

## ADVANCING EMPIRICAL SCHOLARSHIP TO FURTHER DEVELOP EVALUATION THEORY AND PRACTICE

Christina A. Christie  
Graduate School of Education & Information Studies  
University of California, Los Angeles

**Abstract:** Good theory development is grounded in empirical inquiry. In the context of educational evaluation, the development of empirically grounded theory has important benefits for the field and the practitioner. In particular, a shift to empirically derived theory will assist in advancing more systematic and contextually relevant evaluation practice, as well as lead to the development of contingency theories that specify the conditions under which particular evaluation practices are optimal. In contrast to the historical *outside-in* stance, empirical research on the field of evaluation must acknowledge practitioners as “knowers,” allowing for unique insights into the intersection of theory and practice.

**Résumé :** L'élaboration adéquate d'une théorie se fonde sur la recherche empirique. Dans le contexte de l'évaluation éducationnelle, l'élaboration d'une théorie empiriste comporte des avantages importants pour le domaine et le praticien. En particulier, l'adoption d'une théorie reposant sur l'empirisme favorise la promotion d'une pratique d'évaluation plus systématique et pertinente au contexte. En outre, elle donne lieu à l'élaboration de théories de la contingence qui spécifient les conditions optimales de pratiques d'évaluation particulières. Contrairement à la position historique prônant l'influence extérieure [*outside-in*], la recherche empirique sur le domaine de l'évaluation doit reconnaître le statut de « savants » des praticiens, qui permet de jeter un éclairage unique sur le recoupement de la théorie et de la pratique.

One widely accepted aim of evaluation is to determine the value of social, behavioural, and educational programs that are designed to improve society and human interactions (Shadish, Cook, & Leviton, 1991). Evaluation, then, is perhaps best thought of as a pragmatic or professional field rather than as an academic one. This presents certain complications, given that professionals and

academics differ in their principal purposes. Academic fields (e.g., mathematics) strive to advance knowledge, and therefore comprise theories that are descriptive in nature and empirically derived. Professional fields (e.g., law, medicine) have more practical objectives and therefore offer theories that are prescriptive in nature—that is, that describe the application of knowledge. Without more in-depth, empirically grounded, descriptive study, however, the development of prescriptive theory is limited at best.

With respect to evaluation, Alkin and Ellett (1985) make a helpful distinction between the uses of prescriptive and descriptive theories. Prescriptive theories are intended to offer principles, rationales, and organization for the procedural choices made by evaluators (Chelimsky, 1998) and to orient practitioners to the issues and problems with which they must deal. They are mostly formulated through the accumulated knowledge and expertise of theorists who engage in the conduct of evaluation, and are intended to guide practice rather than explain phenomena. Descriptive evaluation theories, which are typically empirically grounded, are akin to a more traditional social science notion of theory. They are concerned with evaluation as an intellectual discipline.

Empirical research about the evaluation process can provide powerful insight and offer evidence from which prescriptive theoretical ideas can be developed and enhanced. Empirical studies in evaluation vary in topic, but in general they describe the characteristics and practice of evaluators within a discipline, explore potentially important relationships between related variables, and describe the frequency of occurrence of factors that may be important to evaluation within a program or discipline. Empirical work in evaluation also includes case examples where the history and lessons learned from evaluation experiences are discussed without results from systematic data collection associated with the experience. These reflective narratives offer valuable insight but are subject to recall and reporter bias. More formal prospective empirical analysis is critical to the further advancement of prescriptive theories of evaluation, as demonstrated by studies that have identified factors associated with increased utilization of evaluation results, such as the importance of addressing questions that are relevant to stakeholders (see Braskamp, Brown, & Newman, 1978; Brown, Newman, & Rivers, 1980; King, 1982; Patton et al., 1977; Smith, 1983). As a result, prescriptive theories about utilization were further developed and evaluation practice changed to better facilitate use (Patton, 2008).

Many evaluation scholars, including Shadish et al. (1991) and more recently Mark (2007), have called for an increase in systematic research on evaluation. Several reasons have been offered, arguably the most important of which is that research on evaluation can serve as a guide to subsequent evaluation practice. Looking to other practical fields can help us to better understand how this can take place. Clinical medical practice, for example, requires the conscientious and judicious use of current best evidence in conjunction with clinical expertise and patient values to guide health care decisions (Adams & Titler, 2009). Likewise, I would argue that evaluators should include the current best evidence on evaluation theory and practice in combination with methodological expertise (and, some would argue, substantive field expertise) and stakeholder values to guide practice decisions. Here, an evaluation begins with numerous systematic studies of the new practice, which, when meta-analyzed, offer evidence that it is better at achieving the desired outcomes than some other practice. In addition, practitioners should provide feedback and monitor the conditions under which the practice was implemented as well as the outcomes of the implementation. This can then inform implementation by providing an understanding of the conditions under which the practice works best and the conditions under which it does not work.

This approach points to a need not only to focus on studying our theories and practice, but also to focus our efforts to what medicine calls “translation science.” In other words, a shift in practice that is grounded in empirically derived theory can assist in advancing more systematic and contextually relevant evaluation practices, as well as lead to the development of more contingency theories—that is, theories that specify the conditions under which particular evaluation practices are optimal (Shadish et al., 1991). It will also help to advance evaluation scholarship by identifying the issues of concern to the field, as well as develop a better understanding of the value of evaluation, above and beyond what is offered by other disciplines.

The argument for increased empiricism presented in this article describes a move toward theories and practices derived from systematic observation rather than experience. These approaches use methodologies that position practitioners as knowers, which can offer a new understanding of the intersection of theory and practice. This argument is qualified by the admission that empiricism alone will not offer all of the knowledge necessary to fully establish evaluation as a profession or discipline, nor will our empirical work be without limits. Rather, I much more modestly argue for shifting the discourse

from a place where empiricism is only intermittently addressed by a relatively small group of people, to one where it enjoys more careful consideration, prominence, and significance by the broader evaluation community. Evaluation is called upon to offer other practice-based fields a stronger evidence base for those practices, and we should therefore move toward establishing a firmer evidence base for our own.

## WHAT WE KNOW

Over a decade ago, Will Shadish gave his 1997 American Evaluation Association Presidential address titled “Evaluation Theory Is Who We Are.” In this paper (Shadish, 1998), he argues why evaluation theory is central to our profession, and identifies the contributions of evaluation theory to evaluation as a profession and, I would add, as a discipline. These contributions include a common language, the identification of important issues (e.g., evaluation use), clearly defined themes, an identity defining evaluation from other professions, a face for presentation to the outside world, and, arguably most important, a unique knowledge base.

While evaluation theory is seen as being central to our field, theory development in evaluation has been, by and large, grounded in the experiences of those conducting the evaluation studies, rather than through empirical investigation. In 1991, Alkin described six ways in which theorists come to modify their theoretical thinking: (a) theorists’ confrontation between one’s own theories and others’ interpretations of them, (b) theorists’ confrontation with other theories, (c) theorists’ confrontation with consolidation and categorization schemes, (d) one’s own research, (e) field-based experiment testing, and (f) influences from personal interactions. Notably, only two of these involve empirical study. Historically, evaluation theorists have relied much more heavily on the four non-empirical mechanisms (points a, b, c, and f) for theory development than the two that are empirically based. This has resulted in a knowledge base about evaluation that can best be described as an accumulation of presumptions about the relationships between evaluation processes and outcomes.

Alkin’s discussion also focuses on how theorists’ views develop, rather than on how actual theories are developed, and although these processes are of course interrelated, this is a distinction worthy of note. For evaluation theory to be practical—one of its primary purposes—it should be systematically examined. Doing so will help evaluators

better delimit a theory's operational specificity, range of application, feasibility in practice, discernable impact, and reproducibility (Miller, 2010). For this to occur, the evaluation community must make a decided and sincere commitment to developing and enhancing the empirical scholarship of evaluation.

## WHY WE SHOULD CARE

Because evaluation is a practical field, legitimate questions emerge about what empirically grounded theory might offer it. For example, how might empirical inquiry of theory improve how decisions are made? Are we better off having theories specific to evaluation rather than relying on theories from other disciplines? While Scriven (e.g., 1995, 2003) and others (e.g., Henry & Mark, 2003; Mark, 2007; Shadish, 1998; Shadish et al., 1991) have spent time addressing these “value added” questions, it is important to flesh out some of them here.

Research on evaluation can contribute to at least four areas of relevant knowledge. First, it can help to further define and demarcate the field. There have been some non-empirical attempts at this. For example, Scriven's Logic of Evaluation (1995), Shadish et al.'s (1991) five fundamental issues (social programming, knowledge construction, valuing, knowledge use, evaluation practice), and Alkin and Christie's (2004; Christie & Alkin, 2008; Christie & Alkin, in press) “evaluation theory tree” are all attempts at theorizing about evaluation at a general level. Only recently have there been attempts to empirically describe evaluation in this way—for example, the bibliometric analysis offered by Heberger, Christie, and Alkin (2010). Second, research on evaluation can support the development of knowledge on evaluation issues. For example, empirical work on evaluation use, which began in the 1970s, has influenced our beliefs about and practices for promoting use. At various points, summary analyses of the empirical body of literature on use (e.g., Cousins & Leithwood, 1986; Shulha & Cousins, 1997) and related issues like evaluation capacity building (e.g., Cousins, Goh, Clark, & Lee, 2004) have helped us to better understand these issues.

Research on evaluation can also lead to the revision of current prescriptive theories by clarifying what transpires when evaluation is practiced. Example of this third area of contribution include Christie's (2003) study of Healthy Start evaluators, which describes differences in practice between internal and external evaluators and links this

description to evaluation theory. And fourth, empirical inquiry should lead to the development of comprehensive contingency theories of evaluation. In contrast to general theories about a given phenomenon, contingency theories provide discriminating and contingent explanations and recommendations for practice, thereby offering hypotheses about how variables operate individually and related generalizations on how and under what conditions they behave in specified conjunctive combinations (George & Bennett, 2004). Utilization-focused evaluation has been described as an empirically derived contingency theory, while theory-driven evaluation (Chen, 1990) and emergent realist evaluation (Mark, Henry, & Julnes, 2000) have also been identified as other types of contingency theories (Mark, Greene, & Shaw, 2006). Increased empirical knowledge in these four areas, separately and together, should help to clarify the consequences of following a particular theoretical prescription and thereby support arguments for the use of certain approaches over others. Ideally, this will shift discussions from ideological and political arguments for conducting evaluations in a particular way to conversations grounded in evidence of which approaches are most appropriate for addressing the main evaluation purpose and questions of interest.

## RANDOMIZED CONTROLLED TRIALS: A CASE STUDY

To illustrate the need for and potential benefits of having empirically examined theories and practices that are relevant to evaluation, it is useful to point to a current issue, namely existing U.S. federal policies that identify the randomized controlled trial (RCT)<sup>1</sup> as the only design for generating credible “scientific” evidence of program effectiveness. The identification of a particular evaluation approach as *the* approach from which decisions about what should and should not be considered an “evidenced-based” practice or program has frustrated many in the field of evaluation and has sparked a heated debate among respected colleagues and friends, some of whom have been driven to agree to disagree over this issue (see Donaldson, Christie, & Mark, 2009).

This question of credible evidence has ignited passion and debate among scholars because it implicitly identifies one type of evaluation approach as “strong” and “scientific” while suggesting that all others are “weaker” (if not weak) and “unscientific.” Many argue that the distinction between scientific and unscientific evidence when based solely on the approach used to gather information is false and ill-conceived. Others argue that evaluators must offer the most accurate evidence possible and that RCT is the best approach available to study

questions of effectiveness; for all other questions, alternate evaluation approaches may be legitimate choices.

In an effort to better understand how well nonrandom experiments can approximate results, Shadish, Clark, and Steiner (2008) compared the relative effect sizes of random and nonrandom experiments. The authors found that when pretest (covariate) variables that might predict the condition were chosen by participants in the nonrandomized condition, and when all participants were measured on the outcomes of interest, ordinary linear regression reduced bias in the nonrandomized experiment by 84–94%, using covariate-adjusted randomized results as the benchmark. This study suggests that adjusted results from nonrandomized experiments can approximate results from randomized experiments. The authors urge others to attempt to replicate the findings of this arguably groundbreaking study. It is this type of research on evaluation that lends empirical substance to otherwise theoretical arguments.

Debates focused on particular evaluation approaches are indeed complex and, as already noted, are by and large conceptual, ideological, and political. We can likely move beyond the conceptual arguments if there are empirical investigations of which models work best, under what conditions, and for what purposes, such as in the study conducted by Shadish et al. (2008). To date, however, there is neither little evidence upon which to base arguments for the use of particular models nor sufficient empirical insight into why that might be the case. So for now, these debates will be primarily driven by and centred on the conceptual and political rather than on empirically derived description or evidence.

## BELIEFS AND KNOWLEDGE ABOUT THE EVALUATION PROCESS

Of course, the debates concerning the primary purpose of evaluation might be only partially resolved with empirical inquiry about evaluation practice. Alkin and Christie (Alkin & Christie, 2004; Christie & Alkin, 2008; Christie & Alkin, in press) argue that theorists hold one of three dimensions of evaluation practice central to their work, based on what they believe to be the primary purpose of evaluation work. And these beliefs and values about the primary purpose of evaluation are influenced and informed by the habits of mind as described by philosophers of science, life experiences, and disciplinary training (Christie & Masyn, 2008). In sum, these debates offer the evaluation field rich conceptual ground for developing prescriptions for practice;

an empirically informed framework for how to best determine the purpose of the evaluation would indeed greatly benefit practitioners.

Empirical examination of prescriptions for evaluation practice should inform when, how, and why particular designs are best implemented, as well as what to expect from these procedures under a defined set of contextual conditions. This, in turn, will help to create a stronger evaluation practitioner base with a unique professional identity and a knowledge base that moves beyond “working knowledge.” As others have suggested (e.g., Kennedy, 1982), the term “working,” when referring to working knowledge, has two meanings: “the special domain of knowledge that pertains to one’s work” and “that the knowledge itself is tentative, subject to change, as the worker encounters new situations or knowledge” (Kennedy, 1982, p. 2). In the evaluation literature, the working knowledge of decision makers has been discussed in the context of how one might use knowledge generated by an evaluation study. However, working knowledge can also describe the methodological and domain-specific knowledge base that practicing evaluators draw upon when studying a program.

Most evaluation practitioners have a working knowledge, that is, an organized body of knowledge concerning evaluation procedures that is used spontaneously and routinely in the context of their work (Christie, 2003). There are four components of working knowledge: formal evidence, experience, interests or goals, and beliefs (including values). An evaluator’s working knowledge includes not just what one has learned from one’s work, but also assumptions about what good studies look like (likely derived from a substantive area of study, such as education), knowledge about the social problem that the program under investigation is designed to address, the political and organizational factors surrounding the evaluation and the program under study, and lastly, the interests and predictions of relevant stakeholders.

As Kennedy (1982) points out, cognitive psychologists have found that working knowledge *alone* does not lead to the best reasoning. Based on a fairly recent study (Christie, 2003), it can be hypothesized that working knowledge has a greater influence on evaluation practitioners’ decisions and practices than formal evaluation theories and scholarship. This study found that the evaluator’s area of academic study was predictive of practice, and a subsequent study corroborated this finding (Christie & Masyn, 2008). In other words, evaluators used little evaluation-specific scholarship to inform their

practical work. Evaluation practice is different from applied research on several important dimensions (Fitzpatrick, Sanders, & Worthen, 2003). Therefore, the notion that most evaluators are drawing upon substantive field research methods (e.g., from the field of education) to inform their evaluation practice is problematic. Evaluation has its own set of competencies, standards, and guidelines for practice that should be used to inform evaluation practice.

Beyond the technical aspects of what the evaluator does and encounters, empirical description and examination of evaluation theories is critical to supporting the training of evaluation practitioners. Without this foundation, they are left to speculate based on what may have worked in the past and in different contexts, and to rely on what has been taught in other disciplines. An evidence base from which evaluation practitioners can draw will lead us away from relying on working knowledge alone to a more explicitly systematic practice.

The argument here, then, is for more evidence-based practice in evaluation. Similar calls for evidence-based practice have become common across most practical fields. Donaldson (2009) jokingly points to this in his introductory chapter of the book, *What Counts as Credible Evidence in Evaluation and Evidence-based Practice?* In this edited volume, a group of esteemed evaluation writers address issues related to the use of the RCT as the “gold standard” for studying questions related to the effectiveness of programs and policies. A similar question can be posed about evaluation: What constitutes credible evidence in evaluation theory and practice? To this question, the evaluation field should not take the more limited view of evidence-based practice, one that is arguably a more prevalent view in other fields, established as effective through scientific research according to some set of explicit criteria (Schwandt, 2005). Instead, I suggest that we embrace what Schwandt (2005) describes as a broader view of evidence-based practice, one that emphasizes the importance of integrating evidence with professional judgement, while taking into account what the client considers valuable in terms of the program and what the evaluation intends to address. Examples of this might include the study questions of interest and the values related to inquiry and knowledge. Therefore, investigations of evaluation theory are critical when endorsing such a definition of evidence-based practice.

It may also prove important to shift our theoretical discourse from one centred on theorists to one more focused on the theories themselves. Evaluation theories offer active propositions that have an

“in order to achieve X, do Y” structure, and an organized set of such propositions constitutes a prescriptive theory. Because prescriptive theories deal with values and what one believes to be the best course of action, they do not easily lend themselves to studies designed to verify principles and they are therefore difficult to test empirically. Instead, empirical studies are useful in that they reveal principles that can then serve as the organizing framework for categorizing various prescriptive theories. It is therefore important to empirically describe the principles of individual theorists’ prescriptive theories of evaluation to reveal the common principles from which organizing frameworks (i.e., theories) can be understood.

## WHAT CAN WE DO TO MOVE FORWARD?

Currently there are more than 75 national and regional evaluation associations with a formal international cooperative network; in 1980 there were three (Donaldson, 2009). This growth in professional associations represents a marked increase in the number of people and institutions, including universities, who have an interest in strengthening our empirical descriptions and understanding of evaluation. Indeed, evaluation is serving a variety of purposes and is being practiced across the world. Professional organizations have an important role in facilitating this work by emphasizing and prioritizing research on evaluation in their conference programs, and by highlighting it as central to promoting the discipline of evaluation.

Research on evaluation will not resolve different views about the proper purposes of evaluation. One theorist may hope to improve the evidence available for decision makers, while another seeks to facilitate social justice, and so on. Perhaps empirical study of evaluation should not support one of these goals versus the other. To better understand the interconnectedness between various theories of evaluation that have resulted from our more individualistic pursuits, we should test some of the more unifying statements about evaluation as well as the common dimensions related to its practice. Empirical investigation of our extant prescriptive theories may lead us to a better understanding of the discriminating and contingent explanations and recommendations for practice and of how and under what conditions they work. This in turn should lead to the development of more contingency theories.

To advance the knowledge base of evaluation we should draw upon the knowledge established by other disciplines—an approach used

in other fields with great success. For instance, there has been a marriage between law and psychology, particularly in the areas of eyewitness identification and jury selection and behaviour. Likewise, as evaluators, we may apply psychological knowledge and knowledge from other disciplines and professions to important substantive evaluation issues. An obvious example—and an area where some of this work has already been conducted—is the empirical psychological work on organizational behaviour and leadership. This literature can be applied to help us understand important evaluation issues related to the contexts in which many of our studies are conducted.

Drawing upon what we know about psychology is logical not only because some formative evaluative thinking was derived from this knowledge base, but also because psychology is concerned with issues closely tied to evaluation, such as human behaviour (e.g., of the evaluator or decision makers), organizational behaviour (e.g., program context), and culture. Although it is not common to see explicit and direct connections to the empirical knowledge of other disciplines in evaluation theory literature, it is reasonable to begin drawing upon such knowledge in order to stimulate further understanding of and research on evaluation.

Admittedly, designing and conducting studies on evaluation is difficult. And while other social scientists—including applied psychologists, sociologists, and political scientists—struggle with understanding problems in complex real world settings, they are also afforded ample opportunity to test theories in laboratory settings. For example, cognitive psychologists typically generate and test theories of memory using experiments where study participants are relatively easy to recruit and where assignment to various test conditions can be manipulated with relative ease. The empirical investigation of evaluation, on the other hand, is focused on how we *understand* real world behaviour. Take Lewin (e.g., Lewin, Lippit, & White, 1939), for example, who studied leadership styles, among other things. His work identified three different styles of leadership around decision making—a real world phenomenon. Evaluation theorists, on the other hand, are concerned with the methods and procedures that can be used to study the phenomenon in practice. As a result, those who study evaluation are concerned with designing studies that are one step removed from the phenomenon in question, presenting a unique set of challenges. For example, because evaluation practice is situational, each context offers its own constituency, programmatic elements, and bureaucratic

hurdles, all of which influence the study designs that can be used and generalizations that can be made.

Nonetheless, broad agendas for conducting research on evaluation have been suggested. Henry and Mark (2003) and Mark (2007) offer current and noteworthy discussions on the different types of and ways of conducting research studies on evaluation. They suggest comparative research on evaluation practice, research on evaluation outcomes, meta-evaluation, and analogue studies. They also describe different strategies for carrying out these studies, including incorporating research studies about evaluation into our funded evaluation studies and encouraging funders of evaluation studies to support funding for research on evaluation with the goal of increasing the ability of evaluation to contribute to social betterment.

We should also consider broader strategies for generating empirical knowledge for practice-based fields such as practitioner-informed research, which is based on practitioners' problematization of their own experiences in daily work, their puzzles, concerns, and worries (Styhre, 2009). This is distinct from action research or participative inquiry. In particular, practitioner-informed research does not imply that the researcher is abandoning the ambition to make traditional academic scholarly contributions to the field. Rather, practitioners are seen as uniquely positioned to offer an *emic*, or insider's, view of practice.

Discussions of practitioner-informed research have typically focused on the research intended to guide practice in a professional area. For example, Cochran-Smith and Lytle (1993) question the common assumption that knowledge about practice should be primarily *outside-in* (e.g., generated by university- or centre-based researchers and then used in schools or programs). They point out that this outside-in assumption implies the unproblematic transmission of knowledge from source to destination and instead call attention to practitioners as knowers, experienced in the complex relationships of knowledge and practice as embedded in local contexts. Thus, practitioner-informed research can be uniquely revealing about the intersection of theory and practice.

The practitioner-informed perspective is distinct from more traditional notions of research, which position research as an activity carried out by a select group comprised mainly of academics. In traditional research involving practice fields, researchers disseminate their find-

ings to practitioners. Research-informed practice, sometimes referred to as evidence-based practice, is believed to be an exemplar for many professions, and suggests a hierarchical relationship not only between theoreticians and practitioners, but also between theory and practice. That is, research generates knowledge, which builds, supports, or tests theory, which in turn informs practice. This suggests a rational model for decision making that can be misleading and lacks important contextual insight that is attentive to nuance and the complexity faced when practicing in the “real world” (Greene, 2009).

For evidence-based practice, the relationship between theory and practice can be characterized as unidirectional, with the expectation that research dissemination will be followed directly by implementation in practice. Schön (1983, 1987) has referred to this unidirectional, hierarchical relationship as “the positivist epistemology of practice” and claims that it has resulted in a growing mistrust by practitioners as to whether academic knowledge can offer anything of relevance to practice situations. This divide is often referred to as a theory-practice gap, and similar descriptions of such a divide exist in the evaluation literature (e.g., Christie, 2003; Mark, 2007). To make sense of the task of increasing the influence of evidence on evaluation practice we must understand the nature of practitioner decision making in real everyday practice.

## CONCLUSION

There is a growing and justifiable concern that theories of evaluation are not more empirically informed, and that the tools of evaluation are not regularly applied to evaluation itself. There are many aspects of evaluation that would benefit from a more empirical foundation and continued empirical development. In particular, this type of attention might move us toward more overarching theories of evaluation and facilitate the integration of our theories into the lexicon, thinking, and practice of those in the field who are conducting evaluation studies.

Evaluation as a field could improve upon the reporting of the empirical research that is conducted and on the topics that are addressed. Therefore, it is unclear where greater emphasis should be applied and how we might develop methods for incentivizing evaluation practitioners and scholars to engage in more empirical work. This gap in our understanding should be an additional focus of empirical inquiry and of meta-analyses of extant empirical work because it would inform leaders in the field about the issues that various constituencies believe

to be important (O'Fallon, Wolfle, Brown, Dearry, & Olden, 2003). A research profile can then be developed that articulates the hierarchies of research areas as prioritized by different groups (Ramirez, Waldman, & Lasser, 1991).

It is important to acknowledge that evaluation theory as it stands today has had a significant impact on our field. Yet, introducing empiricism more prominently will better situate it to achieve even greater influence on the field and the practice of evaluation. While pursuing this empirical work, we must be mindful that in order to bridge the evaluation theory-practice gap, the academic community must take a "realistic view of the limited, indirect, and yet important impact of scholarly knowledge" on evaluators' practice (George & Bennett, 2004, p. 284). Only in this way can we most meaningfully contribute to our goal of improving society and our interactions within it.

## NOTE

- 1 Randomized controlled trials require the random assignment of individuals into a treatment group (the group receiving the intervention or program) or a control group (the group not receiving the intervention or program). Random assignment arguably reduces most threats to internal validity (e.g., selection bias, maturation, history), and the use of a control group provides data to determine the relative performance of the treatment group, helping to answer the "compared to what?" question when measuring program impact (Azzam & Christie, 2007).

## REFERENCES

- Adams, S., & Titler, M. G. (2009). Implementing evidence-based practice. In M. A. Mateo & K. T. Kirchhoff (Eds.), *Research for advanced practice nurses* (1st ed., pp. 329–361). New York, NY: Springer.
- Alkin, M. C. (1991). Evaluation theory development II. In M. McLaughlin & D. Phillips (Eds.), *Evaluation and education at quarter century* (pp. 91–112). Chicago, IL: University of Chicago Press.
- Alkin, M. C., & Christie, C. A. (2004). An evaluation theory tree. In M. C. Alkin (Ed.), *Evaluation roots* (pp. 12–66). Thousand Oaks, CA: Sage.
- Alkin, M., & Ellett, F. (1985). Evaluation models: Development. In T. Husén & T. N. Postlethwaite (Eds.), *The international encyclopedia of education* (Vol. 3, pp. 1760–1766). Oxford, UK: Pergamon Press.

- Braskamp, L. A., Brown, R. D., & Newman, D. L. (1978). The credibility of a local educational program evaluation report: Author source and client audience characteristics. *American Educational Research Journal*, 15(3), 441–450.
- Brown, R. D., Newman, D. L., & Rivers, L. (1980). Perceived need for evaluation and data usage as influencers on an evaluator's impact. *Educational Evaluation and Policy Analysis*, 2(5), 67–73.
- Chelimsky, E. (1998). The role of experience in formulating theories of evaluation practice. *American Journal of Evaluation*, 19(1), 35–55.
- Chen, H. (1990). *Theory-driven evaluation*. Newbury, CA: Sage.
- Christie, C. A. (2003). What guides evaluation? A study of how evaluation practice maps onto evaluation theory. *New Directions for Evaluation*, 97, 7–36.
- Christie, C. A., & Alkin, M. C. (2008). Evaluation theory tree re-examined. *Studies in Educational Evaluation*, 34(3), 131–135.
- Christie, C. A., & Alkin, M. C. (in press). An evaluation theory tree. In M. C. Alkin (Ed.), *Evaluation roots* (2nd ed.). Thousand Oaks, CA: Sage.
- Christie, C. A., & Masyn, K. E. (2008). Profiles of evaluators reported practice: A latent profile analysis. *Canadian Journal of Program Evaluation*, 23(2), 225–254.
- Cochran-Smith, M., & Lytle, S. L. (1993). *Inside/outside: teacher research and knowledge*. New York, NY: Teachers College Press.
- Cousins, J. B., Goh, S., Clark, S., & Lee, L. (2004). Integrating evaluative enquiry into the organizational culture: A review and synthesis of the knowledge base. *Canadian Journal of Program Evaluation*, 19(2), 99–141.
- Cousins, J. B., & Leithwood, K. A. (1986). Current empirical research on evaluation utilization. *Review of Educational Research*, 56(3), 331–364.
- Donaldson, S. (2009). In search of the blueprint for an evidence-based global society. In S. Donaldson, C. Christie, & M. Mark (Eds.), *What counts as credible evidence in evaluation and evidence-based practice?* (pp. 2–18). Thousand Oaks, CA: Sage.

- Donaldson, S. Christie, C. A., & Mark, M. (Eds.). (2009). *What counts as credible evidence in evaluation and evidence-based practice?* Thousand Oaks, CA: Sage.
- Fitzpatrick, J. L., Sanders, J. R., & Worthen, B. R. (2003). *Program evaluation: Alternative approaches and practical guidelines* (3rd ed.). Boston, MA: Allyn & Bacon.
- George, A. L., & Bennett, A. (2004). *Case studies and theory development in the social sciences*. Cambridge, MA: MIT Press.
- Greene, J. C. (2009). Evidence as “proof” or evidence as “inkling”. In S. Donaldson, C. Christie, & M. Mark (Eds.), *What counts as credible evidence in evaluation and evidence based practice?* (pp. 153–167). Thousand Oaks, CA: Sage.
- Heberger, A., Christie, C. A., & Alkin, M. (2010). Understanding the influence of evaluation on other disciplines: A bibliometric analysis. *American Journal of Evaluation*, 31(1), 24–44.
- Henry, G. T., & Mark, M. M. (2003). Toward an agenda for research on evaluation. *New Directions for Evaluation*, 97, 69–80.
- Kennedy, M. M. (1982). *Working knowledge and other essays*. Cambridge, MA: Huron Institute.
- King, J. (1982). Studying the local use of evaluation: A discussion of theoretical issues and an empirical study. *Studies in Educational Evaluation*, 8(2), 175–183.
- Lewin, K., Lippit, R., & White, R. K. (1939). Patterns of aggressive behavior in experimentally created social climates. *Journal of Social Psychology*, 10, 271–301.
- Mark, M. (2007). Building a better evidence base for evaluation theory. In N. L. Smith & P. R. Brandon (Eds.), *Fundamental issues in evaluation* (pp. 111–134). New York, NY: Guilford Press.
- Mark, M., Greene, J. C., & Shaw, I. (2006). Handbook of evaluation: Introduction. In I. Shaw, J. Greene, & M. Mark (Eds.), *Handbook of evaluation*. London, UK: Sage.
- Mark, M., Henry, G., & Julnes, G. (2000). *Emergent realist evaluation: An integrated framework for understanding, guiding, and improving policies and programs*. San Francisco, CA: Jossey-Bass.

- Miller, R. (2010). Developing standards for empirical examinations of evaluation theory. *American Journal of Evaluation*, 31(3), 390–399.
- O'Fallon, L. R., Wolfe, G. M., Brown, D., Dearry, A., & Olden, K. (2003). Strategies for setting a national research agenda that is responsive to community needs. *Environmental Health Perspectives*, 111, 1855–1860.
- Patton, M. (2008). *Utilization-focused evaluation* (4th ed.). Thousand Oaks, CA: Sage.
- Patton, M. Q., Grimes, P. S., Guthrie, K. M., Brennan, N. J., French, B. D., & Blyth, D. A. (1977). In search of impact: An analysis of the utilization of federal health evaluation research. In C. H. Weiss (Ed.), *Using social research in public policy making* (pp. 141–163). New York, NY: D.C. Health.
- Ramirez, G. G., Waldman, D. A., & Lasser, D. J. (1991). Research needs in corporate finance: Perspectives from financial managers. *Financial Management*, 20, 17–29.
- Schön, D. A. (1983). *The reflective practitioner*. London, UK: Temple Smith.
- Schön, D. A. (1987). *Educating the reflective practitioner*. New York, NY: Basic Books.
- Schwandt, T. (2005). The centrality of practice to evaluation. *American Journal of Evaluation*, 26(1), 95–105.
- Scriven, M. (1995). The logic of evaluation and evaluation practice. *New Directions for Program Evaluation*, 68, 49–70.
- Scriven, M. (2003). Evaluation in the new millennium: The transdisciplinary vision. In S. I. Donaldson & M. Scriven (Eds.), *Evaluating social programs and problems: Visions for the new millennium* (pp. 19–42). Mahwah, NJ: Lawrence Erlbaum Associates.
- Shadish, W. R. (1998). Evaluation theory is who we are. *American Journal of Evaluation*, 19(1), 1–19.
- Shadish, W. R., Clark, M. H., & Steiner, P. M. (2008). Can nonrandomized experiments yield accurate answers? A randomized experiment comparing random and nonrandom assignments. *Journal of the American Statistical Association*, 103(484), 1334–1343.
- Shadish, W. R., Cook, T. D., & Leviton, L. C. L. (1991). *Foundations of program evaluation*. Newbury Park, CA: Sage.

- Shulha, L. M., & Cousins, J. B. (1997). Evaluation use: Theory, research, and practice since 1986. *American Journal of Evaluation*, 18(3), 195–208.
- Smith, N. L. (1983). Citizen involvement in evaluation: Empirical studies. *Studies in Educational Evaluation*, 9, 105–117.
- Styhre, A. (2009). The phantom menace: Conducting practitioner-informed research without losing academic liberties. In J. W. Cox, T. G. LeTrent-Jones, M. V. Voronov, & D. Weir (Eds.), *Critical management studies at work: Negotiating tensions between theory and practice* (pp. 29–39). Cheltenham, UK: Edward Elgar.

**Christina A. Christie** is Associate Professor and Division Head, Social Research Methodology, Department of Education, Graduate School of Education and Information Studies, University of California, Los Angeles.