When I first came to UCLA in 1964 as a new faculty member, I spent many hours in the office of a senior colleague, a man of great practical wisdom. I was fascinated by a giant clothespin that sat on his desk, crammed with papers and notes, which had a brass plate on top inscribed “New Ideas.” I wondered about it and after many months had the courage to comment on it. Erick Lindman gently smiled and turned the clothespin over; on the other side was the inscription “Old Ideas Still Good.”

It took several years for me to fully appreciate that sentiment. As I transitioned professionally into what many of us felt to be the “newly emerging” field of evaluation, I came to read the works of Ralph Tyler in connection with the evaluation conducted on the famous “Eight-Year Study” of progressive education. In Tyler (1942), I found many concepts that I recognized to be the basis of contemporary approaches to evaluation. Below is a sampling of statements from Tyler, followed by references to subsequent theoreticians in the field whose ideas show a striking parallel:

Another important purpose of evaluation that is frequently not recognized is to validate the hypothesis upon which the education institution operates. (p. 492) (Theory-Based Evaluation? See Fitz-Gibbon & Morris, 1975; Chen, 1990)

One purpose of evaluation is to make a periodic check on the effectiveness of the educational institution, and thus to indicate the points at which improvements in the program are necessary. (p. 492) (Formative Evaluation? See Scriven, 1967)
The participation of teachers, pupils, and parents in the processes of evaluation is essential to derive the maximum values from a program evaluation. They all have a stake in the educational program of the school or college. They can all contribute to the formation and classification of objectives, they are all in a position to obtain evidence about the progress pupils are making, they can all benefit from efforts to interpret the results of appraisal. (p. 497) (Stakeholders? See Stake, 1975. Participatory Evaluation? See Cousins & Earl, 1992)

A second basic assumption involved in evaluation is that the kinds of changes in behavior patterns in human beings which the school or college seeks to bring about are its educational objectives. . . . An educational program is appraised by finding out how far the objectives of the program are actually being realized. (pp. 495–496) (Goal Attainment Models? See Popham, 1975)

Clearly, many current ideas in evaluation had their foundation in this early writing on the subject. Likewise, many subsequent words of wisdom have had an enduring impact on the writings of other theorists. In reflecting on the insightful words on the other side of the clothespin, I felt perhaps that the inscription should have read “Old Ideas Revisited and Enhanced—But Still Good.” To understand where we are in the field of evaluation today, it is important to revisit the old ideas—the roots—of evaluation and to see the way they were enhanced over time: the new growth, if you will.

THEORIES

In the following chapter, Christina Christie and I will talk about evaluation “theories.” The reason why the word theories is in quotes is that while it is conventionally used in evaluation literature, in some ways, it would be more appropriate to use the term approaches or models. In this context, there are two general types of models: (a) a prescriptive model, the most common type, is a set of rules, prescriptions, prohibitions, and guiding frameworks that specify what a good or proper evaluation is and how evaluation should be done—such models serve as exemplars—and (b) a descriptive model is a set of statements and generalizations that describes, predicts, or explains evaluation activities—such a model is designed to offer an empirical theory. Regarding descriptive models, Henry and Mark (2003) and others have advocated for conducting research on evaluation. I applaud such efforts. But we are a long ways away from achieving such status. When (and if) we do, then the descriptive models would define what is to be appropriately “prescribed.” Until then, however, we must rely on the prescriptive models generated by knowledgeable members of the evaluation community to guide practice.

Almost everyone who has written about educational evaluation has, in one way or another, made prescriptions. Some have spent more time systemizing their standards, criteria, and principles. A few have tried to defend or justify their prescriptions. None of the approaches is predictive or offers an empirical theory. That is, these “theories” have not been validated by empirical research. Thus, in the strictest sense, what we will refer to as “evaluation theories” do not fully qualify for that status. Nonetheless, we intentionally refer to them in that way to reflect their
most common current usage, and similarly, we refer to those who have developed evaluation approaches and models as “theorists.” In this volume, we identify theories by the name of the theorist prominently associated with it.

THEORISTS

Before we begin the process of discussing evaluation theorists, it is important to indicate the way in which we have distinguished various individuals prominent in evaluation. Obviously, all of those who have written about evaluation cannot be considered to have developed a unique evaluation theory (as we use the term here). While their contributions may nonetheless be quite significant, they are not theorists or model builders. We have chosen to categorize the primary writers in evaluation in four ways.

Some individuals have contributed very substantially to the basic research methodology that forms the essential foundation for much of the work in evaluation. These include people such as Donald Campbell, Julian Stanley, Thomas Cook, Matthew Miles and Michael Huberman, Robert Yin, and Anthony Bryk. Of these individuals, we have written only about Donald Campbell in the discussion of theorists because of the unique impact of his methodological contributions. Let us refer to these individuals as methodologists who have influenced evaluation.

Another category of individuals we identify as evaluation issue analysts. Their work, while not necessarily associated with an evaluation model or theory, has very substantially assisted in the understanding of various aspects of evaluation. Individuals in this category would include Lois-Ellin Datta, Stewart Donaldson, Karen Kirkhart, Jonathan Morell, Michael Morris, Thomas Schwandt, William Shadish, Nick Smith, and many others.

An additional category of individuals we identify as evaluation interpreters and teachers. Individuals in this category might also be prominent analysts of evaluation (and many are), but they are distinguished by having written evaluation textbooks that, while not necessarily expounding a new theory, provide an important resource for teaching about evaluation and helping to interpret its nuances. Individuals in this category include, among many others, Michael Bamberger, Linda Mabry, and Jim Rugh (2011), Jody Fitzpatrick, James Sanders, and Blaine Worthen (2011), Rita O’Sullivan (2004), Emil Posevac and Raymond Carey (2007), and Liliana Rodriguez-Campos (2005). Some of these, arguably, might be considered theorists.

Finally, we have attempted to identify individuals who are definitively associated with a particular theoretical position on evaluation as evaluation theorists. These distinctions are difficult and perhaps subjective, since many in this category may not have presented a full theoretical exposition, but nevertheless, they appear to us to have proposed a particular evaluation orientation. Furthermore, some of those in the category denoted as “evaluation interpreters and teachers” might well be considered as theorists. Space limitations may be the only reason for their noninclusion in the “theorist” category.

One further note about the selection and inclusion of evaluation theorists. I have restricted my consideration of evaluators to those who speak about the field generically and whose writings
are not restricted or focused on a specific field. For example, those who write only about evaluating health programs or education or social welfare have not been included. Furthermore, those who do not consider themselves as evaluators but exclusively assign another disciplinary designation to their name are also not generally included. (Some deviations from this rule are noteworthy—e.g., Donald Campbell.)

CATEGORY SYSTEMS

There have been many prior attempts to look at the ways in which various theoretic perspectives relate to each other. Earlier efforts have taken the form of category (or classification) systems. These simplified structures provided a way to identify a limited set of characteristics for grouping theories. Entries within a category are deemed to belong together in the sense that they can be judged to be similar with respect to the characteristic or configuration of characteristics that define that category. However, in making this judgment, the categorizer is selecting from the many aspects of the approach only those that are considered most essential. In many ways, this is similar to an artist’s creation of a caricature, portraying someone or something by focusing on (even overemphasizing) its most prominent features. Among the earliest evaluation category systems were those provided by Worthen and Sanders (1973) and Popham (1975). Subsequently, category systems have been developed by House (1978), Glass and Ellett (1980), Alkin and Ellett (1985), Williams (1988), Shadish, Cook, and Leviton (1991), Alkin and House (1992), and others.

Category systems are of great value. By grouping theories within a category, frequently a new concept is introduced or an existing concept is reinforced. Thus, for example, Stufflebeam’s context–input–process–product (CIPP) model and my Center for the Study of Evaluation (CSE) model of the early 1970s were defined as decision management (Worthen & Sanders, 1973) and decision facilitation (Popham, 1975), respectively. This provided an easily understood way for practitioners to think about and define their preferred approach to conducting evaluation.

Category systems also were an aid to theorists in understanding perceived relationships with other theorists. Placement within categories led theorists to question whether their views were being portrayed adequately (see Alkin, 1991). Since category systems dramatize those features presumed to be most prominent, illuminating those features and not others might provide discomfort to theorists. Theorists might not have seen certain emphases in their own work. The writing might have led to what they considered to be “misinterpretation.”

Moreover, theorists’ views are not fixed in time, as would seem to be implied by their published works. Theorists typically change their views over time, and their published work often lags behind these changes. Nonetheless, one’s views as perceived by others (whether or not they are still held) have influenced theorists. Whatever the explanations for a perceived portrayal of a theorist’s views, the perceptions provided by category systems may force theorists to reconsider their views and perhaps modify them.
While earlier category systems prior to the first edition of *Evaluation Roots* served evaluation well, they suffered from several deficiencies. These category systems failed to portray the historically derived relationships between theories. That is, they failed to show which theoretical formulations provided the intellectual stimulation for new theories (the “Old Ideas Revisited and Enhanced—But Still Good” notion). How might such relationships be depicted?

**EVALUATION THEORY TREE**

The guiding framework for the analyses in the *Roots* evaluation theory tree is provided by the comparative theory work of Alkin and colleagues (Alkin & Ellett, 1985; Alkin & House, 1992). As early as 1985, Alkin and Ellett maintained that all prescriptive theories must consider (a) the issues related to the methodology being used, (b) the manner in which data are to be judged or valued, and (c) the user focus of the evaluation effort. We represent relationships between theories in the form of an evaluation theory “tree,” with each of the main branches representing one of these main dimensions: use, methods, judgment/valuing. Thus, each of the theorists is presented on the tree on a branch that we believe represents his or her main emphasis among these three. The distinction between evaluation models based on these three dimensions is not one based on exclusivity, or that only one model believes in the use of methodology and others do not. Rather, the category system is based on the relative emphasis within the various models. It might then be possible to ask this question: When evaluators must make concessions, what do they most easily give up and what do they most tenaciously defend (Alkin & Ellet, 1985)?

In the theory description in the first edition of this book (Alkin & Christie, 2004), we have attempted to place theorists in a position on the tree that best reflects the way in which they relate to other theorists of the same orientation. This caused some discomfort to various theorists. Carol Weiss, for example, indicated that she was satisfied with her placement but felt that, to some extent, she belonged on the “use” branch as well. David Fetterman, likewise, agreed with his placement but felt that it did not adequately represent his interest in “social justice.” Jennifer Greene commented that Lee Cronbach is not fundamentally concerned about methods but that his placement on the “methods” branch was probably as good a representation as possible. Nonetheless, the process of categorizing in the form of an evaluation theory tree was helpful.

Furthermore, we have attempted to place theorists in a position that best reflects the way in which they relate to other theorists on the same branch. We considered the extent to which various theoretical ideas influenced the theory. This is sometimes historical, but not always so.

Indeed, we recognize this tree to be a drastic oversimplification of very complex relationships. “New” theoretical formulations are usually derived from a number of influences, both within the evaluation field (as the example of Tyler at the beginning of the chapter shows) and from external disciplines. By depicting the relationship between theorists and other evaluation writers with whom they share a common perspective, we capture only one dimension of the influences on their theoretical perspectives. Typically, this single characteristic is the main
influence on each writer’s work, but theorists differ in other respects. Clearly, theorists also are
influenced by colleagues who have been identified on the basis of another main characteristic—
that is, those who are on another branch of the theory tree. In *Evaluation Roots* (2004), I ana-
lyzed the evaluator influences from other branches of the tree based on the names identified by
the theorists in their respective chapters. Carol Weiss, for example, identified Alkin, Guba, 
House, and Patton as theorists from other branches of the tree. Moreover, she indicated having
been influenced by other evaluators not on the evaluation theory tree: Howard Freeman, Gene
Glass, William Shadish, and others. Robert Stake identified Cronbach, Stufflebeam, and Tyler
with other evaluators including Mike Atkin, David Hamilton, and Tom Hastings. Bradley
Cousins identified Carol Weiss from another branch of the tree and other evaluators including
Jennifer Greene and Michael Huberman.

Also, evaluators are stimulated in their thinking not only by the writing of other theorists in
evaluation but also by a variety of other intellectual stimuli. Theorists from other disciplines,
past and present, often provide insights and inspiration to those thinking about evaluation. An
excellent case in point is provided by Ernest House’s *The Logic of Evaluative Argument* (1977),
which clearly shows the influence of early evaluation theorists such as Michael Scriven and
Robert Stake but which also relies heavily on the work of two Belgian philosophers, Perelman
and Olbrechts-Tyteca (1969) and on the work of Weizenbaum (1976). Numerous other examples
will be provided in the chapters by the various evaluation theorists.

A WIDER PERSPECTIVE

This volume differs from the first edition in many ways. Most notable is the inclusion of more
authors—a widening of the tree if you will. The field has changed, and Tina Christie and I felt
the need to reflect this in the selection of other authors. New invitations to participate have been
extended to the North American authors Eleanor Chelimsky, Jennifer Greene, Henry Levin,
Melvin Mark and Gary Henry, and Donna Mertens. Greene’s work and that of Donna Mertens
have become substantially more prominent and merited inclusion. Mark and Henry were not
included in the first edition because we did not pay sufficient heed to their book with George
Julnes. Eleanor Chelimsky had created an impressive model for evaluating national programs in
her work at the GAO. Finally, it struck us that the inclusion of Henry Levin provided an evalu-
ative dimension not otherwise represented—cost-effectiveness evaluation.

One American author, Elliot Eisner, previously represented by a chapter, will be discussed
and placed on the tree but will not have a separate chapter. His chapter from the previous volume
is available at the SAGE website for this book, www.sagepub.com/alkin2e.

It was my intention that this volume would be titled *Evaluation Roots: An International
Perspective*. To attain this goal, a chapter summarizing evaluation theory in Europe was solici-
ted from Nicoletta Stame, and this chapter was included. Another chapter summarizing work
in Australia/New Zealand was prepared by Patricia Rogers with the assistance of Jane
Davidson. These are the two parts of the world where there has been the most evaluation work
promulgated. To complete the international perspective, a third chapter was commissioned to briefly describe evaluation activity elsewhere in the world. Unfortunately, that author failed to produce a manuscript. Thus, instead of a fully international perspective, I have produced *Evaluation Roots: A Wider Perspective of Theorists’ Views and Influences*.

NOTES

1. Bryk might also be considered to belong in the theorist category for his contribution to the stakeholder evaluation model. We have not included him because his primary contributions have more significantly been to statistical methods than to evaluation theory.

2. A chapter by the Australian theorist John Owen, which appeared in the first volume, is likewise available at the SAGE website for this book, www.sagepub.com/alkin2e.

3. To partially address this issue, Fred Carden and I are developing an article focusing on evaluation in low- and middle-income countries, to appear in an issue of the *Journal of Multidisciplinary Evaluation*.

REFERENCES


