational programs*. Thousand Oaks, CA: Sage.
- Scriven, M. (1991). *Evaluation thesaurus*. Newbury Park, CA: Sage.
- Stufflebeam, D. L. (2000). The methodology of metaevaluation as reflected in metaevaluations by the Western Michigan University Evaluation Center. *Journal of Personnel Evaluation in Education*, 14(1), 95–125.
- Stufflebeam, D. L. (2001). Lessons in contracting for evaluations. *American Journal of Evaluation*, 21, 293–314.