

Introduction

1

The chapters of many texts can be read out of order, but that is not true with our text. We have taken special pains to put our chapters in the order that makes sense according to a logic of evaluation that we hope you will embrace and use in your own evaluation studies. You can start by shedding any fear of evaluation. One should not fear evaluation either as the evaluator or any other educational partner when going under the magnifying glass. This book provides a professional and comprehensive set of concepts guidelines for conducting evaluation studies that are useful for improving the educational enterprise. Our approach is results centered. We avoid reliance on fringe movements in evaluation in favor of building on the solid and proven core concepts and methods.

This is not your “standard” educational evaluation book. This one focuses on asking and then answering useful educational questions in order to make sensible and justifiable data-driven decisions.

Our approach will introduce you to some powerful yet commonsense evaluation tools for planning and doing evaluation. Some traditional topics like testing, measuring, and statistics are placed in the proper perspective so that they serve the purpose of evaluation, rather than control you. We stress that in any evaluation, you must carefully first decide on what to evaluate and why, and then design and carry out the evaluation to answer a set of critical evaluation questions.

Evaluation, at its core, is simple and useful: Finding what works and what doesn't.

Evaluation compares our results to our intentions. Sensible? We think so. We also know that evaluation is a large field with its share of pitfalls. We will point out some traps and problems associated with the conventional practice and language of evaluation in order to make our approach clear. And practical. We place a primary emphasis on results, as well as on collecting and using relevant and useful data.

We start with some commonsense ideas.

2 • Practical Evaluation for Educators

Whenever we do anything, especially when it involves the future of our learners, it is practical to ask, “How did we do?” We spend the time of our learners, our teachers, and our educators to help them master that which will be useful. Education is about learners and the future: their future as well as ours. We should all care enough to find out whether we helped or hindered. Evaluation, then, helps us to find out what works and what should be modified, changed, or stopped.

Evaluation data, therefore, must be used for improving and never for blaming.

The value of results is related to individual, organizational, and societal goals, objectives, and expectations that must be treated as a whole system (rather than just looking at pieces and parts and hoping the whole thing will go together).

To usefully evaluate practices and programs to improve education, it is vital to collect and use data in a sensitive, practical, and efficient manner. To do this, we put ourselves in the business of asking and then answering the right questions.

With common sense as our premise, we suggest that the best way to convey a useful framework for providing our approach to collecting and using important data is in four distinct phases. We consider these phases a useful framework that you can build on to conduct evaluation. We call this

Five Principles of Evaluation:

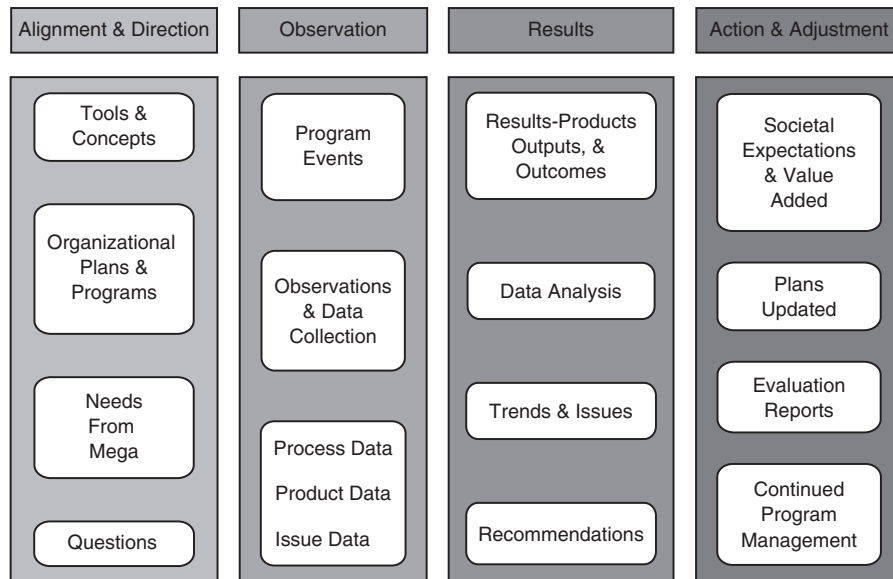
- Systematic Inquiry
- Competence
- Integrity and Honesty
- Respect for People
- Responsibility for General and Public Welfare

framework the Evaluation Action Plan (EAP). We fully expect that the EAP will guide you as you design, develop, and conduct successful evaluations. The plan is reflected in the organization of this book to support planning and doing. The plan simplifies the often overly complex models of evaluation yet emphasizes the major considerations and tasks involved in collecting useful data; based on the right questions, it

will allow you to design, develop, and implement evaluations, as well as act on findings to improve the educational enterprise. We also think it is important to adhere to the ethical and professional guidelines set forth in the five principles published by the American Evaluation Association (AEA).

PHASES OF THE EVALUATION ACTION PLAN (EAP)

Sound evaluation starts by asking the right questions. We build on the proposition that the value of an educational program is based on how well that program meets the needs—gaps in results and consequences—of society. For an evaluation finding to be useful, meaningful, relevant,

Figure 1.1 The Four Phases of the Evaluation Action Plan (EAP)

unambiguous, and acceptable, you must start with some fundamental and essential planning that examines the basis of the educational enterprise. This planning will guide your evaluative inquiry. At each phase of our plan, there are some prime questions to be answered (see Figure 1.1).

Alignment and Direction: Making sure the evaluation questions we ask and answer are useful. To be useful, we ask questions not only about the intended results of educational programs but also about the legitimacy of the results as well. The prime question for *alignment and direction* asks whether the goals and objectives of school programs are based on valid needs—important gaps in results—of society. The consequence of the alignment and direction phase is a set of useful questions to guide the evaluation inquiry. Having that in hand, we can proceed to determine the information required to answer each question. We look for those answers in the observation phase.

Observations: Application of tools and methods to collect valid and useful data. The prime questions for the *observation phase* are: What information will answer the evaluation question set, and what methods will collect the required data?

Results: Comparing what we accomplished with what we wanted to accomplish. The prime questions in the *results phase* are: Did we accomplish what we set out to accomplish, and if not, why not?

Action and Adjustment: Using the data for appropriate decisions on what to keep and what to change. The prime questions for the *action phase* are: What do we report to stakeholders and decision makers, and what should

4 • Practical Evaluation for Educators

we keep or change in programs to bring them into alignment with the needs of society?

Each phase of this plan will equip you to move to the next by providing you with knowledge, tools, and examples, but in this introduction, we would like to set the stage by clearing away some myths and misconceptions about evaluation.

Asking Useful Questions

Educational evaluation must start with educational questions—useful questions. We have provided a typical list below. You will see as we progress in the action plan that data can be collected and used to answer each of these questions. We will stress the importance of the relation between the evaluation questions and the type of data that will support the answer and be both reliable and valid. There are four sets of prime questions that set the stage for each phase of the action plan.

The prime questions for each phase relate to the following areas.

Alignment and Direction: Are the valid needs of society reflected in the results produced by the organizational- and program-level set of solutions? In other words, do educational programs cause students to learn, and are they learning the right things in order to be successful in school and in life? The primary questions yield a host of related supporting questions, each relating to the specific concerns of the evaluation study being planned. This set of questions becomes the focus for inquiry activities, including observant and data collection.

Observation: What information must be gathered to answer the general set of questions? What methods will best gather that information? What support tools are required?

Results: Are the observed results the intended results? What Process factors contributed to the observed results? What issues and trends become apparent over time?

Action and Adjustment: What should be kept, adjusted, or discarded to start new?

You will be pleased when you see how powerful this approach can be. You will also note that building your observation plan will become easier as you look for ways to answer the questions, instead of picking a methodology and then limiting your study to the data that method gathers. A host of study-defining evaluation questions will flow from the primary set of evaluation questions.

Here is an example set:

- What objectives in our educational mission did we achieve?
- Did these objectives add value to our communities and shared society?

- What did learners master?
- Was what the learners mastered worth mastering? In what ways?
- Are methods, tools, materials, and activities effective? Efficient?
- What should we change?
- What should we keep?
- What should we stop?
- What value did we add for all internal and external stakeholders?
- What results justify our continued programs, projects, and activities?
- Do the educational programs have unintended consequences?
- What should we do next?

By asking and answering these vital educational evaluation questions, we can prove our worth to ourselves, our financial backers, and our stakeholders. Is this not all very sensible? We think so, because the questions ultimately define the scope of the evaluation study. The question set must keep us focused on the core purpose at hand. In fact, those fundamental evaluation questions will shape reality, as they will limit our range and scope of the data that help us answer them.

Useful educational questions come from asking, “What it is we want to accomplish in our educational enterprise, and why?” It is also important that this procedure be collaborative, and representative of the full population of stakeholders. Their contributions will provide insights otherwise missed by a narrow evaluation team, and their participation will begin a sense of ownership and buy-in required for using the evaluation results to create the required change.

In this book, we will share with you some frameworks and guidelines for (1) asking the right questions and (2) making sure that your questions are correctly targeted toward really knowing what results were accomplished as well as (3) the usefulness of those results and how to define and get valid data. This entire book is designed so you will be successful at educational evaluation.

Ensuring Commitment

It is vital to establish the “ownership” of education by those who receive and deliver it as well as define its contributions so that we create a lasting partnership for success. Trust, understanding, and agreement on a common destination—results to be achieved—are all key to a successful educational enterprise. If we don’t have the commitment and participation of all stakeholders, our success might be less than it could be.

Evaluation data can sometimes be scary to both educators and stakeholders alike. If there is good news, it should be trustworthy. If there is

6 • Practical Evaluation for Educators

bad news, it is best provided in an environment of trust and the common purpose of continuous improvement and based on real and justifiable data. We should never withhold disappointments, but we best make certain that our successes and shortfalls are based on solid evidence.

Thus this book is based on asking and answering the right questions and collecting the realistic and trustworthy data for comparing our educational objectives with our actual accomplishments. Trust, common purpose, and shared destiny are key to getting and maintaining commitment.

Fear Builders: Actions on the part of the evaluator or evaluation team that cause stakeholders to fear evaluation.

We want you to avoid instilling fear in favor of building trust. From time to time, we will point out evaluation behavior that does both.

Creating the partnership for education also hinges on seriously involving all stakeholders, but also listening to them.

Really listening. While it might be tempting to move ahead with educational plans and evaluations without the stakeholders' involvement and commitment, doing so risks them later seeing your worthwhile efforts as deceptive or worse.

Peter Drucker (1973) had good advice when he suggested we ensure "transfer of ownership" to our stakeholders; people see what is done as

Trust Builders: Behavior on the part of the evaluator or evaluation team that builds trust among stakeholders.

their own rather than belonging to someone else. And the best way to ensure the transfer is to involve the educational partners in setting the objectives and sharing with them the results of our educational successes and shortfalls. Look for ways to

build trust, and your evaluation study will prove to be easier, but your recommendation and findings will have more impact and stand a better chance of leading to meaningful change.

Evaluation provides us the opportunity to have an open and honest relationship with our stakeholders based on performance data, and not just rumors, misunderstandings, and biases. Involving them is the best way to ensure that what we do and deliver is considered in the light of the value we add to learners and our shared communities.

Involving Key Stakeholders in Useful Educational Evaluation

We invite our key stakeholders—our partners in educational accomplishment—to contribute to asking and answering these "right questions."

Key evaluation stakeholders include:

- Learners
- Teachers
- Parents
- Administrators
- Educational staff
- Future employers
- The community (society) in which we live.

The stakeholders we invite to join us should be representative of the communities we serve. They should be seen as a sample of everyone in the educational service area, and they should be seen by their peers as appropriate and representing them. Though as humans, we all have individual interests and concerns, as much as it is possible, do not select stakeholders to help us define useful educational objectives and evaluation criteria—and review our progress—who have axes to grind, who have special “single issues,” or who represent a powerful lobby.

A vital challenge for the stakeholders who will help us all be successful is to keep them focused on results and consequences, not on means, activities, resources, or politics. Single-issue politics from both within and outside of our schools has a chilling effect on defining our objectives, selecting the best curriculum and means, and using our results for improvement only.

Clarifying the Purpose and Nature of Evaluation

Before you start to plan, before you collect data, and before you grapple with our chapters, ask yourself why you are doing evaluation. Is this your initiative, or were you directed to evaluate? What is the motivation for the study? What are you looking to accomplish and contribute as a result of this evaluation? We hope it derives from an outside-in, shared view of societal well-being, which we refer to as “Mega planning.” The aim is to reach an “Ideal Vision,” where the educational programs are an integrated and efficient means to reaching valid societal ends—societal ends that we agree on, not those that are imposed arbitrarily by one group or another.

However, we know that societal/community-referenced evaluation is not always the case. Consider some traditional reasons evaluations are carried out:

- Evaluation to see whether a solution to a problem is working
- Evaluation to discover the cause of a problem
- Evaluation to confirm compliance with a mandate

8 • Practical Evaluation for Educators

- Evaluation to sort out a mess
- Evaluation to bring about a change in social conciseness
- Evaluation to provide feedback to control a program
- Evaluation to satisfy the requirements of law

With careful attention to detailed planning, each of these can be made to be compatible with our data-driven and results-oriented action approach to evaluation. But if taken too narrowly—in isolation and without proper context—each of these has its own narrow set of problems, blind spots, and special data generation and collection troubles. What we are saying is that the way you perceive the evaluation purpose can shape and limit the data that are observed (or not observed), collected or (not collected), and interpreted or ignored. We have taken pains to provide guidance that will help you avoid pitfalls, stick to the right questions, and remain relevant. But the most important understanding that we want to convey in this book is that the value in evaluation stems from a societal view of right and wrong actions as they relate to an Ideal Vision.

We enthusiastically believe that evaluation must directly face the issues that follow from this point: Education is a means to societal ends, and thus adding measurable value to our shared society is critical.

Some evaluations are merely descriptions without judgments of worth, especially value added to our shared society. Such an approach misses the point. We think evaluation practice must confront this issue. This is fundamental to building coherent meaning and integrity in this fractured field.

Our conception of this approach is summed up in the term *Mega*.

WHAT IS MEGA, AND WHY IS IT A PRACTICAL APPROACH?

The concept of Mega is a view of society based on consensus and agreement on how we see and interact with the world. It is a focus on the things we can agree are good and desirable. It avoids the peripheral areas of discordant beliefs that can sidetrack both evaluators and the program actions that lead to societal well-being. It is based on common sense that we all use in our daily lives. In fact, most people around the globe use it for their families and themselves to make daily decisions about their lives and well-being.

For example, here are some guiding principles of Mega thinking, planning, and evaluating:

1. Every time I deal with others, I expect them to treat me the way I treat them. I commit to do no physical, psychological, or financial harm and expect others to do the same.

Introduction • 9

2. When I do business with anyone—fly on an airline, eat in a restaurant, drive in a car, go to the dentist and physician, buy items in a supermarket or pharmacy—I expect them to put my health, safety, and well-being as the first thing on their “to do” list. No exceptions. I do the same in everything that I use, do, and deliver.
3. If you agree with the above, then you are “acting Mega.”

Let’s see what that means as we define Mega and put the idea into action. When you think and act Mega, you commit to add value—real measurable value—to those you deal with as well as to our shared society. This means you don’t cause deadly pollution, you don’t physically hurt yourself or others with dangerous goods and services, and you don’t kill, maim, or poison others. You do what billions of people across the globe do as parents: You ensure the self-sufficiency, self-reliance, and positive quality of life for others.

So what’s so difficult to understand about this? Sociologists call doing so—being Mega—*the social contract*, and it is the basis of much that exists in our shared world:

- The basic and conventional religions around the world
- Rotary and other social service groups
- The Ten Commandments

This all provide guidelines for Mega thinking and doing, and even if you are not a person of formal faith, being Mega oriented is ethical. It is also practical: Why would you want to harm others when such harm might in turn be visited upon you?

In business, the medical Hippocratic oath also is worth adherence: “Do no harm.” Why would you want to harm your clients? Kill them? Injure them? Why make an unsafe or deadly airplane or car or food or medicine? Do so, and you best get out of business and likely also out of the country. Short-term thinking is long-term failure.

Mega thinking, planning, and doing is both practical and ethical: Treat others as you would have them treat you.

Mega is does not mean superimposing your (or anyone else’s) values or procedures on others. It is what is agreed upon by all social partners. And Mega is not code for “really big” or “really important.” Mega means adding measurable value to all, including society.

A few years ago, with some help, Kaufman asked people from around the world and in very diverse cultures, “What kind of world would you want to help create for tomorrow’s child?” They were asked to speak to ends, and not to means, resources, programs, projects, politics, or activities—just

10 • Practical Evaluation for Educators

Figure 1.2 An Ideal Vision

**BASIC IDEAL VISION: The world we
want to help create for tomorrow's child**

There will be no losses of life nor elimination or reduction of levels of well-being, survival, self-sufficiency, quality of life, from any source including (but not limited to):

- war and/or riot
- unintended human-caused changes to the environment including permanent destruction of the environment and/or rendering it non-renewable
- murder, rape, or crimes of violence, robbery, or destruction to property
- substance abuse
- disease
- pollution
- starvation and/or malnutrition
- destructive behavior (including child, partner, spouse, self, elder, others)
- accidents, including transportation, home, and business/workplace
- discrimination based on irrelevant variables including color, race, age, creed, gender, religion, wealth, national origin, or location

Poverty will not exist, and every woman and man will earn as least as much as it costs them to live unless they are progressing toward being self-sufficient and self-reliant. No adult will be under the care, custody or control of another person, agency, or substance: all adult citizens will be self-sufficient and self-reliant as minimally indicated by their consumption being equal to or less than their production.

results and consequences. When these responses were sorted out, an Ideal Vision emerged. This Ideal Vision cut successfully across nations, cultures, religions, and races. Whenever you make a decision, simply review its elements (see it as a fabric and not a bunch of individual strands) and ask, "Will this take us closer or further away from Mega?" The Ideal Vision is described in Figure 1.2.

Did you find anything in the Ideal Vision that you would not want others to apply when dealing with you and yours? If, like us, you focus on the good things society can agree on, why should you not do the same in your life and work?

Mega thinking, planning, and doing are both practical and ethical. We all depend on each other, so why not make everyone Mega thinkers, planners, and doers?

To quote John F. Kennedy, "If not us, who, and if not now, when?"

But desire and motivation to do evaluation are only the first step. You must also have insight and knowledge about what to do and how to do it. Often, you must also know what not to do.

AVOIDING EVALUATION TRAPS AND ERRORS BY FOLLOWING EVALUATION RULES

There are some traps to be avoided. In doing educational evaluation, there are some traps. Common traps. We will show you the major ones in the book as we move along.

Widespread misunderstandings and biases can (but should not be allowed to) creep in from ourselves and from our stakeholders. Collecting useful and complete data should provide answers to the key questions above, but there are some very common and some not-so-common reasons why evaluations can (and often do) get off track. We have identified some of these. We cover these at the end of each section of the book. Evaluation traps are situations that exist in the environment surrounding an evaluation study that involve active forces that cause studies to lose focus, avoid essential information, or gather irrelevant data. Errors are mistakes in planning and executing that trace back to misguided ideas or lack of expertise on the part of the evaluator. For each trap and error, we have provided some rules of the road that will help.

Educational evaluation is far too important to weaken its contributions by defects in our thinking and methods. True, education is “messy” in terms of all of the things that are going on in schools and the system. But still, despite this, our challenge—professionally and personally—is to ask and answer the right questions without mucking up the procedure with errors that can be avoided.

THE GROWTH OF EVALUATION PRACTICE

Building an Approach Based on the Best Features of Accepted Practice

As you plan to do evaluation, you will no doubt encounter some of the vast—the authors would call it “rich”—array of models, methods, and schools of evaluation. We have developed the chapters of this book to help you pick and choose from that assortment when you plan your own evaluation. We have based our own approach to evaluation on practices that have stood the test of time, meet our criteria for common sense, and generally fit within the framework of the AEA principles. The field of evaluation includes much that is useful and much that is distracting, argumentative, and hard to put to practical use. We start with *needs assessment* and *Mega planning*, which provide a framework to set educational goals and to control and analyze solution sets to obtain beneficial educational results. To that core, we have added from the field. Some of the more influential themes include the following.

12 • Practical Evaluation for Educators

Testing and Measurement

Evaluation practice today is a synthesis of several other disciplines that have nourished the growth of the field. The most important of these was the widespread use of testing and measurement to better utilize personnel resource during time of war. This firmly established the discipline as a taproot for conducting evaluation studies to match students to courses of study and help them make career choices. The use of statistical tools soon expanded to include simulation computer models in the discipline of operations research, which applied multidisciplinary approaches to problem solving. This field was also conceived in time of war. But its lasting approach has spread to quality control in organizational/industrial and educational settings.

Evaluators too can now pick from a host of well-developed analysis tools that were developed under the banner of "Operations Research." Parallel to these developments, the study of human behavior blossomed with the work of Thorndike and B. F. Skinner. The advent of general system theory and disciplined study of organizations and group behavior in organizations added more tools and concepts that have contributed to the thinking in the new field of evaluation. Each of the divisions of the social sciences has contributed to evaluation, so that now it is commonplace to see a range in evaluation activity that extends from experimental and measurement-oriented studies to case studies and anthropology. It was only a matter of time, however, before writers influenced by these roots began to identify themselves as evaluators. This began in earnest in the late 1960s and 70s.

The Growth of Models

In the early direction and new language of evaluation, the emergence of a host of evaluation models in the 1970s tended to center around the scholarship of academics, each with some distinctive features but all trying to solve three core problems of evaluation. For every evaluation, the evaluator must solve or answer the following:

1. What should be evaluated? This includes both the target and the limits of the target for evaluation.
2. How do I put myself, the evaluator, in a position to be able to make meaningful value judgments about the merit or worth of the thing being evaluated?
3. Who am I working for? The history of movements and models in evaluation as a field can be traced by the way various evaluators have approached these three questions. In general, the trend has been from narrow focus to broad focus to focus linked to society.

Performance Objectives

Education under the influence of a combination of pragmatism and sensible behaviorism was primed and ready when the use of so-called performance objectives was introduced with the work of Tyler (1966) and Popham (1974) and made rational by the influence of the book *Preparing Instructional Objectives*, by Mager (1997).¹

Evaluation in some early models was the examination of programs to see whether the students met objectives. The evaluators were teachers or material developers intent on showing the utility of educational texts and programs that caused students to reach educational objectives. The model was very school centric. This model was warmly accepted by educators looking for a firm way to anchor the curriculum. However, at its best, the objectives-oriented model of evaluation caused a rather narrow focus, which left some scholars looking for broader conceptions—a flaw in implementation rather than concept. For some, having performance objectives that were in the correct format and measurable did not also guarantee the objectives were addressing the right things. Many also realized that focusing only on objectives could cause the evaluator to miss important data on Process and environment. Educators realized that evaluation could play a role in the development of educational programs to adjust content and procedure along the way to the final Products, Outputs, and Outcomes.² This realization gave rise to a famous distinction when Scriven (1967) introduced the terms “formative” and “summative” evaluation.³ Now, evaluators had a term for the type of evaluation activity used to guide developmental activity in educational programs and another term when evaluation was used to comment on overall final value: What did we accomplish, and was it worth accomplishing?

The Scientific vs. Naturalistic Method

As the conceptual underpinnings of evaluation evolved, other distinctions emerged. Many scholars admiring the progress and precision of the so-called hard sciences conceived of evaluation as an opportunity to conduct research and confirm the utility of programs using experimental and quasi-experimental designs. This was in contrast to a growing number of evaluators who advocated a less disruptive naturalistic form of inquiry that depended heavily on the sensitive skill of the observer, who would through close personal observation grasp the important variables in a program. Still other evaluators saw the organization as the proper focus of evaluation. Plans and programs were consequences of the way personnel conducted the daily business of the organization. The flow of information, resources, and decision making were examined, with the primary task of evaluation to support the decision-making tasks in the organization. The best of these

14 • Practical Evaluation for Educators

models provided coherent linkage of organizational results and purpose with the Process or program (means) used to reach the desired ends. The role of the evaluator was to spot inefficiency or dysfunction and to develop data that would support the decisions that guided the enterprise. Curriculum, planners, and school administrators embraced this model.

The Emergence of Key Evaluation Leaders

Evidence that evaluators were actively searching for answers to the questions posed by the emerging field can be found in the diversity of models described in the works of Worthen and Sanders (1973), Glass and Ellet (1980), and Stufflebeam and Webster (1980), who provided surveys of the field. House (1977) realized making a value claim is a form of argument, and he worked out a system of evaluation logic that took account of the audience for evaluation. Wolf (1979) took this idea a step further in framing evaluation as a form of argument similar to that of the legal system. Guba (1978) contrasted experimental and naturalistic approaches, showing that each had differing advantages.

The experimental approach to evaluation demonstrated a particular relation that was then valued, where the naturalistic approach gathered multiple-relationship data during the course of an evaluation. Some evaluators honed their skills in observation and their ideas about evaluation through years of practice. Stake and his colleagues at the Center for Instructional Research and Curriculum Evaluation at the University of Illinois have refined the case study method into what they refer to as “responsive evaluation.” Here, evaluation teams take an in-depth look over an extended period of time and test their observations by getting multiple frames of reference and the use of “meta-evaluation” discussions. Stake (1974) emphasized the role of audience and the in-depth personal observation available through the case study method.

The Department of Defense

Work for and by the Department of Defense added to the ideas of the early evaluators. Each service developed systematic models for the creation and delivery of instruction. Typical is the Joint Services Model created for the U.S. Army by Florida State, with Branson leading the team (see Branson, Rayner, & Cox, 1975). All of these models stressed the importance of evaluation. These models worked best when coupled with a disciplined investigation of requirements, which the military termed “front-end analysis.” The concept, however, is better illustrated in Kaufman’s work on needs assessment (e.g., Kaufman 1992a, 1995, 1998, 2000; Kaufman & English, 1979).

Evaluation model building gave rise to evaluation theory building. Platt (1982) proposed a general theory that tied key components of the evaluation process to the sources of value in the community, using a framework that outlined 14 key steps. Not all evaluation efforts contain all 14 steps, but Platt argued that many evaluations are incomplete or fail to reach a resolved state when key steps are omitted.

Much later, the issue of theory reemerged in the evaluation literature, not so much as theory about evaluation, but as use of theory related to the thing being evaluated, so as to bring order and criteria to the evaluation—the so-called theory-driven versus black-box debates.

The Proliferation of Movements

As the evaluation field grew larger and as more evaluators filled its ranks, evaluation practice reflected a wide panorama of professional and academic interests. Special interest groups formed so that each special interest in education was also reflected in evaluation. Topical movements with names like Naturalism, Formalism, Individualism, Constructivism, Feminism, and many others became active in writing. A lively array of debates between adherents of differing points of view filled the pages of academic journals. In some circles, evaluation took on the mantle of an art form. Terms like *evaluation portraits* and *word pictures* were used in conjunction with lengthy descriptive narratives. The best of these took on the look and feel of a work of literature in which the insight and merit are as much a function of the writing as the original subject. The work took on value in the same manner that a work of art can show insight into great themes and the human condition. In contrast to this direction in evaluation, many evaluators wanted to address more down-to-earth concerns. Program evaluators (especially where funding was provided by the federal or state government) wanted to answer the mandates of contracts and guidelines. For a brief time, evaluators were common within the ranks of the federal government. That period of promise soon gave way to mounting pressure to reduce the size of government. But a few lasting effects made their way into federal law. The Government Performance and Results Act (GPRA) (1993), followed by the budget law, attempted to shift the focus of evaluation from Inputs and Processes to results and consequences.

Classification Systems

The range of issues and approaches to evaluation was now broad enough to stimulate attempts to classify evaluation and to provide some means of talking about studies to point out what they did and did not accomplish. A very popular schema was devised by Kirkpatrick (1994).

16 • Practical Evaluation for Educators

Kaufman and others have expanded on Kirkpatrick's attractive model and extended it to have a Level V: *Societal impact and consequences*, which is a fundamental dimension of this book (for example, Watkins, Leigh, Foshay, & Kaufman, 1998).

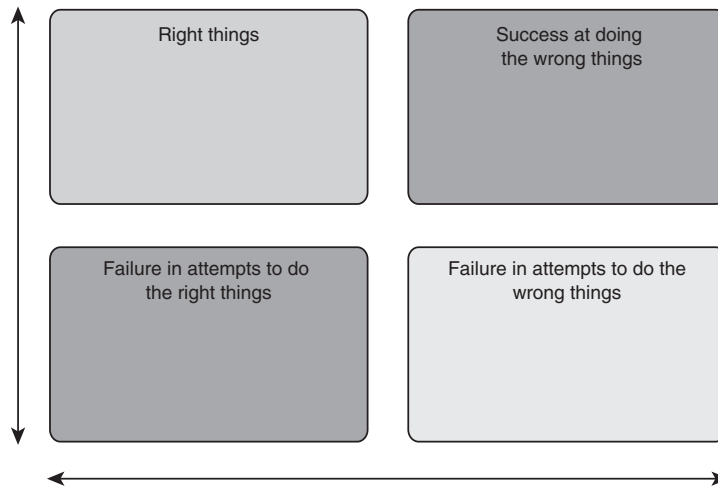
THE PREMISE OF OUR APPROACH

Our premise is that we believe it is possible to meld the best features from the field that have stood the test of time with a societal value-added (or outside-in) derivation of needs (gaps in results) as drivers based on an Ideal Vision for society. We start with societal needs and requirements and roll down inside the educational organization and operation to define what should be done, accomplished, and delivered. (Earlier, we noted that that *Mega* was the term for societal value added and provided a definition, which is the Ideal Vision.) We believe that it is the core job of evaluation to provide results-based data for decision making based on the value or merit of the educational enterprise. Our first focus is on *results*.

Not just any results will do. We believe that the educational enterprise must have legitimacy in a wide societal context. We must not only report on how well our goals are accomplished within the educational agency but also determine that we are pursuing the right goals. The topics and tools presented in the alignment and direction part of the action plan equip the evaluator with the means and perspective to determine that the school system is meeting legitimate needs of society. We stress the importance of asking the right questions, and those questions include, "Do we add value to our communities and shared society?"

From the questions, we derive criteria that tell us where to look for answers. Our second phase is observation. This is where we gather data that will help use answer evaluation questions. The question set determines the methods (and not the other way around). We allow for a wide range of methodology models (models that rely on the central figure of the evaluator as observer, on a system view of information flow, on scientific method on case studies, and on attainment of prespecified ends). The results phase of our action plan has a primary emphasis on educational Outputs and Outcomes and a secondary focus on the Process and environmental factors that produce them. Our action phase connects evaluation results data with the administration of programs and decisions. We believe that continuous improvement is both possible and mandatory. This integrated approach to evaluation is a step in the right direction to solving some of the pressing problems of education. Figure 1.3 illustrates the way we see things.

Figure 1.3 A Basic Job Aid for Assuring Objectives Focus on Ends and Not Means



You can imagine four possible value and worth propositions for any educational program. You must ask yourself: *Will my evaluation efforts be able to detect each of these possibilities?* We suggest that any evaluation that misses one or more of these is incomplete.

As you consider the state of education today, in which many schools fail to prepare their students to enter society and in which there is a growing threat of physical violence, we hope that you agree that seeking an Ideal Vision is not only appropriate but also our moral and ethical duty. How we carry out this responsibility is the subject of this book. In dealing with fear of evaluation, we confront a major obstacle to improved educational evaluation.

If you agree, let's move on to the "Whats" and "Hows" of evaluating educational programs to improve our success with learners and in our shared world.

We intend that this book will help you in this adventure to design and carry out evaluations that can guide the educational enterprise to find what works and what doesn't.

EVALUATION LOGIC

When evaluators set out to evaluate educational programs, the usual focus of the evaluation study is the nature of the program and the results of the program in terms of student learning. Does a reading program

18 • Practical Evaluation for Educators

teach students to read? The desirability of learning to read is often taken for granted.

When a program seems to work, evaluators report the program as being effective, and perhaps even efficient, depending on resources consumed. This common approach to educational evaluation evaluated the means to an end and reported on the effectiveness in reaching the desired result. That is compatible with our approach, but we intend to take it one step further. We also caution that some evaluation studies focus on means apart from ends, and that is not what we recommend. We want to align everything we use, do, produce, and deliver in our schools and educational systems with value added to our shared society. To do less is questionable in terms of practical thinking and ethics.

Our approach has a wider scope because we also want the focus of evaluation to check to see whether the ends satisfy, or meet, valid needs. In the context of education, instructional programs are solutions that are carried out with a mix of stakeholders, resource providers, institutions, and individual teachers. That is why we include a discussion of needs assessment and Organizational Elements Model in a book on evaluation. We have integrated the evaluation of programs and societal needs in our four-phase approach. Evaluation of a program in that setting should flow from four primary sets of questions (a set for each phase of our action plan).

The next step is to determine what information is required to answer each evaluation question and how the information can be obtained, which leads to what to observe.

Observation requirements determine method, and not the other way around. Observations are recorded as data. Some observations can be made more precise by measurement. Some measurements can be further refined using statistics. Descriptive statistics reduces data to a clear picture. Inferential statistics reveals relationships that might not be apparent on the surface of the measurement data. We emphasize that this migration from question to data to statistic is driven from the top down. We caution that picking a statistic and limiting your evaluation inquiry to a single data set that fits the assumptions of the statistical routine can lead to some very striking errors and omissions in your study.

So, when doing educational evaluation, let us both *do it right* as well as *do what is right*. Our learners and stakeholders depend on it.

Let's start *Practical Evaluation for Educators*.

NOTES

1. These were earlier called "behavioral objectives," but practice refocused these on performance (Mager, 1997).

Introduction • 19

2. We will note later that there are three aligned levels of results, something that is almost always missed in conventional planning and evaluation. These three levels are societal, organizational, and individual contributions and accomplishments.

3. Scriven also provided evaluators with the term *goal-free evaluation*, where people ask—without reference to existing goals and objectives—“What really happened here?” in order to reveal unexpected results and consequences. This is a very powerful approach, since it frees the evaluators from being limited to preexisting purposes and intents.