Title
Unbundling evaluation use

Journal
Studies in Educational Evaluation, 29(1)

ISSN
0191-491X

Authors
Alkin, MC
Taut, SM

Publication Date
2003-05-01

DOI
10.1016/S0191-491X(03)90001-0

Peer reviewed

PERGAMON

Studies in Educational Evaluation 29 (2003) 1–12

Studies in Educational Evaluation

www.elsevier.com/stueduc

UNBUNDLING EVALUATION USE

Marvin C. Alkin and Sandy M. Taut

Graduate School of Education and Information Studies, University of California Los Angeles, USA

Evaluation use (or evaluation utilization) refers to the way in which an evaluation and information from the evaluation impacts the program that is being evaluated. Did the evaluation generate a new understanding of certain aspects of the program? Were progressive changes made in the program as a consequence of the evaluation? Did program staff acquire new skills and insights during the course of, and attributable to, the evaluation?

The concern for evaluation use initially developed in the United States during the 1960s. Major social program initiatives, like *Head Start* in education, began to include the mandate for end-of-year evaluation reports. Many of those conducting these evaluations lacked awareness of the unique characteristics of evaluation as opposed to research. The result was a multitude of evaluation reports with little relevance and utility to their potential users, simply fulfilling the requirement of evaluation. These developments triggered the fledgling evaluation profession to strive for a greater understanding of evaluation use, particularly of how the likelihood of evaluation use could be increased.

Although there has been great progress, to this day confusion persists in the literature regarding evaluation use. Indeed, it seems like the evaluation use concept has come under attack from several different directions. On the one hand, there are those who argue about the relevance and appropriateness of the "use" term and prefer "influence" (e.g., Kirkhart, 2000). On the other hand, some have tried to equate or show the strong relationship between evaluation use and knowledge use, generally ignoring the differences in context (e.g., Cousins & Shulha, 2002). In this article, we will provide a conceptual

0191-491X/03/$ - see front matter
PII: S0191-491X(03)00002-6

framework for considering the use – influence distinction, as well as the relationship between evaluation use and knowledge use. We will clarify and expand some of the well-established definitions of concepts related to evaluation use.

Systematic Knowledge

First, as background, it is helpful to discuss the relationship between the following three concepts: information, knowledge, and systematic knowledge (see Figure 1).

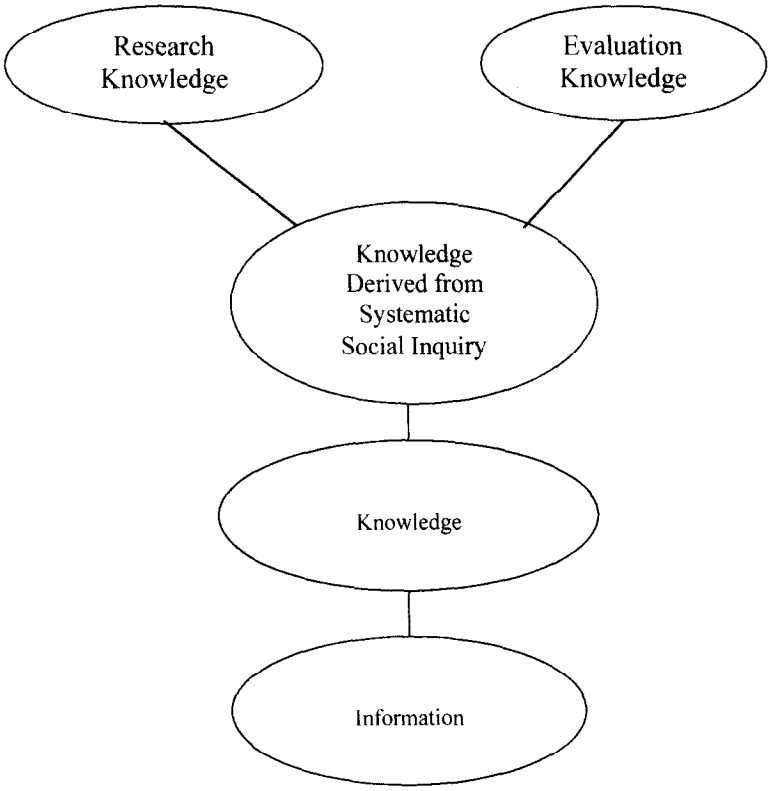

Figure 1: Evaluation Knowledge Basis

Basically, information is data that have been collected in some manner. This could be as informal as an observation of the color and lines of an automobile considered for purchase or as well-structured as data collected using a standardized test or other carefully designed instrument. However, information does not become knowledge until it has been interpreted in some way. Thus, the indication of the preference of a particular color and the extent to which the automobile color matches that preference, or the meaning of the test or questionnaire data in terms of standards that have been established constitute a conversion

of information into knowledge. It is important to note that what one considers knowledge depends largely on the epistemological position one holds. For example, the objectivist notion that knowledge is value-free is challenged by interpretivist and realist perspectives. The former asserts that there are multiple interpretations of the same information, with one view no more valid than the next, while the latter takes into account power inequities and value systems that cause some interpretations to be regarded as more valid than others.

Within both evaluation and research we are concerned not just with any knowledge but with systematic knowledge. Systematic knowledge is based on information collected using systematic social scientific procedures. Systematic speaks to the way the information is collected, analyzed and communicated. The process is based on a logical chain of reasoning linking empirical observations and conclusions in a coherent, shareable, and persuasive way with the aim to ensure objectivity, reliability and validity of findings (National Research Council, 2002). According to Owen (2002), traditional academic social research as well as policy analysis, creative processes producing innovations, and program evaluation are all based on systematic social inquiry.

Research and Evaluation

It is important to focus on the distinctions between knowledge produced as a part of research and evaluation knowledge. Whereas both research and evaluation knowledge are produced in a similar fashion, the purposes and uses of such knowledge are different. In the case of research, the goal is generalizable knowledge—contributions to the body of knowledge in a particular field of research that hold true across all settings, times, and for all individuals represented by the produced knowledge. General theories are built by testing and refuting claims or hypotheses. In the case of evaluation, the purpose is context-specific knowledge—that is, knowledge that is applicable only within a particular setting at a particular point in time, and intended for use by a particular group of people. Cronbach and Suppes (1969) make this same distinction, referring to *conclusion-oriented research* and *decision-oriented evaluation*.

Borrowing from Scriven's (1967) definition of evaluation, evaluation is different from research in terms of its unique purpose to *judge* the worth, merit, or quality of an evaluand, as well as in terms of the context-specific uses of the knowledge produced to make these judgments. This valuing component is the reason why the *applied* versus *basic* research dichotomy that is common in the research community does not adequately capture the evaluation–research distinction.

The Concept of Evaluation Use

After defining the concept of evaluation knowledge, let us now consider the factors influencing the production of such knowledge, and how this relates to evaluation use. Clearly, evaluation knowledge is a product of the conduct of an evaluation, i.e. evaluation knowledge can be created not only as an end-product, but also at numerous points during the course of the evaluation. For example, the evaluation of a youth HIV/AIDS prevention program has the potential of producing knowledge not only by disseminating a report at the end of the first year of program implementation, but also by causing the program staff to

reflect on their practice by asking them to give feedback on survey questions, or in informal conversations about the program's underlying theory.

Furthermore, evaluation knowledge is produced in relationship to a particular program. Thus, while the evaluation itself determines the nature of the produced evaluation knowledge, other factors come into play as well – and all these factors interact with each other. User characteristics (the people involved) and context factors (the situation or background surrounding the program being evaluated) also have an impact on the produced evaluation knowledge (see Figure 2). These same features have been found in prior research to influence the use of evaluation findings. In a systematic examination of factors affecting evaluation use, Alkin (1985) differentiates between "evaluation factors," "human factors," and "context factors."

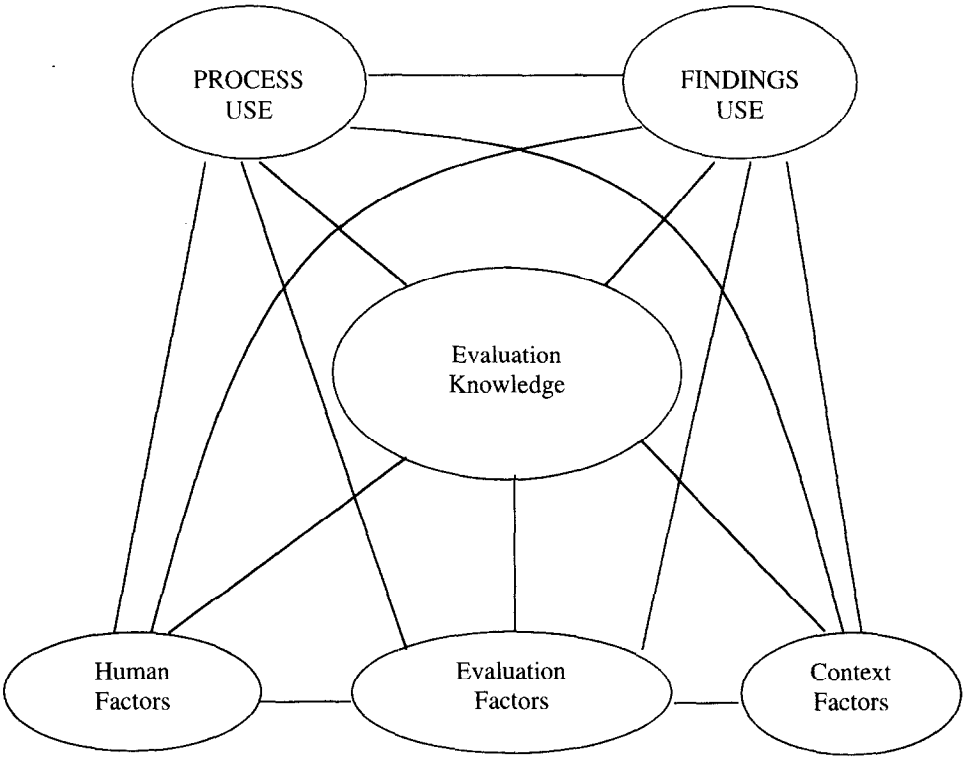

Figure 2: Factors Affecting Evaluation Use

Let us first discuss how features of the evaluation itself (evaluation factors) shape the production of evaluation knowledge. First of all, the particular evaluation procedures are of concern. What kind of design was used? Was there a comparison program? What kinds of instruments were used for data collection? What kinds of analysis or aggregation procedures were employed? Second, the consideration of the evaluation includes

information transmission, both during the conduct of the evaluation as well as at its conclusion. Third, there are varying degrees of information dialogue that can take place during the course of the evaluation between the evaluator and potential users (stakeholders). Fourth, the way in which evaluation knowledge is presented and its reporting also determines the nature of evaluation knowledge being produced. Finally, included also in the evaluation factor are the personal characteristics and actions of the evaluator, for example, the evaluator's commitment to use, willingness to involve users, political sensitivity, credibility, rapport with users etc.

Both the kind of evaluation knowledge an evaluation produces, and how this knowledge is used, is also a function of the particular program context (context factors). Context factors encompass such characteristics as pre-existing evaluation bounds, including contractual obligations and fiscal constraints, inter- and intra-organizational characteristics, as well as external community factors. The characteristics of the individual project or program, like its maturity or age, would also be considered as context factors.

The human factors are characterized by Alkin (1985) as both evaluator and user characteristics. In this paper, we chose to consider the evaluator characteristics as part of the evaluation factors (and have already discussed this) and will focus on user characteristics as the third major element. User characteristics subsume such features as the identity of the user including organizational responsibility and various personal and professional style characteristics. Most important, however, is the potential users' interest in evaluations generally (and in this evaluation in particular), and their commitment to use.

Summaries of research on factors associated with evaluation use are provided by Cousins and Leithwood (1986), Shulha and Cousins (1997), and Hofstetter and Alkin (in press). This extensive body of research, however, has focused on the use of evaluation *findings* and has not, to any great extent, considered the way that the evaluation process is used. Thus, the major research on use has in essence addressed the following question: How has evaluation knowledge (findings) been converted into what are called *instrumental* and *conceptual* use?

This distinction between use of instrumental and conceptual findings was first suggested by Rich (1977). Instrumental use refers to instances where knowledge has been used to impact on direct action such as making particular decisions about a program. Conceptual use (sometimes referred to as *enlightenment* use) describes instances where no direct decision has been made, but where particular conceptual understandings have been modified, relating to changes in the way users think about particular aspects of a program. Also prevalent in the evaluation literature is the concept of *symbolic* use addressing situations where evaluation was used to either justify a decision that had been previously made or to demonstrate that a program was willing to be evaluated, thus enhancing the reputation of the program manager or decision-maker. Owen (1999) has distinguished between these two kinds of evaluation use (justifying a prior decision versus enhanced reputation), referring to the former as *legitimative* – that is, legitimating a prior decision. The term symbolic evaluation has been reserved for those instances where evaluation is conducted simply as a symbol or a status action.

More recently, the term *process use* has come into prominence, addressing not the use of the evaluation findings, but referring to the manner in which the conduct of the evaluation (the evaluation process) impacts on individuals or organizations. Patton (1997,

p. 90) defines process use as "individual changes in thinking and behavior, and program or organizational changes in procedures and culture, that occur among those involved in evaluation as a result of the learning that occurs during the evaluation process." If, for example, the program is altered due to the thinking process that the evaluation triggers, rather than as a consequence of the findings, then this is an indication of process use. Others, like Preskill and Torres (2000), focus more specifically on the learning effect that the evaluation process has at the individual, team, and organizational levels.

Does the evaluation process produce evaluation knowledge? If so, how is process use of evaluation knowledge different from findings use of evaluation knowledge? In the case of process use, unlike findings use, it is not only the evaluation knowledge that is or is not being used for learning. Consulting the social psychological literature (Fischer & Wiswede, 1997; Aronson, Wilson & Akert, 1999), we found that learning has two components that interact with each other: knowledge acquisition and accumulation on the one hand, behavior acquisition and modification on the other hand. In the case of findings use, the former type of learning dominates. In the case of process use, however, the conduct of the evaluation itself also enables potential users to acquire new skills and to modify their behavior. For example, staff of a substance abuse program might during the course of the evaluation (because of the questions asked by the evaluator and the meta-perspective introduced through evaluation) get a chance to reflect on their interactions with their clients. They might detect certain behavior patterns that turn out to be ineffective, resulting in a reshaping of these interaction patterns as the evaluation unfolds. We presume that the way in which the evaluation is conducted, the context, and the users (as described above) have impact on whether and in what form process use takes place (see Figure 2).

Refining Evaluation Use Distinctions

We have already referred to the distinctions between findings use and process use. Typically, use, referring to findings use, has been examined in terms of the categories instrumental, conceptual and symbolic (or legitimative). These concepts are fairly well understood when considering use of outcomes or findings (see above). We maintain that process use is not another category on a par with these types of use but rather another domain of use. That is, findings may be used or process may be used (see Figure 3). Moreover, process use may occur instrumentally or conceptually. Thus, as a result of the process, an instrumental change or a decision may be made. For example, questions by the evaluator about the nature of the program may lead to reconsiderations, reexaminations and modifications of the program. This we believe to be instrumental process use. Process use may also occur conceptually. The evaluation process may change the attitudes of users about the importance of evaluation or about the potential role of multiple stakeholders in decision-making. Greene (1988, p. 111) describes such "conceptual uses of the evaluation process." She posits that a participatory evaluation process enhances the likelihood of utilization of the evaluation results *by way of* conceptual process use, e.g. by contributing to the stakeholders' "readiness for utilization." Further research is necessary to explore the nature of possible links between conceptual / instrumental process use and findings use.

Symbolic use and the legitimative use counterpart suggested by Owen (1999) provide another interesting distinction. Legitimative, as suggested, refers to the

legitimization of previous decisions. Presumably, this refers to the findings of the evaluation study demonstrating that a previously made decision was in fact a correct one. Thus, legitimative use is clearly within the findings use category. On the other hand, symbolic use refers to the process of conducting evaluation as a symbolic act, which might add to the stature of the decision maker or the program or whatever. Symbolic, thus, refers specifically to a kind of legitimization through engaging in the process without particular heed to the findings. Thus, it seems reasonable to us to differentiate between legitimative and symbolic in terms of findings use, and process use respectively. Figure 3 summarizes these distinctions.

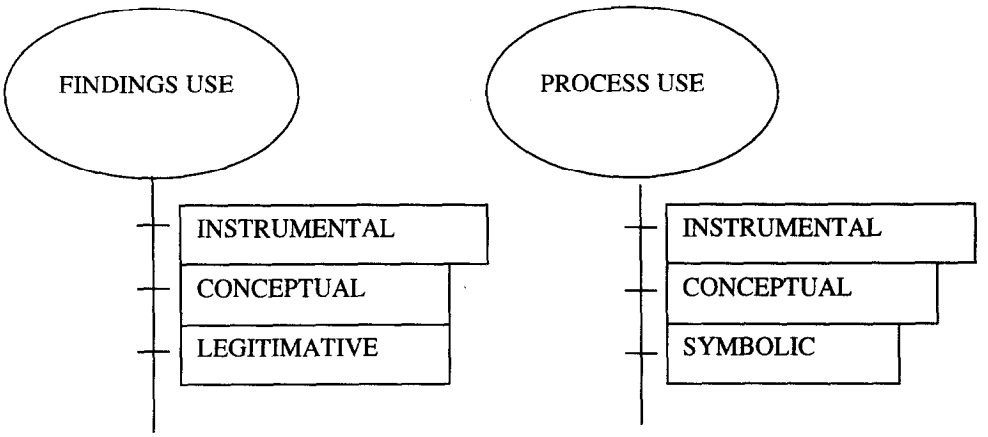

Figure 3: Types of Evaluation Use

A further distinction may be in order. Instrumental and conceptual findings use may refer to evaluation knowledge produced either formatively or summatively. Legitimative use, however, is by its very definition only summative. It is easy to mistakenly conclude that some particular instance of formative findings use is process use. An example of this is provided in an article by Russ-Eft, Atwood and Egherman (2002). The authors describe program changes that occurred due to the involvement of the evaluators in program delivery. The authors call these program changes examples of either "unplanned instrumental use of a formative nature" or "process use." It is important to clarify the distinction between these two. The former is the use of interim evaluation findings produced and communicated by the evaluators to make program changes, the latter is the learning that takes place with stakeholders through their engagement in the evaluation process, as opposed to their reactions to evaluation findings produced and shared by the evaluators. To rightly talk about process use, the stimulus that causes instrumental or conceptual changes to take place must be the *participation in the evaluation process*, not the acknowledgement of evaluation findings.

Moreover, process use, as we have discussed it, can take place at different points in time. Instrumental or conceptual process use may occur during the course of the evaluation

(see examples above) and at the conclusion of an evaluation. In the latter instance, we are considering the net result and impact of the participation in the complete evaluation process. In a participatory evaluation, for example, one might expect that evaluation skills are acquired by the conclusion of the evaluation process (instrumental process use), or that attitudes have changed that could be applied to other situations (conceptual process use). The impact of the evaluation process on organizational development (the area of concern of numerous recent theorists; see Preskill & Torres, 1999) might likewise occur during the evaluation process or at the conclusion of the process. Further, increased appreciation of the value of engaging in the evaluation process may be acquired both during the process, as well as by the conclusion of the process.

Evaluation Influence

Now let us consider *influence*. Kirkhart (2000) introduced the notion of evaluation influence as an important area of study. She defines influence as "the capacity or power of persons or things to produce effects on others by intangible or indirect means" (p. 7). She further notes that influence extends the notion of use to include "effects that are multidirectional, incremental, unintentional, and noninstrumental, alongside those that are unidirectional, episodic, intended, and instrumental (which are well represented by the term *use*)" (p. 7). Her integrated theory of influence encompasses three dimensions: intention (unintended vs. intended), source (process vs. results), and time (immediate, end-of-cycle, long-term).

It is important to note that Kirkhart (2000) did not suggest influence as a replacement for the notion of use. Instead, she argues that the question of how and to what extent evaluations "shape, affect, support, and change persons and systems" (p. 7), can only be adequately addressed by broadening the understanding of the impact of evaluation from narrowly defined use to influence; in her words, "... the term use is an awkward, inadequate, and imprecise fit with non-results-based applications, the production of unintended effects, and the gradual emergence of impact over time" (p. 6). We agree with Kirkhart that the very broad question she poses concerning all possible impacts of an evaluation calls for a very broad framework of evaluation influence, one that depicts the notion of "indirect and intangible" as specified in her definition of influence. However, we argue that the question of how and to what extent the processes and findings of an evaluation lead to intended use by intended users is a more narrow, yet equally important question. We do not concur with Kirkhart that a shift in terminology from use to influence is necessary to more accurately express the process vs. results distinction, i.e. to correctly position process use at the same level with findings use, instead of as "erroneously tacked on to the recognized typology of results-based use" (p. 9). In fact, we would assert that both use and influence can be characterized in terms of a process-results dimension. Likewise, both use and influence can happen immediately, end-of-cycle, or long-term – depending on the definition of these terms. In our understanding, end-of-cycle evaluation does not end at the presentation of the report. The evaluator interested in use can anticipate and work towards impact in a finite period following the conclusion of the evaluation. It is difficult to precisely define this period, but we think of perhaps a one-year term, depending on the program context and the evaluation users.

Regarding the unintended – intended dimension of evaluation impact, instead of *intention* we would talk about *awareness* of evaluation impacts. Along this dimension we would then distinguish between *aware / intended*, *aware / unintended*, and *unaware / unintended*. In our understanding, the concept of influence adds to the concept of use in instances in which an evaluation has *unaware / unintended* impacts. These impacts constitute an important aspect of an evaluator's work – that merits attention from evaluation research and practice. However, it seems not as essential to the evaluation profession as the impacts that are of a conscious (either intended or unintended) nature, in the eyes of the users, and hopefully of the evaluator as well.

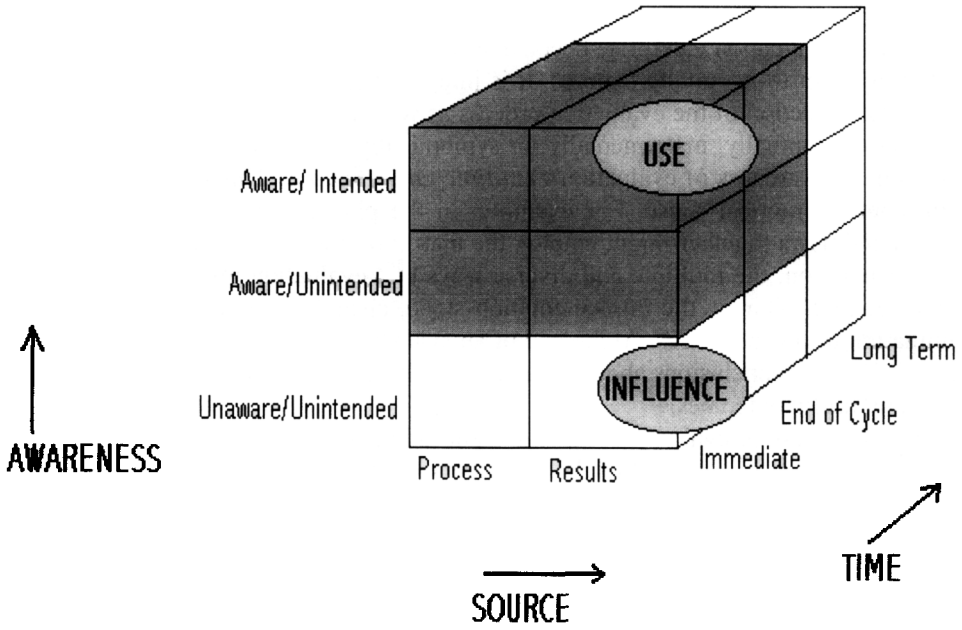

Figure 4: Evaluation Impact Within the Program Context (adapted from Kirkhart, 2000)

Our argument is based on the conviction that the informed, user-oriented evaluator as a part of the evaluation is able to impact users and contexts and therefore has some potential for increasing the possibility of use (of all kinds, including process use). Because, by definition, influence is unintended, it is outside the domain of the evaluator to affect such possible evaluation influences. Thus, while the likelihood of evaluator actions in increasing evaluation use is great, the likelihood of evaluators' actions increasing influence is not. This does not mean that all potential uses must be anticipated from the onset of an evaluation; therefore the aware / intended – aware / unintended distinction. That is, the variety of uses that evaluation can have at some point during the course of the evaluation

must at each specific point in time be of a conscious nature, from the perspective of the then intended users, but they can be ones not intended by the evaluator or the users at the outset of an evaluation. By aware or conscious we mean that the primary intended users, if asked and willing to do so, can specify the uses that the evaluation attempts to have, at a particular point in time and in a particular context. All the impacts that users are not aware of, i.e. that they cannot name, we would consider influence. To summarize this section on influence, we used Kirkhart's (2000) influence framework figure and adapted it to our arguments (see Figure 4).

Implications

Isn't it enough to know that evaluation should be useful (see Program Evaluation Standards, Joint Committee, 1994) – why are the above distinctions and redefinitions important? Most fundamentally, unbundling the notion of evaluation use can help everyone involved in the evaluation enterprise become aware of its diverse nature, while at the same time alerting them to its unique features. It is important to make clear that evaluation use can occur as a function of the evaluation process as well as due to its findings, and in both cases either conceptually, instrumentally, or symbolically / legitimatively. To do so raises user and client awareness of evaluation's *multiple contributions*. This knowledge can also help in a more practical sense. For example, in the planning stage of the evaluation, evaluators and users together might employ the matrix of potential evaluation uses to plan for, or speculate on, the multiple and diverse ways that evaluation might be of use in the particular context. During the implementation stage, changing contexts might call for a reexamination of potential uses. At the end of an evaluation cycle, all those involved could come to more valid conclusions about the extent to which the evaluation served the purpose of producing useful outcomes.

An awareness of evaluation's intended and unintended impacts is another defining characteristic of *use*, as opposed to the unawareness and unintentionality of evaluation's *influence*. This distinction is important because evaluators can only try to achieve those impacts that can be addressed and discussed together with potential users, at any point in time during the evaluation process. Influences of evaluation are undoubtedly of importance, but they are unintended and cannot be addressed until after they have occurred. Practicing evaluators have to do their best in actively ensuring and promoting evaluation use, while at the same time noting evaluation influences that might occur but which are outside their sphere of action.

Perhaps the most fundamental issue to be raised is the distinction between *evaluation* knowledge on the one hand, and knowledge produced for *research* on the other hand. Based on our own experiences, confusion still exists regarding this distinction. This is problematic because of the different expectations associated with knowledge produced for research and evaluation. Research knowledge does not have a primary emphasis on usefulness; its main characteristics are scientific rigor and generalizability. If evaluation is measured against these standards it is bound to fail. In addition, such standards make us miss the great opportunity to ensure useful, context-specific, yet systematic data-based evaluation processes and findings that are tailored to stakeholder needs. Therefore,

educating evaluation clients and users about these distinctions is an important task for the evaluation profession.

References

Alkin, M.C. (1985). *A guide for evaluation decision makers*. Beverly Hills, CA: Sage.

Aronson, E., Wilson, T., & Akert, R. (1999). *Social psychology* (3rd ed.). Englewood Cliffs: Prentice-Hall.

Cousins, J.B., & Leithwood, K.A. (1986). Current empirical research on evaluation utilization. *Review of Educational Research, 56* (3), 331-364.

Cousins, J.B., & Shulha, L.M. (2002, November). *Recent developments in inquiry about knowledge and research utilization*. Paper presented at the Annual Meeting of the American Evaluation Association, Washington, DC.

Cronbach, L. J., & Suppes, P. (1969). *Research for tomorrow's schools: Disciplined inquiry for education* (pp. 197-222). Boston, MA: Kluwer.

Fischer, L., & Wiswede, G. (1997). *Grundlagen der Sozialpsychologie* [Foundations of Social Psychology]. München: Oldenbourg.

Greene, J.C. (1988). Stakeholder participation and utilization in program evaluation. *Evaluation Review, 12* (2), 91-116.

Hofstetter, C.H., & Alkin, M.C. (in press). Evaluation use revisited. In D.L. Stufflebeam, T. Kellaghan, & L. Wingate (Eds.), *International handbook of educational evaluation*. Boston, MA: Kluwer.

Joint Committee on Standards for Educational Evaluation (1994). *The program evaluation standards* (2nd ed.). Thousand Oaks, CA: Sage.

Kirkhart, K.E. (2000). Reconceptualizing evaluation use: An integrated theory of influence. *New Directions for Evaluation, 88*, 5-23.

National Research Council (2002). *Scientific research in education*. Committee on Scientific Principles for Education Research. Shavelson, R.J., & Towne, L. (Eds.), Center for Education. Division of Behavioral and Social Sciences and Education. Washington, DC: National Academic Press.

Owen, J.M., with Rogers, P. (1999). *Program evaluation: Forms and approaches* (2nd ed.). London: Sage.

Owen, J.M. (2002, November). *Linking evaluation use to the research utilisation literature*. Paper presented at the Annual Meeting of the American Evaluation Association, Washington, DC.

Patton, M.Q. (1997). *Utilization-focused evaluation: The New Century text*. Thousand Oaks, CA: Sage.

Preskill, H., & Torres, R.T. (1999). *Evaluation inquiry for learning in organizations*. Thousand Oaks, CA: Sage.

Preskill, H., & Torres, R.T. (2000). The learning dimension of evaluation use. *New Directions for Evaluation, 88*, 25-37.

Rich, R.F. (1977). Uses of social science information by federal bureaucrats: Knowledge for action versus knowledge for understanding. In C.H. Weiss (Ed.), *Using social research in public policy making* (pp. 199-211). Lexington, VA: Lexington Books.

Russ-Eft, D., Atwood, R., & Egherman, T. (2002). Use and non-use of evaluation results: Case study of environmental influences in the private sector. *American Journal of Evaluation, 23* (1), 19-31.

Scriven, M. (1967). The methodology of evaluation. In R.E. Stake (Ed.), *Curriculum evaluation.* American Education Research Monograph Series on Evaluation, No. 1. Chicago, IL: Rand McNally.

Shulha, L.M., & Cousins, J.B. (1997). Evaluation use: Theory, research and practice since 1986. *Evaluation Practice, 18* (3), 195-208.

The Authors

MARVIN C. ALKIN is a professor and chair of the Social Research Methods Division in the Graduate School of Education and Information Studies at the University of California Los Angeles (UCLA), USA. He has written extensively on evaluation utilization and evaluation theory. He founded and directed the UCLA Center for the Study of Evaluation.

SANDY M. TAUT received her *Diplom* in Psychology from the University of Cologne, Germany, and is currently a Ph.D. candidate under the mentorship of Professor Alkin. Her country-comparative dissertation deals with use and participation in the context of school evaluation. She gained practical evaluation experience in diverse settings, with a recent focus on international development assistance.

Correspondence: <alkin@gseis.ucla.edu> or <staut@ucla.edu>