American Journal of Evaluation

The Power of Why: Engaging the Goal Paradox in Program Evaluation

Victor J. Friedman, Jay Rothman and Bill Withers

The online version of this article can be found at:

Published by:

$\text{\$}$SAGE Publications

http://www.sagepublications.com

On behalf of:

American Evaluation Association

Additional services and information for *American Journal of Evaluation* can be found at:

Email Alerts: http://aje.sagepub.com/cgi/alerts

Subscriptions: http://aje.sagepub.com/subscriptions

Reprints: http://www.sagepub.com/journalsReprints.nav

Permissions: http://www.sagepub.com/journalsPermissions.nav

Citations (this article cites 6 articles hosted on the
SAGE Journals Online and HighWire Press platforms):
http://aje.sagepub.com/cgi/content/abstract/27/2/201#BIBL

The Power of Why

Engaging the Goal Paradox in Program Evaluation

Victor J. Friedman

Max Stern Emek Yezreel College

Jay Rothman

The Aria Group, Inc.

Bill Withers

R&B, Inc.

Abstract: Clearly defined and measurable goals are commonly considered prerequisites for effective evaluation. Goal setting, however, presents a paradox to evaluators because it takes place at the interface of rationality and values. The objective of this article is to demonstrate a method for unlocking this paradox by making goal setting a process of evaluating goals, not simply defining them. Goals can be evaluated by asking program stakeholders why their goals are important to them. Systematic inquiry into goals also prepares the ground for setting consensual goals that express what stakeholders really care about. This article describes the method, provides a case illustration, offers guidelines for practice, and discusses the method in the context of the evaluation literature on goals and goal setting.

Keywords: goal setting; evaluating goals; values inquiry

As evaluation emerged as an independent field within the social sciences, it became closely identified with the measurement of goal attainment (Morris & Fitz-Gibbon, 1978; Patton, 1997). Most mainstream approaches to evaluation regard specific, clear, and preferably measurable goals as an essential prerequisite for a effective evaluation (Rossi, 1972; Weiss, 1972; Worthen & Sanders, 1973). As a consequence, part of evaluators' role has been to help define and clarify program goals to provide proper foundations for evaluations (Patton, 1997).

The process of goal clarification, however, has been portrayed as a problematic part of program evaluation. Patton (1997), for instance, wrote that "there may be no more deadly way to begin an evaluation effort than assembling program staff to identify and clarify program goals and objectives" (p. 149). He pointed out that goal setting often leads either to a manipulative and futile "goals-clarification game" or to intense "goal wars" in which program stakeholders become locked in battle over the control of a program's direction.

Victor J. Friedman, Senior Lecturer, Max Stern Emek Yezreel College, Emek Yezreel 19300, Israel; phone: 972-4-6423060; fax: 972-4-6423422; e-mail: victorf@yvc.ac.il.

201

Patton (1997) attributed the difficulty of goal setting to the fact that evaluators "live inside" the following paradox:

> Statements of goals emerge at the interface between the ideals of human rationality and the reality of diverse human values and ways of thinking. Therein lies their strength and their weakness. Goals provide direction for action and evaluation, but only for those who share in the values expressed by goals. (p. 174)

The strength of goals is that they direct programs by focusing actions on specific outcomes. The weakness of goals is that they are a source of potential conflict because different program stakeholders may hold very different values. Patton implied that these conflicts present a particular problem because goals are rooted in values that do not lend themselves to rational analysis or assessment. The paradox appears to be that the very source of rationality and focus in program design is also the source of irrationality and divisive conflict.

The objective of this article is to offer evaluators a simple but powerful method for working through this goal paradox. It argues that program goals should be evaluated, not simply clarified, at the outset of a program. Goals can be evaluated by asking program stakeholders why their goals are important to them. As will be seen, a structured, systematic, and deep inquiry into the "why" question provides a rational means for deliberating about human values. This inquiry process provides a means for goal refinement and value alignment that also fosters team building and collaboration.

Because of our own commitment to linking theory with practice, this article addresses two audiences: those with a theoretical orientation as well as those with a practical orientation to evaluation. The first section has a theoretical focus, describing how our evaluation approach deals with the goal paradox by drawing on ideas from the theory and practice of conflict resolution (Rothman, 1992, 1997; Rothman & Friedman, 2002) and action science (Argyris, Putnam, & Smith, 1985; Friedman, 2001a, 2001b; Schön, 1983). The second and third sections have a more practical focus, providing an illustration of our approach and concrete guidelines for putting it into practice. The final two sections return to a more conceptual analysis of the method and compare the approach to goal evaluation presented in this article with other evaluation theories and approaches, such as "goal-focused" evaluation (Peled & Spiro, 1998), "goal-free" evaluation (Scriven, 1991b), "responsive" evaluation (Abma & Stake, 2001), "program theory" evaluation (Chen, 1990; Renger & Titcomb, 2002), and others.

The Why Question: From Identity Conflict to Program Evaluation

The use of the why question for evaluating program goals originated in the theory and practice of resolving "identity" conflicts (Rothman, 1992, 1997), such as the struggle between Israel and the Palestinians or the tensions between the Black community and the police in Cincinnati that erupted in civil unrest in 2001 (Rothman & Land, 2004). Patton (1997) alluded to this kind of conflict when he wrote that "[goal] wars involve highly emotional issues and deeply held values, such as conflicting issues on abortion or sex education" (p. 151). Identity conflicts may manifest themselves as competition over resources, interests, or goals, but they are rooted in people's individual and collective purposes, senses of meaning, and definitions of self. They are particularly intractable because they involve threats to, or the frustration of, fundamental human needs, such as dignity, recognition, safety, control, purpose, and efficacy (Azar, 1990; Burton, 1990).

Dealing with identity conflict requires interventions that go deeper than a search for settlements through zero-sum, power-politics models of negotiation (Banks, 1984) or even through "principle-based" bargaining approaches that search for cooperative solutions (e.g., Fisher & Ury, 1981; Tjosvold, 1991). When parties to such a conflict are at an impasse, one way to move ahead is by asking people on each side why they hold so tenaciously to their positions and why they feel so passionate about them (Rothman, 1992, 1997; Rothman & Friedman, 2002). Each party to a conflict is asked to tell its "story" in the presence of its adversaries. With the probing of a skilled mediator, all parties to a conflict reflect on and clarify the needs and values that are driving them. Conducting this inquiry process in an open and structured way enables each side to hear the other side, often for the first time. This approach enables each side to understand, though not necessarily agree with, the other side's viewpoint.

The importance of this process lies in the fact that an individual's or group's sense of self is often forged in opposition to others: Being an Israeli means that I am not a Palestinian Arab, being a woman means that I am not a man, being a worker means that I am not a manager, and so on. The why question make this process of joint identity formation explicit and positive. It acknowledges that I am who I am, in part, because of who you are, and vice versa. Most important, it creates conditions under which we can jointly and consciously define who "we" are and what kind of relationship we want to have. Under the right circumstances, engaging identity conflict actually offers opportunities for mutual growth, adaptation, and learning as well as a foundation for reaching viable and sustainable solutions (Lederach 1995; Mayer, 2004; Rothman, 1997; Rupesinghe, 1995).

The application of identity conflict engagement to program evaluation began when a foundation officer asked one the authors, Jay Rothman, whether he could clearly define successful conflict resolution in a way such that his board could better evaluate its programs. To answer this question, Rothman and Ross (2000) undertook an empirical study of definitions and measures of success in a dozen conflict resolution initiatives throughout the world. They found that in practice, there were no common definitions of success across all conflict resolution programs. Rather, definitions were varied, local, and contingent on the nature of the conflict and the desires of the stakeholders. In addition, they found that different stakeholders in the same conflict resolution programs—such as participants, funders, professionals, administrators, and evaluators—often defined success differently. If parties to a conflict and other program stakeholders hold different implicit definitions for success, it is unlikely that any outcome will satisfy them all. Without goals that reflect genuine agreement about what stakeholders really want, it will be very difficult to conduct meaningful evaluations.

Logic dictates that stakeholders make their definitions of success explicit and resolve potential conflicts at the outset of a program and reassess them throughout the process of program implementation. However, this logic brings us back to Patton's (1997) goal paradox. The desire to get programs off the ground creates strong incentives to avoid potential minefields by ignoring, circumventing, or smoothing over differences among stakeholders (Weiss, 1993). Wholey, Scanlon, Duffy, Fukumoto, and Vogt (1971) found that federal programs typically had multiple objectives and that interpretations of both ends and means varied widely among implementers. During implementation, these differences often surface as tensions and conflicts among stakeholders (Friedman, 2001b). These conflicts are often difficult to resolve or even to discuss openly precisely because of the avoidance strategies that prevented them from being dealt with in the first place.

For example, in a formative evaluation of a major intervention program for helping teachers work more effectively with "at-risk" students, the evaluators discovered conflicting views among the major program sponsors (a government ministry, a school system, and a nongovernmental

organization) about what the program should actually accomplish (Cohen, Friedman, & Eran, 1996). The official, written program goals were stated at a level of generality that sidestepped these conflicts, but the professionals delivering the program and the teachers kept getting different messages about what was expected from them. These multiple and conflicting messages created a great deal of anxiety among the teachers in particular, who felt that someone would inevitably blame them for failing to deliver. Although the evaluators were aware of these goal differences, they were reluctant to bring these findings into open discussion lest they ignite a conflict in which they would find themselves in the middle.

The research by Rothman and Ross (2000) indicated that effective programs and effective evaluation require a process for forging common goals that engage different stakeholders' definitions of success. This process must not only clarify the form of program goals (i.e., specific, clear, and measurable) but evaluate their content as well. To develop such a process, we began to experiment with systematically applying principles of identity-conflict resolution to program goal setting.

This approach is not based on the assumption that all programs involve intractable conflict but rather that in many, if not most, programs, differences among stakeholders require forging some degree of common identity. Program goal setting as identity formation involves asking three fundamental questions:

1. What are your definitions of success, or goals, for this program?
2. Why are these goals important to you?
3. How do you think the program should go about achieving these goals?

These questions are asked on the individual level ("Who am I?"), the stakeholder group level ("Who are we?"), and the program or intergroup level ("Who are we as a coalition of groups?"). Rather than minimize or obscure differences, this approach aims to make differences at all levels as visible as possible so they can be seriously engaged and resolved before a program actually takes shape.

The main innovation in this approach is the focus on the why question: Why are these goals important to you? Why do you care passionately about them? In our opinion, the why question is the essence of goal evaluation. It treats goals as an object of inquiry. In the goal-setting process, we ask the why question even before we talk about the "what" (i.e., the goals themselves).

When we ask an individual why she or he feels passionate about his or her goals, we are not asking for explanations, justifications, or rationalizations. Rather, we want the individual to reflect on and articulate the underlying values, experiences, and contexts in which these goals are rooted. The aim of the why question is to enable different stakeholders to authentically express themselves so that they can discover deep commonalities and appreciate differences. This process creates conditions under which stakeholders more easily reach agreement on goals to which they are truly committed. It also enables them to question their own goals, reframe them, and discover new ones.

As Patton (1997) framed it, the goal paradox is rooted in the tension between rationality and values. Thus, goal clarification is a bit like standing on the edge of an abyss of "irrationality." We advocate leaping into the abyss of the goal paradox by encouraging stakeholders not only to make their goals explicit but also to give full voice to the passions underlying them. On the basis of his experience with mediating identity conflicts, Rothman hypothesized that giving voice to these passions in a controlled, structured way would create openness for consensus building in program goal setting and deepen stakeholders' internal commitment to the goals themselves. It is important to note, however, the difference between the why question in conflict resolution and its use in goal setting. The former inquires into needs that are

threatened or violated, whereas the latter inquires into needs that drive commitment. When groups are experiencing identity conflict, they need to look to the past before they can collaborate on the future. The method described in this article is not intended for general conflict resolution and is not recommended for groups that are in the midst of conflict over the very existence of a program. The theoretical underpinnings behind this approach are discussed later in this article. In the next two sections, we turn to a case illustration of this kind of inquiry and guidelines for putting it into practice.

The Power of Why: A Case Illustration

The formal process we have developed for putting these ideas into action has been called "action evaluation" (AE) and, more recently, C-3. AE draws on "action science," a set of concepts and strategies for combining reflection and inquiry with action (Argyris et al., 1985, p. 65; Friedman, 2001a, 2001b). AE involves conducting this kind of inquiry at all stages of program design and implementation. In the case of goal setting, it means systematically inquiring into program goals that people choose during the goal-setting process itself. Prior to a face-to-face goal-setting meeting, individual stakeholders are asked to respond to the what, why, and how questions in writing, either through a Web-based data collection and analysis system (http://www.ariac3.com) or through a written questionnaire. These individual responses become the database for forging a common, agreed-on set of goals among members of each separate stakeholder group in a face-to-face meeting. Finally, the goals of all the stakeholder groups become the databases for reaching consensus on a set of common goals that give expression to what various stakeholders really want.

C-3 focuses exclusively on generating or reflecting on program goals that give expression to stakeholder identity. The name C-3 is a play on words intended to emphasize that program stakeholders should "see" programs goals through three lenses of identity: individual, group, and intergroup. Goal evaluations using AE or C-3 have been conducted in over 50 programs in 7 countries.[1] AE or C-3 has been used in programs ranging in size from projects with a few participants in single stakeholder groups to the Cincinnati Police-Community Cooperative, which involved 3,500 participants in eight stakeholder groups (Rothman & Land, 2004).

The following "why discussion" took place as part of an AE process in an experimental high school that attempts to bring different streams of Judaism—orthodox, conservative, reform, secular—together under one roof. This case was chosen because it was the instance in which one of the authors first experienced what we call the "power of why." It is based on field notes hand recorded by the evaluators.

The school had been awarded a grant to develop and implement an approach to teaching Jewish text (the Bible, the Talmud) that gives concrete expression to pluralism. The school set up a project team consisting of teachers and advisers from a local university, but this team failed to make any progress for an entire year. The problem, at least in part, was that the teachers held powerful and seemingly incompatible views about the meaning of these texts and how they should be taught. As long as they taught separately, direct conflict over these differences could be avoided. By bringing these issues to the fore, the project threatened this avoidance strategy.

To get the project off the ground, the school's principal and the team leader decided to try AE. The first step was for the team leader to identify the main stakeholders in this project. He identified two groups: faculty members and students. This illustration deals with the faculty group, which consisted of nine teachers and two university advisors. After receiving an explanation of the process by the team leader, each member of the faculty group filled out a brief questionnaire consisting of three open-ended questions: "What are your goals for this project?

(i.e., How would you define success?),)," "Why are these goals important to you?" and "How would you go about achieving these goals? (What should be done? What concrete steps would you suggest be taken?)." Each participant was asked to name at least three definitions of success (goals). Their responses to these questions were entered directly into the AE database on the Web.

The evaluators analyzed the 20 goals that were entered by the 11 faculty members and grouped them into categories according to similarity of content. This analysis yielded 3 goals that were shared by at least 2 of the participants and 7 "unique" goals that were mentioned only by 1 participant each. The analysis also identified 2 potential goal conflicts or dilemmas.

The results of the initial analysis were presented at a meeting of all the members of the faculty group. The purpose of this meeting was to generate a common statement of project goals. It was divided into two distinct parts, each lasting approximately 2 hours. The first part was a discussion that focused only on the why question. The second part was a presentation of the goals, as analyzed by the evaluators, followed by a decision-making process to craft a set of final program goals.

During the first part, the teachers were seated in a circle. There were two evaluators, one to facilitate the discussion and the other to record what was said. There was considerable tension in the air. After the evaluators introduced themselves, one of the teachers expressed doubts as to whether goal setting was the right focus and asked whether the whole exercise wasn't just a "waste of time." The evaluator acknowledged that it might in fact turn out to be a waste of time but suggested that the only way to test that claim was to engage in the process itself.

The evaluators handed each participant a printout of his or her individual answer to the three questions. He then asked them to temporarily leave the "what" question (their goals) aside and to focus on the why question. He asked each one to state why the goals he or she had defined were important to him or her: why he or she felt passionately about them.

The discussion began slowly. One participant said, "I was instrumental in setting up the Bible program and I want to see how it works." The evaluator probed into this response, asking the teacher why it mattered so much to her, and she responded, "It's my job." Then the team leader spoke up:

> *Project leader:* I was in the same position and thought it would be good for me to see how my ideas are working out. I am curious to know how it is developing. I am highly invested in this project—being in charge of something like this was intellectually stressing. And I like being in charge. I'd rather rule than be ruled. Also I am a bit weird and it is an opportunity for me to bring my weirdness into the curriculum.
>
> *Evaluator:* In what way are you "weird"?
>
> *Project leader:* I believe that language skills are the most important aspect, and it is important to invest time in learning the language of the text. In that way they can learn to hear the voice of the text. We don't need to focus on what the text says. The students will hear that for themselves if they know the language and how to read the text.

The teacher who had suggested that the process would be a waste of time read aloud the why statement he had entered in the database:

> *Teacher:* It is more important to discuss our beliefs, values, and motives regarding the material we teach than to simply decide which texts to teach. We are role models and must walk in the path of holiness so we can bring these teachings to the school community.
>
> *Project leader:* That was provocative! Read it again. [The teacher reads it again.]
>
> *Evaluator:* [To teacher] But why is that so important to you?
>
> *Teacher:* [With strong emotion] So the students will get the best Jewish education possible. I have a Messianic view of this work. I want the world to be a better place . . . that's why I am a teacher!

At that point, there was a change in the atmosphere in the room. The participants began leaning toward one another, anxious to hear what the others were going to say. One by one, they talked about why, expressing the values and motivations that made the success of this project important to them. The evaluator no longer probed very much, but he occasionally intervened to keep the focus on why rather than what. The other participants listened and also asked questions. Once the discussion gained momentum, the evaluator's main role was simply to ensure that each speaker had ample time and space to have his or her full say without being interrupted. Occasionally, the evaluator intervened to steer the discussion away from debate, pointing out that the aim of this discussion was not agreement but rather understanding.

According to the project leader, these teachers had never actually talked with one another in this way before, even though most of them had been working together in the same school for 2 years. They had rarely found the space to speak publicly and fully about what underlies their teaching and their worldviews without having to be ready to defend themselves. They had few opportunities to hear one another express himself or herself so personally, deeply, and passionately.

This dialogue sharpened differences among the stakeholders but also revealed deep commonalities, such as their commitments to Jewish learning, to the students, and to making the school work. In their questions and their listening, the teachers expressed an appreciation for one another, even for the differences they had among themselves. This discussion went on for almost 2 hours, until all the participants had an opportunity to talk about their whys. No one was forced to share his or her why, and 1 participant chose to remain silent.

After a short break, the group moved to another room for the analysis of their whats (definitions of success). Each participant was given a printout with the full analysis and a summary of shared, unique, and potential conflicts and dilemmas was taped to the wall on flip chart paper. The facilitator first asked whether all of their goals were included in the printed analysis and whether the grouping of goals made sense. The group responded by looking critically at the goals, each of which was hotly debated. Nevertheless, participants also inquired, considered, and looked for common ground. The evaluator pushed the group and kept them focused on the task. He encouraged consensus but also tried to prevent compromises for the sake of convenience. Some goals were discarded or reworded.

After 2 hours, the group reached consensus on four goals for the project. The team leader and the participants expressed surprise at how little time it took them to reach agreement on goals to which they were genuinely committed. The teacher who wondered whether it would be a waste of time also expressed surprise and satisfaction with the outcome. The positive effect of this goal-setting process continued as the project developed over the next year.

Guidelines for Facilitating a Why Dialogue

One of the reasons that we have been so impressed by the power of why is that we have each independently experienced (a) how this question can stimulate a discussion that is similar to the one described above and (b) how this discussion appears to help stakeholders reach consensus without smoothing over their differences. We believe that a why discussion is a useful way of beginning a goal-setting process at the outset of a program or for reevaluating program goals as a part of a formative evaluation. It can be applied to projects of practically any size, though the amount of time and effort increases with the number of stakeholder groups. The greater the uncertainty, ambiguity, or differences about program goals, the more important a systematic why (and what) discussion is. Finally, this approach is appropriate to the extent that program organizers or formal leaders are open to a participative, stakeholder

approach. The degree of participation may vary. For example, in some school projects students have been involved in the face-to-face goal setting, whereas in other schools only their written goals were taken into account.

When properly structured and facilitated, the why discussion provides an efficient and relatively nonthreatening way of preparing participants for a more productive goal-setting process (the what discussion). The AE or C-3 process also includes very specific guidelines for the what discussion, but in this article we focus only on the why. Information on the full AE or C-3 process can be found elsewhere (e.g., Friedman & Rothman, 2001; Rothman and Friedman, 2002). Although we strongly advocate following the AE or C-3 process, the why question can be integrated into other goal-setting methods.

In general, the starting assumption in applying the why question is that stakeholders see one another as potential partners and want to strengthen cooperation through an open engagement of differences. A why dialogue as described in this article is not recommended when partners see one another as adversaries and when there are feelings of deep enmity and mistrust. Such situations probably call for a more systematic process of conflict resolution. Unless evaluators have the mandate and the expertise to mediate a serious conflict, they are advised to refer the stakeholders elsewhere until they are ready to seek cooperation.

The challenge in facilitating a why dialogue is moving the discourse from explanations to passions. In the case example, the first teacher responded to the why question and to the facilitator's probing with explanations: "I was instrumental in setting up the Bible program and I want to see how it works. . . . It's my job." Participants' written responses to the why question on the questionnaire are usually of an explanatory nature; they rarely write deep or passionate statements of why they care about their goals. Nevertheless, the questionnaire does open up an avenue of thought that can be deepened during the subsequent why discussion.

There is no single, surefire method for facilitating a why dialogue. Each new situation presents a unique context and set of challenges. Furthermore, each of us facilitates these discussions somewhat differently, in tune with our different personal styles. Nevertheless, on the basis of our collective experience so far, we have identified a number of guidelines that we have found helpful for facilitating a productive why dialogue:

1. Collect what, why, and how data from each individual in advance.
2. Give participants their individual data and a few minutes to silently review them and to think more deeply about their why responses. We have found it useful to ask participants to think of one word (a "passion point" or keyword) that summarizes one or more of their why responses and to think of short stories or anecdotes that illustrate their whys.
3. Invite participants to share their why responses, passion points, and/or stories. Make sure all participants understand that the time provided for each participant (7 to 10 minutes) is "protected time." Each speaker is given the uninterrupted time and space to give voice to his or her whys and to feel heard. It is often helpful if the evaluator models choosing a passion point and telling a story that makes it come alive.
4. Give participants a choice. Invite them into this process, but make active participation optional. Honor silence as a legitimate form of participation rather than as a sign of resistance.
5. Inquire more deeply, when necessary, into participants' why statements. The facilitator often needs to gently probe to challenge speakers to go deeper and to help them be understood. Other participants are encouraged to join into this inquiry process, as long as they refrain from criticism or argument. The aim is understanding, not agreement. They can test their understanding of the speaker by offering passion points they infer from what has been said.
6. Record passion points and stories. There are distinct advantages to a scribe (rather than tape recording), because it obviates the need to transcribe and because the data are immediately available. The scribe may actually assist the facilitator by asking for clarifications or for a speaker to slow down.

7. Set clear boundaries for participation. Only stakeholders who have entered data in advance should participate in the why discussion. Participants should commit themselves in advance to coming on time and staying until the end. The process should last no more than 2 hours, which limits the number of participants to about 15.

Collecting data in advance ensures that each individual's voice is included in the goal-setting process and that each stakeholder has an opportunity to think carefully and independently before having to express themselves in a group. It enables us to analyze the what data (the goals) and to present this analysis immediately after the why discussion, making the goal setting significantly more efficient. However, as with any analysis, evaluators risk biasing goals in the way they interpret and organize the data. We believe that this risk is worth the advantages as long as evaluators treat their analysis as a kind of "hypothesis" to be tested with the stakeholders themselves. The evaluator should provide the specific responses from which each goal was derived and invite participants to reinterpret the data for themselves. Evaluators should not defend their interpretations but rather invite criticism and encourage participants to craft the goal statements in ways that express what they really mean.

As illustrated in the case, we conduct the why discussion prior to the discussion of goals and explicitly ask participants not to connect their whys to their whats. Participants are sometimes taken aback by this approach, which seems counterintuitive, but it has significant advantages. First, it frees people from having to express why in defense of what (fixed positions) or in opposition to someone else's what. The fact that there is no need for debate, agreement, or decision making is extremely liberating. It relieves people of the burden of defending themselves or having to persuade others. The very strangeness of why without what often shakes participants out of their resistance to goal setting and gets them thinking in unconventional ways.

This process aims at achieving shared goals but not necessarily shared passions. Clearly, the passions that commit people to program goals may be very different. Passions are individual, and people may hold passionately to the same goals for different reasons. Separating what and why can be compared with the distinction between "discussion" and "dialogue" (Isaacs, 1999). According to Isaacs (1999), dialogue is a process in which people "think together" in "exploring the nature of choice . . . evoking insight . . . reordering knowledge, particularly the taken-for-granted assumptions that people bring to the table" (p. 45). Dialogue improves the quality of talk by "helping to create an atmosphere in which we can perceive what really matters most to us, and to one another" (p. 47). Discussion, on the other hand, is about "making a decision . . . which seeks closure and completion" (p. 45). AE or C-3 creates a structured process for interweaving dialogue and discussion. It creates a separate time and space for addressing the why question. In most cases, this discussion creates conditions for more productive goal setting (what) and action planning (how).

The first few minutes of a why discussion are often uncomfortable. The psychotherapy and counseling literature warns that the why question elicits reason or intellectualizing (Cormier & Cormier, 1991) and can be offensive to a client and create defensiveness (Ivey, Ivey, & Simek-Downing, 1980; Pedersen & Ivey, 1993). And, in fact, people usually begin at the level of explanation and even banalities. Getting to passions requires gentle probing: "Yes, but why do you care so much about that?" or "Why do you feel passionately about the issues at hand?" Once participants realize that they are not being asked for reasons but rather to reflect on deeper motivations, it can also be discomfiting, because they are simply not accustomed to being asked to think about why they hold the values that they hold.

Storytelling and passion points are convenient tools for getting over this discomfort. Furthermore, passion points provide a way of linking whats with whys. After participants define their consensual goals in the what discussion, the evaluator can ask participants to

identify at least one passion point for each goal. This process is a way of testing whether the suggested program goals actually give expression to what the stakeholders said they value. If goals do not stand this test, they ought to be reconsidered. In this way, the why discussion provides a basis for actually evaluating goals and including this information in the database for ongoing evaluation.

The Power of Why

The above guidelines are intended to generate an inquiry process, on the basis of the why question, that creates conditions conducive to meaningful and productive collaborative goal setting. Something happens when this question is asked and pursued seriously in a group of people who come together to create a program for action. Each of us independently and repeatedly experienced how this process leads stakeholders to meet one another in a way that is very unusual and often quite moving. It is the kind of experience that led us to sit down and send e-mail messages to each other that said "Wow! Let me tell you what just happened. . . ."

What accounts for the power of why? Patton (1997) framed the goal paradox in terms of an opposition between human rationality and human values. This framing builds on an understanding of rationality as the optimum means for achieving certain ends. Means are viewed as relatively flexible; they can be critically examined, tested, and changed through logical and empirical deliberation. Patton's framing implies that rational, or at least thoughtful, deliberation stops when it comes to ends; values and goals are treated as givens that lend themselves to strong emotions and power struggles.

Although Patton's framing has face validity, it does not necessarily follow that there cannot be a thoughtful and deliberative discussion about values underlying goals. Argyris and Schön (1996) argued that under conditions of value conflict, people can engage in a kind of inquiry (i.e., "double-loop learning") that "gets underneath the members' initial commitments" and in which participants "ask why they hold the positions they do and what the positions mean" (p. 21). To be effective, however, this kind of inquiry needs to be guided by a set of higher level values: valid information, free and informed choice, and internal commitment to choice and constant monitoring of its implementation (p. 118).

An effective why discussion promotes all three of these higher level values. It generates valid information by making stakeholders' underlying motivations public and observable. Rather than suppressing or advocating what they think, people display their thinking—literally holding it up as if suspended before them—so that they and others can see and understand it (Isaacs, 1999). Suspension means acknowledging one's thoughts and feelings as they arise without being compelled to act on them. Suspension opens people to new possibilities for understanding and action.

The degree of validity of this information, of course, depends on the extent to which participants are open and honest. A rigorous inquiry process, however, provides opportunities for testing one's understanding and improving the quality of the information. In this context, goal inquiry is intended to clarify, make sense of, and understand one's own values and those of others. It does not mean accepting, rejecting, or judging values as right or wrong. This feature of the why dialogue is closely related to "appreciative inquiry" that gets "beyond superficial appearances to deeper levels of the life-generating essentials and potentials of social existence . . . to affirm, and thereby illuminate, the factors and forces involved in organizing that serve to nourish the human spirit" (Cooperrider & Srivastva, 1987, p. 131).

The why dialogue enhances free and informed choice by asking people to step back to consider why they want what they say they want. This process brings people's own needs, values,

and desires into consciousness, where they can be explored, questioned, and subject to choice. Because they can observe and understand the motivations of others, their eventual choice of goals will draw on a wider base of information. We have seen participants rethink, reframe, and discover new goals in the course of listening to themselves and to others talk about their whys. As illustrated in the case example, participants in a why discussion come to appreciate one another's differences in ways that lead them to seek common ground.

Internal commitment is closely related to the issue of passion in the why dialogue. Giving voice to passion is clearly one of the sources of the power of why. Passion is underexplored in the literature, perhaps because it is so out of step with the dominant espoused values of rationality in organizational life (Boverie & Kroth, 2001; Hirschhorn, 2003). Maybe passion is the ultimate stretch into discomfort that all learners need to take. Maybe we dodge passion because it is mistakenly associated with irrationality, a loss of self-control, and the heated emotions that sweep people up and lead to tragic consequences. There is, however, no reason to equate passion with a loss of control or a kind of "let it all out" emotional expression. Max Weber regarded passion as a decisive quality for both politicians (Weber, 1919/1946a, p. 115) and scientists (Weber, 1919/1946b, p. 135). However, he made a distinction between passion "in the sense of *matter-of-factness* of passionate devotion to a 'cause'" and passion as a "sterile excitation . . . devoid of all feeling of objective responsibility" (Weber, 1919/1946a, p. 115). Passions need not be dampened to ensure productive work and relationships. To the contrary, we argue that passion, when linked to responsibility, is essential to good work for both program stakeholders and evaluators alike.

When program goals are linked to the underlying whys (e.g., through the use of passion points), it provides a basis for ongoing monitoring of the relationship between goals and underlying motivations. Thus, a why discussion provides a means for reassessing goals at any point in the life of a program:

> I was working with the medical staff at an inner-city AIDS clinic. These people were under great stress. It was the mid-90's and we were just beginning to see a high number of infected infants. People were sniping at one another. The administrator hired a grief counselor to run some workshops and me come in to "fix" the team. This was a group with a clear, established, and proven set of goals and procedures. They were even achieving much of what they had agreed to. I asked two simple why questions: "Why did you become a nurse?" and "Why on earth did you choose a field of practice where every one of your patients dies?" The responses were powerful. It was an emotional discussion as people piggy-backed on what had been said. In the follow-up, some tried and true "what's" and "how's" ended up being changed and some were re-embraced driven by the resonance of shared why's.

As this vignette illustrates, the question is not simply whether a program is proceeding toward its goals, but whether these are the right goals. In this way, the why question provides a powerful means for evaluating goals themselves and realigning them with what stakeholders really want.

Finally, we believe that the power of why stems at least in part from a fundamental change it fosters in identity; that is, the way one thinks about oneself and about one's relationship with others. The essence of the why discussion is articulating and sharing one's passionate devotion to a cause in the presence of others. Needless to say, this kind of meeting between people is unusual. It is intended to set into motion a process of "reflexive" dialogue in which the experience of the other touches and resonates with something in ourselves. As a result, both sides develop a kind of "analytical empathy" (Rothman, 1997, p. 45) in which they discover similarities and commonalities at the level of their deeper, existential needs.

Martin Buber (1937/1970) claimed that an individual's most basic "I" (needs, values, desires) is articulated through an encounter with a "thou." Buber (1966) suggested that in such encounters, people "respond" rather than "react" to one another:

> In our life and experience we are addressed; by thought and speech and action, by producing and by influencing we are able to answer. For the most part we do not listen to the address, or we break into it with chatter. But if the word comes to us and the answer proceeds from us then human life exists, though brokenly, in the world. The kindling of the response in that "spark" of the soul, the blazing up of the response, which occurs time and again, to the unexpectedly approaching speech, we term responsibility. (p. 19)

In why dialogues, we have frequently observed what we believe is this kind of response and the emergent feelings of responsibility. It is not a feeling of in-group intimacy but rather the experience of coming together to accomplish purposes valued by the stakeholders in a wider environment (Hirschhorn, 1990). Paradoxically, the why discussion, by making these values very personal, enables people to move beyond the interpersonal, or even the political, and focus on their joint tasks and purposes.

There are, however, risks involved in bringing differences to the surface. One or more stakeholder groups might discover that there are not sufficient grounds for collaborating on a project and end its participation. We would not necessarily see such an outcome as a "failure" of the process as long as stakeholders made their decisions on the basis of free and informed choice. Such an outcome would, in our opinion, be preferable to achieving false consensus in which sides cover up or smooth over differences, because this strategy would very likely lead to destructive conflict and even less trust later during implementation.

Goal Evaluation in Evaluation Theory and Practice

As pointed out earlier, evaluation theory has focused mainly on improving the form (specific, clear, measurable, useable) rather than the content of goals (Patton, 1997). Evaluability assessment, in particular, aims at determining whether program goals are well formulated and realistic given the available resources (Smith, 1989; Wholey, 1994). However, there is no a priori reason that the content of program goals should not be an object of evaluation. Scriven's (1991a) definition of evaluation as the practice of determining the worth, value, and merit of something can be applied to ends as well as to means. There is no reason why evaluators should refrain from helping stakeholders become aware that they have made implicit choices about interests and goals that might otherwise be regarded as givens. This process of goal evaluation is especially important when programs involve different and potentially conflicting goals. The more stakeholders discover these implicit choices, the more they create potential leverage points for generating desired outcomes for all sides.

As pointed out earlier, one main criterion for assessing the "worth, value, or merit" of a particular goal is passion: the extent to which stakeholders are internally committed to it. What happens if stakeholders do not feel passionate about the project goals they have defined? On this issue, we take a frankly normative stance that reflects Max Weber's (1919/1946b) statement that "nothing is worthy of man as man unless he can pursue it with passionate devotion" (p. 135). Good program goals are ones that stakeholders feel passionate about. People should not invest their time, talent, money, or authority into programs they do not care that much about. Passion, as Weber pointed out, is not simply the expression of strong emotion, and it need not be dampened to ensure productive work and relationships. Passion can be rational,

in the sense that it helps people carry out difficult, mundane, or sometimes even distasteful action steps because, as Weber pointed out, it is linked to responsibility.

Of course, this standard applies only when stakeholders come to the table with some positive investment in a program even if they might hold divergent program goals. If people are stakeholders in a program with which they disagree in a fundamental way, it most likely is because they have not been involved in its design or have been involved only nominally, withholding genuine participation. If the why discussion is carried out after a program has begun and some of the stakeholders have felt uninvolved or coerced, it is a conflict engagement scenario for which the method we have described is not appropriate. In addition, one danger of the why discussion is that it may force an expression of passion that is not really there. Another danger is that it might give unfair advantage to people who are especially articulate or disadvantage people who have passions but are not skilled or comfortable at expressing them.

Our analysis of the power of why suggests two additional criteria: (a) the extent to which program goals accurately reflect what stakeholders define as success and (b) the extent to which consensual goals are based on free and informed choice. It is important to note that these criteria are by no means morally or culturally neutral. They reflect values we hold for human dignity, participation, and learning. These criteria and the procedure for goal evaluation described above correspond closely to three requirements for what House and Howe (2000) called "deliberative democracy evaluation": the inclusion of all relevant interests, dialogic interaction that enables stakeholders to discover their real interests, and a deliberative process that subjects values and ends to rational examination.

Despite Patton's (1997, pp. 167-169) criticism of the goals-clarification game, his 10 "guidelines for working with intended users to identify meaningful and useful goals" all relate to the clarification and crafting of goal statements. They are very reasonable guidelines for creating goals that are outcome oriented, specific, and focused, but they treat the content of goals as givens. None of these guidelines involve asking program staff members why they hold these particular goals.

Patton (1997) did suggest that one way out of the paradox might be to circumvent it through goal-free evaluation (Scriven, 1991b). Goal-free evaluation argues that evaluation should focus on actual program outcomes rather than intended goals. In fact, it regards intended goals as being of little practical use to evaluators because they are sources of perceptual bias, blindness to actual outcomes, and threats to objectivity. Indeed, Patton's critique of the goals-clarification game is that evaluators use it to covertly impose goals they already have in mind onto program staff members. This criticism is valid and certainly a danger in our approach to goal inquiry, which might be criticized for involving an evaluator too deeply in program goal setting. However, we argue that the way to avoid bias is to make evaluators' role explicit and open to scrutiny and criticism by stakeholders.

Goal-free evaluation also raises the question of whether our approach to goal setting and the why discussion focuses too much on stakeholders' wants, needs, and desires over those of program beneficiaries. One way of dealing with this kind of bias is to ensure that beneficiaries are themselves stakeholders in the process, which is usually the case in AE or C-3. Even so, it is important to make sure that the voices of program recipients do not get drowned out by those of other stakeholders. In some programs, such as the Cincinnati Police-Community Cooperative, in which there were eight distinct stakeholder groups, all of whom were potential beneficiaries, systematic joint inquiry into goals might be the only way to ensure that beneficiaries' needs are addressed by the programs.

Peled and Spiro (1998) advocated goal-focused evaluation, which addresses gaps between a program's declared goals and the operative goals implicit in its implementation. They argued that a comprehensive evaluation of a program's effectiveness should identify these

gaps and allow stakeholders to refocus program goals and channel resources accordingly. Our approach is consistent with goal-focused evaluation because both evaluate the content of goals and advocate the periodic calibration of declared and operative goals. Unlike Peled and Shapiro, however, we argue that goals should be evaluated during program design rather than waiting to discover goals that are implicit in program implementation. Furthermore, in formative evaluation, we would not necessarily take operative goals as givens but subject to inquiry guided by the why question.

On the conceptual level, the "theory-driven evaluation" approach to goal evaluation is consistent with, but more extensive than ours (Chen, 1990). It focuses extensively on "normative outcome evaluation," which is defined as an attempt to assist stakeholders in identifying, clarifying, or developing the goals or outcomes of program for program improvement. Chen (1990) addressed goal setting from both the political and substantive perspectives. He also distinguished between the "desirability aspect" (what stakeholders want) and the "plausibility aspect" (what social science research suggests) of program goals. Chen advocated making goals an object of inquiry through "goal revelation evaluation" aimed at "examining and uncovering the major goal dimensions underlying stakeholders' beliefs and views" (p. 92).

Both approaches use data collection and analysis in the goal-setting process, but there is a significant methodological difference. Chen (1990) advocated circumventing direct confrontation through the use of statistical tools and nominal group methods for constructing stakeholders' normative theory of goals. We, on the other hand, advocate face-to-face engagement of differences through why and what discussions.

Logic modeling emphasizes the clarification and plausibility of goals but usually aims the why question at understanding the causal links between certain program interventions and program outcomes (e.g., Brinkerhoff & Apking, 2001; McLaughlin & Jordan, 1990; Renger & Titcomb, 2002; Turnbull, 2002). Evaluability assessment studies often focus on desirability through extensive stakeholder consultation in the building and testing of logic models (e.g., Dwyer et al., 2003). These approaches attempt to identify and prioritize stakeholders' goals, but they rarely inquire into why stakeholders have chosen these ends.

Our approach to goal evaluation, on the other hand, focuses mainly on desirability, which could lead to the danger of setting unrealistic goals. Therefore, this approach to goal evaluation needs to be balanced against the evaluability of program goals. Passionate commitment, for example, does not justify setting program goals that are clearly unrealistic given the available resources (Smith, 1989; Wholey, 1994). One way of dealing with this problem is to ensure that experts are included as stakeholders in the process. Indeed, the process is designed to generate productive dialogue for working through potential conflicts between experts and beneficiaries. Another strategy is for stakeholders to invite experts to review goals that emerge from the deliberative process before making final decisions about implementation. Nevertheless, our approach values free choice over expert knowledge as long as stakeholders are well informed and understand the potential implications of their choices. This approach makes particular sense when expert knowledge falls short and programs deal with problems characterized by uniqueness, uncertainty, and instability (Friedman, 2001a).

Abma and Stake's (2001) responsive evaluation shares our emphasis on making value differences among stakeholders an object of inquiry. Another commonality is framing the evaluator's role as "discovery learning teacher" and "facilitator" (p. 14). Responsive evaluation explicitly refrains from pushing for consensus where none exists. Our approach does not advocate pushing for false consensus, but it does consider consensus to be a criterion for a good goal. Thus, we put effort into inquiry and a search for common ground before coming to the conclusion that no possibility for consensus exists.

The degree of consensus also distinguishes our approach from empowerment evaluation (Fetterman, 1994, 1998; Fetterman & Bowen, 2003). Both approaches advocate an open, democratic, and creative process that allows stakeholders an opportunity to voice their visions of a program. However, when members of a group disagree on a goal or mission statement, empowerment evaluation seeks to help them reach an agreement they can "live with" but to which they may not be fully committed (Fetterman & Bowen, 2003). This pragmatic approach also reflects concern that goal setting can be a slow process that is perceived as a burden by busy stakeholders. We believe that this concern is quite valid and try to structure the process in a way that is both efficient and oriented toward achieving consensus. Although empowerment evaluation takes motivation into account in the choice of goals, it has not yet included the why question as an explicit part of its goal-setting approach.

Finally, our approach to goals enters into the discussion of the role of dialogue in evaluation theory and practice. Ryan and DeStefano (2000) offered a model that yields four "genres" of dialogue: conversation (inclusive-divergent), inquiry (inclusive-convergent), instruction (critical-convergent), and debate (critical-divergent). The why discussion does not fit neatly into any of these categories. It is clearly inclusive but also critical in the sense that it does not take goals for granted. It aims at neither convergence nor divergence but rather suspension and appreciation. On the other hand, conversation, inquiry, and instruction all manifest themselves in the goal-setting process (the what discussion) that follows the why dialogue.

There may very well be programs for which this approach to goal evaluation is not appropriate because the goals are simply too mundane and obvious to warrant inquiry or passion. However, it is difficult to make this determination in advance. Program goals that seem very straightforward and obvious to one stakeholder might not seem so straightforward and obvious to others. In addition, some people get passionate about what seems to others mundane. The question for evaluators is whether for any given program there is a real need for goal evaluation and whether the conditions exist for a productive process. The why discussion, as an approach to goal evaluation, is most essential when (a) there are multiple stakeholders; (b) stakeholders are likely to have very different viewpoints, values, and so on; (c) definitions of success in the subject area are uncertain, ambiguous, or contested; (d) stakeholders have free choice about participation in the program; and (e) the program initiators want a participatory approach to program design. On the other hand, it is least likely to be productive when (a) there is only a single stakeholder, (b) goals are mandated and nonnegotiable, (c) stakeholders have no choice about participation, and (d) program initiators oppose participation in program design. There are, however, many possibilities between these two extremes. For example, a why discussion can be very useful even when there is a single stakeholder if only to help that stakeholder clarify his or her level of commitment.

There are also costs to the why discussion and to goal setting as conducted in AE or C-3. It requires stakeholders to devote at least 4 hours to this process. Members of stakeholder groups who cannot participate in the process may feel left out. The AE or C-3 method can be carried out with very large groups and complex programs with multiple stakeholders; the Cincinnati Police-Community Cooperative involved over 3,500 citizens in eight stakeholder groups. Because the why discussion needs to be carried out in relatively small groups, this large a program requires parallel discussions. There is no need to merge the whys, but reaching agreement on goals can be carried out only through representatives. Naturally, this process demands considerable time and resources, but the costs of carrying out this systematic process need to be weighed not just against benefits but also against the cost of not doing so.

Given the intensely political nature of program environments and competing interests among stakeholders, the strongest claim against this approach may simply be that it is politically

naive. Although it is pure folly to ignore politics, we believe that politics too can be a self-fulfilling prophecy. Inconsistencies and maneuvering are not necessarily indications that people are consciously hypocritical. People do not regularly work to align values with actions or to minimize the inconsistency between words and deeds. In many cases, they are simply not conscious of what they really mean and care about. Making this kind of alignment an explicit, intentional, integral part of evaluation creates conditions under which program stakeholders are most likely to learn from the process (Preskill & Torres, 1999; Preskill, Zuckerman, & Matthews, 2003).

Evaluators do not necessarily have to live within a goal paradox. Systematic, structured inquiry into the why question provides a means for engaging differences among stakeholders in a way that increases the likelihood of mutual understanding, appreciation, and consensus. This kind of inquiry integrates the evaluation of goal content with the goal-setting process in a way that aims at defining and improving goals at the same time. It is consistent with and can be adapted to many existing evaluation methods to generate program goals that more accurately reflect the outcomes stakeholders really want and to which they are truly committed. The power of why lies in its potential for harnessing the passion inherent in values to the usefulness of goal setting for directing action.

Note

1. It is beyond the scope of this article to describe all of these projects. However, a brief description of all projects, including names, numbers of participants, stakeholder groups, and outputs will be supplied by any of the authors on request.

References

Abma, T., & Stake, R. (2001, Winter). Stake's responsive evaluation: Core ideas and evolution. *New Directions for Evaluation*, *92*, 7-22.

Argryis, C., Putnam, R., & Smith, D. (1985). *Action science: Concepts, methods, and skills for research and intervention*. San Francisco: Jossey-Bass.

Argyris, C., & Schön, D. A. (1996). *Organizational learning II: Theory, method, and practice*. Reading, MA: Addison-Wesley.

Azar, E. (1990). *The management of protracted social conflict: Theory and cases*. Aldershot, UK: Dartmouth.

Banks, M. (Ed.). (1984). *Conflict in world society: A new perspective on international relations*. Sussex, UK: Wheatsheaf.

Boverie, P. E., & Kroth, M. (2001). *Transforming work: The five keys to achieving trust, commitment, and passion in the workplace*. Cambridge, MA: Perseus.

Brinkerhoff, R. O., & Apking, A. M. (2001). *High impact learning: Strategies for leveraging business results from training*. Cambridge, MA: Perseus.

Buber, M. (1966). *The way of response*. New York: Schocken.

Buber, M. (1970). *I and thou* (W. Kaufman, Trans.) New York: Scribner. (Original work published 1937)

Burton, J. (Ed.). (1990). *Conflict: Human needs theory*. New York: St. Martin's.

Chen, H. T. (1990). Theory-driven evaluations. Newbury Park, CA: Sage.

Cohen, M., Friedman, V., & Eran, M. (1996). *Evaluation of the "New education environment in technological schools of the Amal Network": First interim report of school characteristics and the initial stages of program implementation* [in Hebrew]. Jerusalem, Israel: Brookdale Institute.

Cooperrider, D., & Srivastva, S. (1987). Appreciative inquiry in organizational life. In W. Passmore & R. W. Woodman (Eds.), *Research in organizational change and development* (Vol. 1, pp. 129-169). Greenwich, CT: JAI.

Cormier, W. H., & Cormier, L. S. (1991). *Interviewing strategies for helpers: Fundamental skills and cognitive behavioral interventions*. Pacific Grove, CA: Brooks/Cole.

Dwyer, J. J. M., Hansen, B., Barrera, M., Allison, K., Ceolin-Celestini, S., Koenig, D., et al. (2003). Maximizing physical ability: An evaluability assessment to plan a community-based, multi-strategy approach in an ethno-racially and socio-economically diverse city. *Health Promotion International*, *18*(3), 199-208.

Fetterman, D. (1994). Empowerment evaluation. *Evaluation Practice*, *15*(1), 1-15.

Fetterman, D. (1998). *Empowerment evaluation: Collaboration, action research, and a case example*. Available at http://www.aepro.org/inprint/conference/fetterman.html

Fetterman, D., & Bowen, C. (2003). Experiential education and empowerment evaluation: Mars Rover Educational Program case example. *Journal of Experiential Education*, 25(2), 286-295.

Fisher, R., & Ury, W. (1981). *Getting to yes: Negotiating agreement without giving in*. Boston: Houghton Mifflin.

Friedman, V. (2001a). Action science: Creating communities of inquiry in communities of practice. In P. Reason & H. Bradbury (Eds.), *The handbook of action research* (pp. 159-170). London: Sage.

Friedman, V. (2001b). Designed blindness: An action science approach to program theory evaluation. *American Journal of Evaluation*, *22*, 161-181.

Friedman, V., & Rothman, J. (2001). Action evaluation for knowledge production in social-educational programs. In S. Shankaran, B. Dick, R. Passfield, & P. Swepson (Eds.), *Effective change management through action research and action learning: Frameworks, processes and applications* (pp. 57-65). Lismore, Australia: Southern Cross University.

Hirschhorn, L. (1990). *The workplace within: Psychodynamics of organizational life*. Cambridge, MA: MIT Press.

Hirschhorn, L. (2003). *Passion and group life*. Philadelphia: Center for Applied Research.

House, E., & Howe, K. (2000). Deliberative democracy evaluation. *New Directions in Evaluation*, *85*, 3-12.

Isaacs, W. (1999). *Dialogue and the art of thinking together*. New York: Doubleday Currency.

Ivey, A., Ivey, M., & Simek-Downing, L. (1980). *Counseling and psychotherapy: Skills, theory, and practice*. Englewood Cliffs, NJ: Prentice Hall.

Lederach, J. P. (1995). *Preparing for peace: Conflict transformation across cultures*. Syracuse, NY: Syracuse University Press.

Mayer, B. (2004). *Beyond neutrality: Confronting the crisis in conflict resolution*. San Francisco: Jossey-Bass.

McLaughlin, J. A., & Jordan, G. B. (1990). Logic models: A tool for telling your program's performance story. *Evaluation and Program Planning*, *22*, 65-72.

Morris, L., & Fitz-Gibbon, C. (1978). *How to deal with goals and objectives*. Beverly Hills, CA: Sage.

Patton, M. (1997). *Utilization-focused evaluation* (3rd ed.). Thousand Oaks, CA: Sage.

Pedersen, P., & Ivey, A. (1993). *Culture-centered counseling and interviewing skills: A practical guide*. Westport, CT: Praeger.

Peled, E., & Spiro, S. (1998). Goal focused evaluation: Lessons from a study of a shelter for homeless youth. *Evaluation*, *4*(4), 455-468.

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

Preskill, H., Zuckerman, B., & Matthews, B. (2003). An exploratory study of process use: Findings and implications for future research. *American Journal of Evaluation*, *24*, 423-442.

Renger, R., & Titcomb, A. (2002). A three-step approach to teaching logic models. *American Journal of Evaluation*, *23*, 493-503.

Rossi, P. (1972). Testing for success and failure in social action. In P. Rossi & W. Williams (Eds.), *Evaluating social programs* (pp. 11-65). New York: Seminar.

Rothman, J. (1992). *From confrontation to cooperation: Resolving ethnic and regional conflict*. Newbury Park, CA: Sage.

Rothman, J. (1997). *Resolving identity-based conflict: In nations, organizations and communities*. San Francisco: Jossey-Bass.

Rothman, J., & Friedman, V. (2002). Action evaluation for conflict management in organizations and projects. In J. Davies (Ed.), *Second track diplomacy: Concepts and techniques for conflict transformation* (pp. 285-298), Boulder, CO: Rowman & Littlefield.

Rothman, J., & Land, R. (2004). The Cincinnati Police-Community Relations Collaborative. *Criminal Justice*, *18*(4), 34-42.

Rothman, J., & Ross, M. (2000). *Theory and practice in ethnic conflict resolution: Conceptualizing success and failure*. London: Macmillan.

Rupesinghe, K. (Ed.). (1995). *Conflict transformation*. New York: St. Martin's.

Ryan, K., & DeStefano, L. (2000, Spring), Disentangling dialogue: Issues from practice. *New Directions for Evaluation*, *85*, 63-76.

Schön, D. A. (1983). *The reflective practitioner*. New York: Basic Books.

Scriven, M. (1991a). Beyond formative and summative evaluation. In M. McLaughlin & D. Phillips (Eds.), *Evaluation in education: At quarter century* (pp. 19-64). Chicago: University of Chicago Press.

Scriven, M. (1991b). Pros and cons of goal free evaluation. *Evaluation Practice*, *13*, 55-62.

Smith, M. F. (1989). *Evaluability assessment*. Boston: Kluwer Academic.

Tjosvold, D. (1991). *The conflict-positive organization: Stimulate diversity and create unity*. Reading, MA: Addison-Wesley.

Turnbull, B. (2002). Program theory building: A strategy for deriving cumulative evaluation knowledge. *American Journal of Evaluation*, *23*, 275-290.

Weber, M. (1946a). Politics as a vocation. In H. H. Gerth & C. W. Mills (Eds.), *From Max Weber: Essays in sociology* (pp. 77-128). New York: Oxford University Press. (Original work published 1919)

Weber, M. (1946b). Science as a vocation. In H. H. Gerth & C. W. Mills (Eds.), *From Max Weber: Essays in sociology* (pp. 129-156). New York: Oxford University Press. (Original work published 1919)

Weiss, C. (1972). *Evaluation research: Methods of assessing programs*. Englewood Cliffs, NJ: Prentice Hall.

Weiss, C. (1993). Where politics and evaluation research meet. *Evaluation Practice*, *14*(1), 93-106.

Wholey, J. (1994). Assessing the feasibility and likely usefulness of evaluation. In J. Wholey, H. Hatry, & K. Newcomer (Eds.), *Handbook of practical program evaluation* (pp. 15-39). San Francisco: Jossey-Bass.

Wholey, J., Scanlon, J., Duffy, H., Fukumoto, J., & Vogt, L. (1971). *Federal evaluation policy: Analyzing the effects of public programs*. Washington, DC: Urban Institute.

Worthen, B., & Sanders, J. (1973). *Educational evaluation: Theory and practice*. Worthington, OH: Charles A. Jones.