Action Evaluation for Knowledge Creation in Social-Education Programs

By Victor J. Friedman, Ed.D.
Chairman of Behavioral Sciences Department
The Ruppin Institute,Israel
and
Jay Rothman, Ph.D.
Director, The Action Evaluation Project
The McGregor School of Antioch University
Yellow Springs, Ohio 45387

www.aepro.org

victorf@ruppin.ac.il

jrothman@mcgregor.ed

Abstract
This paper describes "action evaluation," a method for integrating knowledge production into the design, implementation, and assessment of social-educational programs. Action evaluation builds on an earlier vision of program evaluation as social experimentation but also incorporates concepts from conflict resolution and action science for promoting productive learning. These include program theories of action, goal inquiry, stakeholder deliberation, team building, conflict engagement, and web-based research. The paper concludes by raising a number of questions about the implications of action evaluation for knowledge creation in the social-human services.

Program evaluation represented an intentional, systematic attempt at knowledge creation long before the concept of organizational learning came into good currency. In the 1960's, U.S. government agencies and private foundations developed "programs" aimed at solving specific social problems through specific means of intervention. Legislators, administrators, and practitioners wanted to know whether programs "worked" and the academic community, especially social scientists and statisticians, quickly responded to this need for policy and evaluation research. They believed they possessed the methodological tools to provide administrators and legislators with valid information about program effectiveness. Donald Campbell, one of the fathers of the field, envisioned an "experimental society" in which social scientists and administrators would closely collaborate in using evaluation to gradually build a scientific knowledge base for social action (Campbell and Russo, 1999, p. 9).

This early promise of program evaluation as knowledge creation ran aground on methodological and political obstacles. Evaluators faced the thorny problem of establishing causality in the "real world," which rarely lent itself to the laboratory conditions necessary for conducting true experiments. Innovative approaches, such as the "quasi-

experimental" method (Campbell and Stanley 1963), placed high demands both on both program and research design, making evaluation a burden or simply unfeasible for many programs. Ironically increased research on an issue led to less of an ability to make clear and unambiguous judgments about success or failure (Cohen and Weiss, 1977), leading to the growing public impression that "nothing works" (Patton 1997, p. 13). Policy research often shifted from a focus on the issues to a focus on methodology and statistics. Most legislators and administrators showed little interest in methodology or in even using evaluation in decision making at all, tending to ignore unwanted evidence or co-opt/manipulate evaluation for their own political purposes (Patton, 1997). By the 1970's most program evaluators came to accept their field as inherently political (Weiss, 1993) and many were highly skeptical about evaluation utilization, much less social experimentation and learning (Patton 1997).

Action evaluation (AE) represents a recent attempt to revive the vision of program evaluation as social experimentation.. Action Evaluation is a new method of evaluation, one that focuses on defining, monitoring, and assessing success. Rather than waiting until a project concludes, Action Evaluation supports project leaders, funders, and participants as they collaboratively define and redefine success, as a project unfolds, until it is achieved.(for further definitions and examples of its application, see the AE website: www.aepro.org)AE has much in common with other evaluation approaches such as "utilization-focused evaluation" (Patton, 1997), "empowerment evaluation" (Fetterman, 1994), and "theory-driven evaluation" (Chen, 1990). [such as what?] It attempts to break new ground, however, by introducing concepts from conflict resolution (Rothman, 1997) and action science (Argyris, Putnam, & Smith, 1985; Friedman, 2000; Schon, 1983).

"Action evaluation" began when the loan officer of a major foundation asked Rothman (1997) if he could demonstrate that conflict resolution can be "successful." Rothman had to answer "no" because the very meaning of "success" in conflict resolution is unclear. This puzzled him and led to a study (Ross & Rothman 1999) which found that (1) definitions of success in conflict resolution programs are varied, local, and contingent upon the nature of the conflict and the desires of the stakeholders and (2) conflict resolution programs involve multiple stakeholders (e.g. participants, funders, professionals, administrators, evaluators) who themselves define success differently. Furthermore, Rothman noted that his own theory of conflict resolution (1997a) was constantly evolving through a process of conceptualization and testing in each new situation. In order to address the uniqueness, instability, and uncertainty inherent in programs like his own, Rothman began developing action evaluation (Rothman, 1997b).

Each action evaluation project is somewhat unique but they are all share the following theoretical and practical elements:

Three phases of theory building and testing. Mainstream evaluations tend to focus on the relationship between program inputs and outcomes, treating programs themselves as "black boxes" (Chen, 1990). In order to get inside the black box and to generate useful information, Chen (1990) proposed "theory driven evaluation," which looks at the goals, treatments, and contextual factors in both a program's plan/design ("normative theory") and its implementation ("causative theory"). Chen's approach was based on the observation that project stakeholders are often unaware of a program's implicit logic. Thus, the evaluator's role is to make this tacit knowledge explicit, providing stakeholders a more accurate understanding of how a program works and how it can be improved. Chen's (1990) distinction between normative and causative theory directly parallels the distinction between "espoused theory" and "theory-in-use" in the organizational learning literature (Argryis and Schon 1996) The latter, however, have gone much farther in explaining gaps between the two theories and the ways in which individuals and systems keep themselves unaware of them. From the perspective of action science, helping program stakeholders discover their tacit theories also enables them to achieve greater control over the choices they have made in their perceptions, their goals, and their strategies. By discovering their own causal responsibility for outcomes, they possess some leverage for producing change (Argyris et al, 1985).

The baseline phase of AE focuses on clarifying definitions of success and on making a program's "theory of action"

explicit by asking: What are the goals of relevant stakeholders? Why are these goals important to them? How are these goals to be achieved? Data on these questions are collected, analyzed, and fed back to program stakeholders (See Rothman and Friedman, 1999). The outputs of the baseline are both an explicit statement of program theory and an action plan for putting that theory into. The formative phase views program implementation as an iterative process of experimentation aimed at testing the program's theory of action. Data on program implementation and initial outcomes are collected, analyzed, and fed back to stakeholders in order to compare its espoused theory with its theory-in-use. Inquiry at the formative stage follows the logic of the spiral of organizational knowledge creation, making tacit knowledge explicit at the individual, program, and multiple stakeholder levels (Nonaka & Takeuchi, 1995, p. 73), leading to new problem settings, new goals, and new strategies. The summative phase focuses on (1) making judgments about the overall merit of a project and on (2) generalizing from the program's cumulative experience to other similar situations – not in the sense of establishing general laws, but in the sense of building a repertoire of exemplars which enable planners and practitioners to recognize both the similarities and the uniqueness of related problem situations (Schon 1983).

Goal inquiry. Early on evaluators realized that effective program evaluation depends upon the existence of clear, specific, and measurable goals, but this often led to complex goals clarification processes with all of their well-known pathologies (Patton, 1997). Recognizing that different stakeholders may have different goals, Chen (1990) made the identification and clarification of goals not just an antecedent, but also an object of evaluation itself. Peled and Spiro (1998) advocated a "goal-focused evaluation," which identifies both a program's "declared" and "operative" goals as a means of refocusing programs and rechanneling resources. The common denominator of all these approaches is the assumption that ends are "out there" to be found and specifically linked with means. AE, on the hand, challenges stakeholders to consider why these goals are important, facilitating "double-loop learning," which aims at inquiring into and changing underlying assumptions, goals, values, and standards for performance (Argyris and Schon 1974, 1996). Goal inquiry is also essential for addressing conflicts that inevitably surface when the goals of multiple stakeholders are taken into account.

A deliberative stakeholder approach. Learning depends upon engaging "people who want to know something" (Patton 1997, p. 50); that is, identifying and involving the stakeholders for whom a program is important and for whom information about it is relevant to their needs (Patton 1997). In addition program evaluation can function as an organizational learning mechanism to the extent that knowledge created for one stakeholder group reaches and is used by others (Lipshitz, Popper, & Oz 1996).

Most social-educational programs are produced through a bureaucratic process in which programs are developed at a center and implemented by a distributed population of practitioners, with evaluation acting as a feedback loop. Schon et al (1984) advocated "inductive planning," in which practitioners design, implement, and research their own policies with the center documenting and disseminating cases of intervention as well as maintaining the network of practitioners. Neither approach, however, takes into account fundamental discontinuities in the institutional structures of the social-educational services (Lipsky, 1980). Program effectiveness is often inhibited by gaps between planning and implementation (e.g. Bowen, 1999) or contradictions between the goals held by different stakeholders (e.g. Cohen, Friedman & Eran 1996). In these cases stakeholders were either unaware of these gaps and contradictions or acted as if they did not exist. AE is conflict-engaged, seeking to ensure that internal gaps and contradictions are surfaced and engaged (see below).

Furthermore, AE attempts to deepen stakeholder involvement and awareness through a systematic process of deliberation and consensus building around goal setting, program design, and monitoring. At the baseline, AE inquires into the goals of each stakeholder group separately and then brings the different stakeholder groups together to reach consensus on program goals and strategies. During the formative stage this deliberative process within and among stakeholder groups continues in reflection on the gap between espoused theory and actual practice. Reasoned, reflective deliberation fosters Model II values, which are considered critical for double-loop learning (Argyris and Schon, 1996). It

promotes "internal commitment" by giving each individual stakeholder and stakeholder group "voice" (Hirschmann, 1970). It also attempts to maximize the "free and informed choice of all participants"[where is this quote from?] regardless of their formal role or position in the hierarchy.

Team building. The goal of AE is to produce research in practice, not research on practice (Friedman, 2000) with no clear division of labor between those who produce knowledge and those who use it. Team building here means developing the roles, relationships, and norms for making theories explicit, testing them, and generating valid information (Argyris & Schon, 1996). Lipshitz et al (1999) have identified four behavioral norms which promote the production of valid information: transparency (exposing ones actions and the reasoning behind them), integrity (providing and encouraging full and accurate feedback even when it means admitting one's errors or pointing out those of others), an issue-orientation (focusing on the relevance of information to the issues regardless of social standing of the recipient or the source), and inquiry (asking questions for the purpose of genuinely understanding an uncertain or puzzling situations/behaviors). As discussed below, team building can also result from open and safe inquiry in to differences and disagreements within and across stakeholding groups in order to produce learning and change, when useful.

Conflict engagement. The evaluation community widely accepts that evaluation processes and purposes are often fraught with conflict, , but few evaluators address the issue of conflict directly. Patton (1997), for example, advocates the creation of "evaluation task forces" (pp 353-357) for putting competing interests on the table, but does not specify how they can be reconciled aside from using empirical evidence for "reality testing." There is a price for not fully engaging conflict. Because each stakeholder hopes that a program will meet its needs, there is an incentive to seek pragmatic, short-term solutions that enable programs to progress while leaving contradictions unaddressed. Programs often consume valuable resources but fail to get off the ground, fail to meet expectations, and/or explode when conflict can no longer be contained or avoided. It may often be program evaluators who light the match by exposing internal contradictions and who bear the brunt of built-up frustration and disappointment. Thus, evaluators also have an incentive to collude in the cover-up and avoidance of conflict such that programs not only fail, but also little can be learned from the failure as well.

In this common approach to conflict, it is viewed as a zero-sum game played out through force and/or bargaining. The zero-sum approach has been challenged by conflict theorists who suggest that clarifying and analyzing conflicting interests can reveal opportunities for win/win, integrative, or cooperative (e.g. Fisher & Ury 1981). However, many controversies involving social-educational programs can not easily be reduced to interests because they are rooted in deeply held "frames" involving perception, belief, and individual and collective identity (Schon and Rein 1994; Rothman 1997). Despite good communication and good intentions, identity conflicts frustrate attempts at achieving integrative solutions because they involve deep, and often passionate, differences as to what constitutes reality and what really matters.

Although AE acknowledges the destructive potential of conflict, it also recognizes that seemingly intractable conflicts offer opportunities for growth, adaptation, and learning (Rothman 1997; Schon and Rein 1994). Since social-educational programs are planned and delivered by people, they cannot be reduced to purely technical skills. Knowledge creation in this sector means producing deep changes in personal and professional identity as well as in the relationships among stakeholders Therefore, rather than focusing on solutions, action evaluators engage conflict by creating "reflexive dialogue" in which stakeholders critically inquire into their frames and underlying identity issues (Rothman 1997). Integrating conflict engagement with program evaluation provides stakeholders with opportunities for "negotiating reality," by testing their frames in concrete situations rather than in abstract ideological or political terms.

From the perspective of organizational learning, the ability to engage conflict is also important because evaluation inevitably involves providing information that is embarrassing and threatening. Evaluators need to be prepared to deal effectively with defensiveness, particularly when they themselves are in the position of revealing gaps between

espoused theory and theory-in-use (Argyris et al 1985).

Web-based strategies for data collection, analysis, and dissemination. From its inception AE has employed a site on the world wide web (see www.aepro.org), the heart of which is a database built on a Filemaker Pro platform. At the baseline stage the database has served primarily as means of collecting, organizing, analyzing, and storing data as well as generating and disseminating findings. In the formative stage, a customized and web-based discussion forum is used as stakeholders monitor and revise their goals and action plans as they seek to implement them.

The use of the Internet enables action evaluators easy access to programs and participants scattered around the world. Each project receives an account with its own password. Project participants log into their project area and enter their baseline data (What? Why? How?). Participant data is aggregated at the individual, stakeholder group, and project level. The results of the analysis are be posted on the website, but feedback sessions have mostly been held face-to-face.

In addition to data management, the website serves a number of functions directly related to organizational learning. First it provides an organized, easily accessible, cumulative record of program development and implementation. Second, it offers program participants who rarely meet face to face an opportunity to converse or exchange information either synchronously or asynchronously. Finally, by bringing together detailed evaluation research on programs in conflict resolution, education, and social services, it offers potential for comparative study, cross-fertilization, and cumulative learning.

Conclusion
Action Evaluation is a relatively new social technology (assisted by computer technology) having been launched about 8 years ago and now piloted in about 3 dozen projects. It is now moving in to a more mature phase in which partnerships with practitioners are beginning to yield experience that will either affirm or disconfirm the basic hypothesis at its core: enhancing participation of all relevant project or program stakeholders in defining, monitoring and adapting goals (or definitions and operationalizations of success) will enhance the utility of evaluation in helping success become self-fulfilling, which as Argyris et. al. suggest, is the real purpose of an "action science."

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

Argyris, C., & Schon, D.A. (1996). Organizational learning II: theory, method, and practice. Reading, MA: Addison Wesley.

Bowen, K. (1999). Development of local program theory: using theory-oriented evaluation to make a difference. Paper presented at the Annual Conference of the American Evaluation Association, Orlando, Florida.

Campbell, D. and Russo, J. (1999). Social experimentation. Thousand Oaks, CA: Sage.

Campbell, D. and Stanley, J. (1963). Experimental and quasi-experimental designs for research," in Handbook of research on teaching. Skokie, Il: Rand McNally.

Chen, H. (1990). Theory-driven evaluations. Thousand Oaks, CA: Sage.

Cohen, M., Eran, M., & Friedman, V. (1996). Evaluation of the "New Education Environment" program in Amal technological schools, First Interim Report. Jerusalem, Israel: Brookdale Institute.

Cohen, D. and Weiss, J. (1977). Social science and social policy: schools and race, in C. Weiss (ed.), Using social research in public policy making. D.C. Heath: Lexington, MA, 67-84.

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

Friedman, V. (2000, forthcoming). Action science: creating communities of inquiry in communities of practice, in H. Bradbury and P. Reason (eds.) The handbook of action research. Thousand Oaks, CA:

Sage. Fisher, R., and Ury, W. (1981). Getting to Yes: Negotiating Agreement Without Giving In. Boston: Houghton Mifflin.

Hirschman, A. (1970). Exit, voice, and loyalty. Cambridge, MA: Harvard University Press.

Lipshitz, R., Popper, M., & Friedman, V. (1998). Facets of organizational learning. Unpublished paper.

Lipshitz, R., Popper, M., & Oz, S., (1996). Building learning organizations: The design and implementation of organizational learning mechanisms. Journal of Applied Behavioral Science, 32, 292-305.

Lipskey, M. (1980). Street-level bureaucracy: dilemmas of the individual in public services. New York: Russel Sage.

Morgenthau, H. (1948). Politics among nations: The struggle for power and peace. New York: Knopf.

Nonaka, I. & Takeuchi, H. (1995). The knowledge creating company. New York: Oxford University Press.

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

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

Ross, M. & Rothman J., eds. (1999). Theory and practice in ethnic conflict management: conceptualizing success and failure. London: MacMillan.

Rothman, J. and Friedman, V., (1999), Action Evaluation: helping to define, assess and achieve organizational goals. Presented at the American Association Conference in Orlando, Florida, and Published in the Action Evaluation Website http://www.aepro.org/inprint/papers/aedayton.html

Rothman, J. "Articulating Goals and Monitoring Progress in a Cyprus Conflict Resolution Training Workshop." In Ross, Marc and Jay Rothman (eds.) Theory and Practice in Ethnic Conflict Resolution: Conceptualizing Success and Failure, MacMillan Press, London. 1999

Rothman, J. "Action Evaluation and Conflict Resolution in Theory and Practice", Mediation Quarterly , Vol. 15, #2, Winter, 1997a.

Rothman, J. "Action Evaluation and Conflict Resolution Training: Theory, Method and Case Study" The Journal of International Negotiation Vol. 2, #3, 1997b.

Rothman, J. (1997). Resolving Identity-Based Conflict: in Nations, Organizations and Communities. San Francisco, Ca.: Jossey-Bass.

Schon, D., (1983). The Reflective Practitioner. NY: Basic Books.

Schon, D. Drake, W., and Miller, R. Social experimentation as reflection-in-action. Knowledge creation, diffusion, and utilization. 6(1), 5-36.

Schon, D. & Rein, M. (1994) Frame reflection: toward the resolution of intractable policy controversies. New York: Basic Books.

Weiss, C. (1993). Where politics and evaluation research meet. Evaluation Practice 14(1), 93-106. (Originally published in 1973)