Reducing Anxiety and Resistance in Policy and Programme Evaluations

A Socio-psychological Analysis

IRIS GEVA-MAY
Simon Fraser University, Canada

WARREN THORNGATE
Carleton University, Canada

Evaluators of an organization whose programmes or policies are under evaluation frequently encounter anxiety and resistance from the evaluatees. Literature abounds with suggestions for developing collaborative interactions. Few studies, if any, show circumstances where the evaluatees' need for co-operation and support transcends the inherent anxiety and resistance and triggers for the evaluatees' desire for moral support and advice. An analysis of the socio-psychological factors involved in the circumstances described in this study may shed light on how the evaluator can develop the evaluation process to enhance the collaboration with the evaluatees. Here we report and analyse three case histories of such evaluations to determine the critical features of the evaluations that made them collaborative rather than conflictual. The common issues raised shed light on practices that alleviate anxiety in the evaluation process.

KEYWORDS: collaboration; policy; politics; programme; resistance

Introduction

Classical organizational, policy or programme evaluation usually prescribes strict and objective social science research methods to assess an organization's programmes, procedures, budgets or personnel, and to influence the implementation of policy. Evaluators attempt to avoid bias by divorcing an evaluation from political influences and by setting in advance the evaluation targets, questions, methods and measures. Though the evaluation is designed to assist the commissioning client in policy making, policy decisions based on its findings usually affect many others such as employees, sponsors, and the public utilizing the evaluated policy or programme.

Evaluation can be applied to institutional or governmental functions, organizations, policies or programmes. Yet organizational arrangements, policies and

programmes are inter-related, so evaluating one may be meaningful for reshaping or reformulating the others. Formative or summative evaluation of a programme may provide feedback for the modification or termination of a programme or may shed light on the modification or reformulation of the policy that has triggered it. DeLeon (1978) and Daniels (1997) assert that many ongoing programmes are terminated as a result of an evaluation, while the policy that triggered it is merely modified and followed by newly proposed programmes.¹

Several authors (Brewer, 1978; Cameron, 1978; Daniels, 1997; deLeon, 1978; Kaufman, 1976, 1987; May and Wildavsky, 1978) note that the object of the evaluation affects complexities and difficulties in the evaluation process. Programmes are the easiest targets to change; they can be more easily evaluated, dropped and replaced. Policies are harder to change because they codify general approaches or strategies of programme design and function. Yet policies are less resistant to evaluation than are organizations because organizations can sacrifice or change a policy to sustain their existence. The implementation literature, of which evaluation is an important component, widely discusses the inter-relations, interactions and distinct roles of policies and programmes (Bardach, 1977; Ingram, 1990; Mazmanian and Sabatier, 1989; Pressman and Wildavsky, 1984). The success of a policy depends on the extent to which the programmes that carry it out manage to survive in the real world and beyond the policy maker's office. Hence the evaluation of a programme does not necessarily reflect only the programme's effectiveness, efficiency or sustainability, but also the viability of the policy that generated it.

Perhaps because of the stakes involved in evaluation, evaluatees involved in any evaluated organization, policy or programme often feel threat, apprehension or anxiety about the evaluation process, and often try to influence or resist it (Chelimsky, 1985; Geva-May and Pal, 1999; Palumbo, 1987; Weiss, 1983; Wildavsky, 1979).² Skilled evaluators recognize such reaction and seek collaboration with the clients and stakeholders to reduce it, a goal more frequently recommended than met (e.g. Patton, 1997; Weiss, 1975). Yet some evaluations may produce the opposite reactions of co-operation and collaboration between stakeholders and evaluators, positive feelings of mutual interest, benefit, moral and psychological support. These are, of course, the noble goals of participatory and collaborative evaluations, goals that have inspired a large body of literature on participatory and collaborative evaluation methods (e.g. Whitmore [1998] and any of Google's 168,000 hits for 'participatory evaluation'). There is, however, scant literature on the social psychological factors that determine when stakeholders will resist or facilitate an evaluation. A better understanding of the social psychological factors underlying such responses could improve future evaluations.

Three cases of spontaneous collaboration among evaluators and evaluatees are reported below (Geva-May, 1991a, 1991b, 1991c). All are related to emergency circumstances in which, to the pleasant surprise of the evaluators, the stakeholders did not view the evaluation as an interference, burden or intrusion into their work, nor as a prelude to threatening changes. The social psychological factors identified in these cases might provide guidance for reducing anxiety from

and resistance to the evaluation processes beyond the vague prescription of seeking evaluatees' or stakeholders' collaboration. This article will:

- describe the evaluation processes undertaken, the circumstances of the evaluation, and the evaluatees' responses;
- analyse in social psychological terms why the evaluations elicited far more collaboration than reactance; and
- offer lessons about how to attenuate feelings of evaluation anxiety and resistance.

Literature Review

As long ago as 1950, Ralph Tyler defined evaluation as a process of determining to what extent the aims and the objectives of an ongoing programme or project have been attained. Other definitions refer to rigorous social science methods for the study of projects and programmes (e.g. Bernstein and Freeman, 1975; Rossi et al., 1979) or for comparing the relative cost and benefit of several programmes in order to expand the narrower Tylerian view (see Alkin and Ellett, 1984). Summarizing prevalent definitions, Nevo (1983: 5) regards the evaluation process as a 'systematic description of . . . objects and/or an assessment of their merit and worth'. The valuative aspect of evaluation, 'the systemic investigation of merit and worth', is illustrated by the Joint Committee (1981) or, for instance, in articles by Eisner (1979), Glass (1969), House (1980), Scriven (1967) and Stufflebeam (1974).

Although their definitions vary, all those above are related to classical evaluation functions. There are, of course, many more aspects of evaluations than these functions, depending on the information an evaluation seeks. The CIPP Model (Stufflebeam, 1971, 1972) proposes that evaluation gathers information related to the *goals, design and implementation* processes, and to the outcomes of the evaluation object. Such evaluations emphasize the *merit and worth* of these evaluation targets. Guba and Lincoln (1981) refer to descriptive information about the target object and its setting, to standards of worth, to *information that is relevant to certain audiences*, to issues of *interest*, or to *values of concern*. Patton (1987) broadens the scope of programme evaluation, regarding it as:

a systematic collection of information about activities, characteristics, and outcomes of programs, personnel, and products for use by specific people to *reduce uncertainties, improve effectiveness*, and make decisions with regard to what those programs, personnel or products are doing and affecting. (Patton, 1987) (authors' own italics)

All agree that the information obtained in evaluation studies can assist policy formulation, policy execution or accountability processes in public decision making (Chelimsky, 1985).

Cronbach (1963), Stufflebeam (1971) and others emphasize the role of evaluation in decision making. Cronbach's definition of evaluation directly points to information collection and use in deciding about a programme under review. Stufflebeam (1971), following Braybrook and Lindblom (1970), regards the role of evaluation as 'the process of delineating, obtaining, and providing useful

information for judging decision alternatives' (Braybrook and Lindblom, 1970: 128). This implies that decision makers should ascertain the relative values of two or more alternatives and that evaluation findings can provide this type of comparative information.

Evaluation has been charged with being *unrealistic* because it sets standards too high, *irrelevant* because its answers are not of immediate, practical interest, and *unfair* because it adheres to its sponsors and does not provide information to less powerful stakeholders (Rog and Fournier, 1997; Whitmore, 1998). Other objections to evaluation note the anxiety and rejection it creates (e.g. Weiss, 1975), both because of its valutive connotations and because, as Chelimsky (1985) noted, some evaluations can be used as tools of political infighting. With regard to anxiety and rejection, Behn (1978), deLeon (1983), Geva-May (2001), Lynn (1987) and Wildavsky (1979) note the unwillingness of those benefiting from current organizational arrangements to 'rock the boat'. All valutive efforts – measures of worth or merit, assessments of goal achievement, or recommendations for change – have the potential to produce such reactions.

Hence, the definition of evaluation as a process that involves both description and judgment (e.g. Guba and Lincoln, 1981; Stake, 1967) has been rejected by Cronbach and the Stanford Evaluation Consortium (1980) because of the dangers of anxiety or resistance it can engender. Instead, they define evaluation as a 'systemic examination of events occurring in and consequent of a contemporary programme – an examination conducted to assist in improving this programme having the same general purpose'. This definition will be assumed in our case analyses.

Many authors point to the importance of collaboration among the people who commission or perform an evaluation and the people whose policy or programme is evaluated (Chelimsky, 1985; Cousins and Leithwood, 1986; Geva-May and Pal, 1999; Palumbo, 1987; Patton, 1987; Weiss, 1983, 1985, 1998). Collaboration is seen as one way of reaching consensus to ensure implementation, rather than rejection or ignoring, of findings. It is sensible to assume that, if clients and evaluators communicate well about all aspects of an evaluation, anxiety should decline.

The evaluator's primary stakeholder is typically the client who asks for recommendations based on evaluation research. It is said that if the client helps define the problem, set the evaluation goals, and design the evaluative framework, then the client has at least a nominal stake in the findings. In 1985, for example, Weiss (1985: 211) listed seven different meanings of the 'utilization' of policy research. When the client is involved, the evaluator has less need to be concerned about resistance. Nevertheless, evaluation results can still be threatening to evaluatees, especially if the evaluation is commissioned by an external client. Moreover, since evaluation is primarily retrospective and looks at context as factors that facilitated or impeded policy success, anxiety is bound to emerge.

Case Studies

Although case studies are limited in span, they help develop research questions and answers to major problems in evaluation. The three case studies presented below contribute to the existing theoretical literature by exemplifying evaluation approaches that can be borrowed or adapted to reduce resistance and alleviate anxiety. The studies all pertain to nascent policies and their implementation as related programmes.

Our three case studies address the thesis that evaluation anxiety and resistance can be reduced when:

- the evaluation is perceived as a formative assessment of nascent programmes and policies regarded as pilots needing refinement and perfection;
- evaluations do not require all evaluation criteria in advance and thus allow for early collaboration with the evaluatees on the refinement of assessment standards;
- policy makers, administrators and front-line implementers see themselves as having common goals or being 'in the same boat';
- the goal of the evaluation is seen as assessing policies rather than people;
- the evaluation gives the stakeholders, including the evaluatees, a voice in policy improvements that would affect its implementation at the programme level;
- evaluators serve as conveyors of information, supporters and advocates and induce feelings of reliance, dependence, interest and trust in the evaluatees, thus promoting a willingness to collaborate.

The evaluatees are more co-operative when they believe that they will get something in return (e.g. useful and frequent feedback and information) for their participation.³

Two of the three case studies date from 1990–2 following the large and sudden wave of emigration from the former Soviet Union (FSU) to Israel after the fall of the communist empire. Although the same wave brought emigrants to other shores, 400,000 arrived in Israel, increasing the population by 20 percent between 1989 and 1996. Forty percent of these immigrants were academics or professionals (160,000 people), 4 percent were teachers, and almost 2 percent were scientists or research and development engineers (more than 6,000 people). Since Israel is a country of immigrants and its *raison d'être* is to be a homeland for any person of the Jewish faith who wishes to come to Israel, the Law of Return enacted in 1950 by the Knesset (the Israeli Parliament) states that 'every Jew has the right to immigrate to the Land of Israel' (*Book of Laws*, 1952). Thus, the Israeli Government adheres to the view that immigrant absorption of adults and children is its responsibility, as is employment, which serves as a major tool in the social absorption process.

The first case study concerns steps taken by the Government to solve immigrant employment problems (Geva-May, 1991b). Rates of absorption in the technical and liberal arts professions were only about 35 percent in the first few years of the programme.⁴ The main thrust of the Government's effort was in training

or retraining immigrants and finding suitable employment for special groups such as scientists, specialist engineers, physicians and artists in the 45–64-year-old age group (Ministry of Immigrant Absorption, 1996).

The second case study concerns a centralized effort to absorb immigrant children into the Israeli school system (Geva-May, 1991c). Between December 1989 and September 1991, Israeli schools had to accommodate about 70,000 students of various ages, more than 6 percent of the total student population in Israel at the time. This case study relates to an evaluation initiated by the Ministry of Education that examined the effectiveness of policies and programmes intended to accommodate these students.

The third case study relates to the situation produced by the Gulf War in Israel of mid-January to February 1991 and the policy made to protect the educational system at that time (Geva-May, 1991a). The Flexible School Program (FSP) of the Ministry of Education was intended to ensure that the students were kept busy and occupied during the days of the war by creating appropriate learning opportunities, providing for their safety, and minimizing the disruption of normal education activities.

Case Study 1: Evaluation of the Retraining of Immigrant Scientists, 1990–3

Faced with the inability of the Israeli job market to absorb special groups such as scientists and R&D engineers, the Ministry of Science and the Ministry of Immigrant Absorption began offering retraining possibilities. Scientists who could not obtain a position in academia were encouraged to join retraining programmes given in Israel's six universities to become high school teachers.⁵ It was clear to all concerned that although these universities had accepted the challenge, it was the first time they had undertaken such a training project, on very short notice and for a very special population.

In creating this retraining programme, the Government's policy had three goals:

1. to provide future stable employment of the scientist immigrants
2. to offer an immediate absorption solution
3. to benefit the high schools needing qualified science teachers, especially in development towns, low socio-economic areas and localities with disadvantaged populations.⁶

A three-year evaluation study was commissioned to examine the feasibility and potential of such a retraining policy (Geva-May, 1991b). The information needs dictated a Tylerian evaluation investigating the attainment of the three main policy goals above. Eventually, the study was awarded by the Ministry of Science and the Ministry of Immigrant Absorption to a Technion researcher and was conducted in five of the six universities taking part in the project.⁷

The success of the retraining policy could be gauged by observing its execution and by detecting attitudes towards its implementation during the first year, 1991–2. The main evaluation tools were semi-structured interviews and attitude questionnaires designed to gather factual information and assess the attitudes

of the programme directors, lecturers and immigrant trainees toward each other and toward the suitability of the retraining programmes. The questionnaires were distributed to the five university programme directors, to a random sample of 30 lecturers and to the immigrant scientists. The evaluation examined the profile of 193 immigrant scientists and the suitability of the retraining models, and offered formative process information that provided extensive technical feedback and primary inter-institutional information (e.g. Chelimsky, 1985; Scriven, 1967).

One of the main issues of concern pertained to each institution's criteria for acceptance into the retraining programme. For instance, the profile of the immigrant scientist trainees varied from PhD scientists with renowned research experience but no teaching experience, to MSc researchers with publications in their field of expertise. Some of the trainees clearly needed to be updated in content areas taught in the West and in complementary areas such as the use of computers. Other questions related to mastering Hebrew. Some institutions required that participants be in the country for a minimum of two years, assuming that this would indicate proficiency; others required proficiency exams. Age was a controversial issue in the institutions and a major moral and professional dilemma.⁸ The age range of the participants was 30–55, but some institutions did not accept trainees over 40 years of age. The findings of the evaluation suggested that it was advisable to give preference to candidates who had previous teacher training, that the number of participants in the programme should be limited to small groups of no more than 20 in order to offer the trainees individual attention, and that age should not be a criterion of exclusion.

The retraining programme profile was the other main evaluation area of concern. In all departments the programme ran parallel to the academic year: October–July. Most of the lecturers were staff members in the department. Methodology courses were given by experienced teachers. Nevertheless, tutors had no prior preparation about the culture and habits of the immigrant trainees. In each university the final assessment of the immigrant scientists was determined by the standards set for Israeli students. The academic requirements concerned content matter and teaching performance. The major subjects in all institutions were discipline-oriented and general education subjects, including methodology, the use of computers for teaching, basic review and content matter updating in mathematics, physics, chemistry, mechanics or electronics. The general-education subjects offered were an introduction to teaching, an insight into the educational system of Israel, social and cognitive psychology, and the psychology of the child. The universities chose to teach the same subjects and provide the same assessments for certification as in their regular teacher-training programme because it did not discriminate between the populations of future teachers.

Nevertheless, the evaluation pointed to the need for substantially more hours devoted to subject-matter updating. Scientists who had specialized in a narrow field of expertise did not always master basic aspects necessary for teaching in schools. Only one institution allocated enough hours (360) for this purpose. Findings also suggested that the psychology classes should emphasize the mentality and social behaviour of the host country to enable teachers to develop tools

Evaluation 9(2)

for social interaction and communication with students and parents. Major pedagogical issues such as deductive and inductive learning, co-operative versus individual learning, and group work procedures were new to most trainees. Also, the evaluation showed that trainees needed more exposure to evaluation and assessment methods and practices, teaching approaches for different learning levels and student populations, and exposure to books and materials. The only demand imposed by the Ministry of Education and Culture, additional weekly hours for subjects such as the language, culture, history, literature and geography of Israel, received different degrees of emphasis in each institution. The evaluation clearly showed that the social and cultural integration of the immigrants, their orientation and identification were helped by lessons in history and geography and by the organized tours of Israel.

Mastery of language was viewed as necessary for both teaching and professional integration; so was the feedback that many more hours should be dedicated to language learning. It was suggested that professional language, children's slang and listening comprehension should be emphasized in the retraining programme.

The prerequisites for practicum sessions were the same as the requirements imposed on the Israeli teacher-training students, namely 50–80 hours over two months. During this time each immigrant teacher was assigned to a school and worked under the supervision of a teacher-tutor. The evaluation showed that these sessions were important for future professional integration and that the programmes should allow for considerably more such exposure. Moreover, the immigrant trainees should receive follow-up counselling once they started teaching.

All of the programme design decisions stated above had been determined intuitively prior to the policy's implementation on an individual, ad hoc basis by each institution, and with no previous experience in retraining immigrant scientists to become teachers. Ultimately, in the minds of the evaluatees, the main role of the evaluation was to validate assumptions about the feasible characteristics of such a programme, sharing concerns and learning from what was done in other institutions.

It is interesting to note the difference of approach taken by the director who declined to take part in the evaluation. She knew and befriended the evaluators and was willing to attend an initial meeting but, as discussions about the programme and the evaluation plans continued, she showed clear signs of apprehension and a desire to distance herself from the project. She finally announced that she would not be willing to take part in the evaluation study because the retraining programme was still at a formative stage. Apparently, she feared any type of intervention while developing the new programme and was worried about damaging the reputation of her institution. This is precisely the type of response (reluctance, anxiety, rejection) that evaluation literature discusses, and evaluators expect.

For the five institutions that did take part in the evaluation study, responses to the evaluation study were unexpected. All lecturers and programme directors were keen to meet with the evaluators. They sought information on the practices,

decisions, materials, criteria for success, and coping strategies in the other institutions. They valued the fact that the evaluators knew the other programmes and hence were able to compare initiatives. Calling the evaluators, sometimes at home, or inviting them for coffee so that they could further share or discuss issues of concern became a common practice. Coping strategies, attitudes, rationales, and pitfalls were openly discussed. Personal incidents with trainees, or ways of helping trainees with personal problems, were also subjects of discussion. The evaluators often sensed that the evaluatees were looking for opportunities to meet and talk. They were keen to succeed, considered the programme to be of significant national and humanitarian importance, were genuinely interested in exposure, sharing and support, and were open to feedback on what they regarded as a new, ad hoc, programme to develop an absorption policy.

Case Study 2: Organizational and Pedagogical Aspects in the Absorption of Immigrant Children in Schools

This evaluation was one response to the sudden wave of immigration to Israel in the 1990s, undertaken to assess how well the schools were absorbing over 70,000 school children into the educational system (Geva-May, 1991c). The Ministry of Education was interested to know how the schools were coping with this acute problem, particularly in view of the lack of time for preparation. The aim of the evaluation was to provide the decision makers who commissioned it with baseline data that could be used in a longitudinal assessment of the ways and means of student absorption and of the problems that arose in the process. The evaluation was undertaken in 1991. The evaluators came from the main evaluation office in Jerusalem and from the branch office in each of the Ministry's six districts. All were senior education civil servants, mainly senior teacher counsellors or superintendents trained as evaluators. The subjects of the evaluation were the organizational and pedagogical arrangements that schools had made to absorb immigrant students into the educational system.

The evaluation examined 32 elementary and high schools in each of the four districts representing different geographic and socio-economic regions of the country. Each school had at least 10 percent immigrant students. In total, 24,124 students were enrolled in these schools, of whom 3,087 (12.7%) were recent immigrants.

The evaluation focused on the following four themes.

1. The type and extent of learning activities, e.g. the number of hours allocated for this purpose and their source and purpose, the educational frameworks for immigrant students, the subjects and topics provided within these frameworks, support and guidance for preparation for the final matriculation exams.
2. The type and extent of socialization activities.
3. The training and counselling of teachers to deal with immigrant students, i.e. degree of preparation, institutional or individual counselling, etc.
4. Problems encountered by the heads of schools, teachers, and students, e.g. Hebrew language classes, the transfer of students to mainstream classes or

Evaluation 9(2)

special immigrant classes, the degree of integration of the more 'veteran' students in the mainstream class, the willingness of teachers to absorb new immigrants, and the native students' willingness to assist.

To gather data on these four themes, an attitude questionnaire was administered to the heads of schools and the teachers, and interviews were also conducted to clarify and cross-validate the questionnaire findings. The findings showed a high degree of effort and investment from heads of schools, teachers and native Israeli students. Despite a general feeling of cautious preparedness, there was also a high degree of puzzlement and concern about the process in what was viewed as a completely new situation.

Respondents shared concern about a lack of classrooms and problems in acquiring suitable books and educational equipment. At both school levels, the respondents reported difficulty in recruiting and employing suitable teachers who could teach immigrant classes, teachers who knew the students' language and understood their mentality and behavioural patterns. The need to train native teachers to handle immigrant student populations was obvious, but no such programmes existed then.

Key concerns in all the schools included setting criteria for the allocation of time in the immigrant language class or questioning whether immigrant students should be separated, setting thresholds for the level of linguistic mastery allowing a child to move from the language class to the mainstream class, and the extent and type of individual tutoring. This was especially true in the high schools where mastery of Hebrew was viewed as more problematic than in elementary schools.

Schools encountered communication problems in two main areas. The first was between the school staff and the students' parents, some of whom were embarrassed to reveal their poverty by visiting the school or being visited at home, or had difficulty in accepting how children were disciplined in that system or did not understand the significance of matriculation exams. The school staff had difficulties in coping with the attitudinal differences of immigrant students. Yet, while school psychologists tried to address these issues, students, teachers and school administrators still faced daily communication problems and solving them was a major concern.

In both high schools and elementary schools, the main pedagogical difficulties could be traced to inadequate learning materials. Elementary schools lacked suitable and motivating programmes. Many high school immigrant students excelled in mathematics, physics, sports and the arts, but lacked challenging materials in these areas at a linguistic level they could comprehend. They had no academically adequate and linguistically suitable textbooks, which would challenge and prepare them well for the matriculation exams. Providing extra-curricular materials for socialization to the Israeli culture was also found to be an important issue of concern.

The evaluators felt welcome when they visited the schools. On many occasions, when new projects or new materials were initiated, or when problems arose, school directors, programme chairs and sometimes the teachers, would phone the evaluators and ask to meet. Many stayed on after the evaluation

meetings had ended. They felt relieved by the opportunity to share their frustrations, complain, pass on requests to the Ministry, make suggestions, and ask for information about what was being done in other schools. They expressed a strong commitment to the programme but felt overwhelmed by the pressure of the problems and the responsibility that the circumstances imposed. They found support in sharing their concerns and doubts, consulting about standards, criteria and courses of action, crying on the evaluator's shoulder or sharing feelings of achievement.

Case Study 3: Organizational Arrangements of the Israeli School System During the Gulf War

During the Gulf War, Israeli fear of being attacked was high and, indeed, several destructive Iraqi missiles were launched into heavily populated areas of the country. The normal course of life was disrupted and security matters were of high priority. In a state of uncertainty, the Ministry of Education decided to implement the Flexible School System policy designed to minimize the number of children in the schools. Except for those in grades 11 and 12, students were required to stay at home and were provided with learning materials to study at home. This involved coming to school once a week with their parents to obtain these materials. Small group discussions were held either during those school visits or in homes visited by itinerant teachers.

An evaluation study of the Flexible School System Program was commissioned by the head office of the Ministry of Education in Jerusalem, and was performed by the Evaluation and Planning Division (see details in Case Study 1) between 27 January 1991 and 2 February 1991 (Geva-May, 1991a). A first study was undertaken between 27 January and 31 January 1991, while the Flexible School System Program initiated by the Ministry of Education was in process. A second survey was undertaken between 3 February and 10 February 1991, while the school system was slowly returning to normal.

The evaluation focused on the organizational and pedagogical arrangements of the Israeli school system in order to keep the students safe, occupied and involved in the learning processes. The key questions pertained to:

- learning programmes (to what extent the schools can rely on existing materials or need to prepare new ones, to what extent schools prepare materials related to ongoing events);
- keeping in touch with students (to what extent the schools have an 'open line' system and whether it is used, organizational arrangements to meet with small groups of students, etc.); and
- teacher and staff performance.

The tools employed were feedback survey-type questionnaires and interviews addressed to school directors. To validate and expand on the findings, content analysis of documents and Ministry directives to the field was also undertaken. About 15 schools in each of the seven districts of the Ministry of Education took part in the study.

The general findings addressed more than the questions first raised by the

Evaluation 9(2)

client; additional issues such as budgetary problems and safety concerns were also raised by the respondents and became additional evaluation questions.

The study found that in all the schools the large majority of teachers came to school as usual. It was suggested that additional manpower should be available to the school in emergency times and that a drawer name-list of volunteers be drafted as well as a list of nurses and psychologists.

Teachers and educational advisors kept in touch with students and their parents even before the Ministry officially declared the Flexible School System Policy. Special educational materials were based on the regular curricular programme, but games and fun activities were also provided to dissipate tension. Teachers requested that special disciplinary counsellors be available for advice on the materials that needed to be prepared. Schools encountered budgetary problems because of the amount of new materials that had to be developed, duplicated and distributed, the need to purchase plastic window covers for classrooms, by the cost of the ongoing flow of phone calls, and the purchase of teaching materials.

Safety issues were another concern. The school directors referred, for instance, to the need for one authoritative body either at the state or the local level to issue directives, as it was not always clear which age groups were allowed to come to school (especially in six-year comprehensive high schools). Moreover, the directors noted that additional telephone lines were needed and that parents found it a burden to accompany their children to school. Additional issues concerned the need for programmes and policy at the level of standby 'drawer directives' and 'drawer materials' prepared by pedagogical centres for emergency circumstances.

In general, school managers preferred to receive instructions from the Ministry directly and not through the media, and pointed to the need to receive guidelines about ways to cope with administrative problems. The managers did not feel the need for directives on pedagogical issues as they felt that they were capable of coping with these issues adequately; and they required significantly more involvement by educational counsellors, nurses and psychologists.

Faced with this myriad of budgetary, security and pedagogical problems requiring adequate solutions, the heads of schools and co-ordinators felt grateful for the appearance of a new player on the field whose role was to negotiate the criteria for what was right, wrong or feasible regarding the implementation of an ad hoc policy that needed to be tested in real time. The evaluators' intervention was regarded as an attempt to provide support and advice, to share concerns, to consult about materials and procedures, to convey demands to those in charge at the national or local level, and to share information gathered from other schools.

General Observations of the Evaluation Process

The scientist retraining courses provided in the universities at the request of the Ministry were designed by their Schools of Education under tremendous time pressure. Despite their willingness to participate in the project, the Schools had no previous experience with retraining programmes for new immigrant scientists and were forced to design them ad hoc. We expected resistance to our

intervention because of the stakes involved regarding their prestige. Indeed, one of the Schools of Education responded 'normatively' – as most evaluation literature suggests – with apprehension and resistance, and proffered reasons for not participating in the evaluation study.

In Case Study 2, school directors, overwhelmed by the numbers of immigrant students in the K-12 (kindergarden to grade 12) educational system, were happy to have someone to listen to their problems and concerns, even the 'evaluators'. All attempts to serve this large immigrant population were new, experimental and had no previous normative criteria by which to judge them. Commitment to effectively promote the absorption policies/programmes, together with doubts, concern and puzzlement over their respective programme design and potential outcomes, led directors to confide in and accept the external evaluators as a source of support.

In Case Study 3 the co-operation and collaboration of the school directors with the evaluators were particularly poignant. Despite the reluctance of the evaluation team to enter the schools under the special circumstances, its members were warmly welcomed. The evaluation team expected modest collaboration at best because an evaluation could place an additional burden on the system while the school directors were busy and concerned about the activation of the Flexible School System programme and about students' safety. Nevertheless, the schools showed a genuine desire to use the evaluators' services in order to co-ordinate their educational and organizational activities with the head office, to share concerns and problems, and to set standards for performance. Indeed, the evaluatees showed a particularly pro-active attitude by suggesting additional evaluation items that had not been foreseen by the evaluators, offered sample copies of the materials they had developed in order to be shared with other schools, suggested ways for programme improvement and feedback for policy reformulation. In all three cases the evaluation process was regarded as useful and important, and the evaluators were viewed by the evaluatees as sympathetic listeners and providers of advice.

Analysis

It is, of course, difficult to document the absence of anxiety or resistance in the three case studies. It is possible that anxiety and resistance did exist but were too weak or infrequent to be detected. It is also possible that the methods of detecting anxiety and resistance were so insensitive that nothing short of an evaluatee's revolt would have indicated that anything was amiss. What can be said is that the senior author's involvement in the three evaluations gave her and her evaluation teams the memorable impression that these cases were unlike almost all of the dozens of other evaluations in which they had participated. Return rates of the evaluation questionnaires approached 100 percent. Most interviews were long and sincere, rather than perfunctory, and produced high suggestion/criticism ratios. Attendance at meetings of evaluators and evaluatees was high, as was the evaluatees' interest in the findings. Rumours of an evaluator's possible ulterior motives never surfaced. Evaluatees initiated additional official or unofficial

meetings, phoned or stayed on in order to share their views, fears, achievements, consult, seek or pass on information. These are all indicants of collaboration and co-operation rather than resistance, of evaluative comfort rather than anxiety.

Yet if these three evaluations exhibited less anxiety and resistance, more collaboration and co-operation, than other evaluations, they raise the question: what made them different? The prescriptive literature reminds us of the importance of an evaluator's professionalism and listening skills in the success of the data-collection process, but the evaluators in the three cases did not exhibit these strengths to an extraordinary degree. The literature also reminds us of the importance of developing trust and rapport with the evaluatees prior to data collection. There was certainly trust and rapport between the evaluators and the evaluatees in the three cases, but the evaluators made no extraordinary efforts to develop it.

To account for the extraordinary lack of anxiety and resistance, and the high degree of co-operation and collaboration, let us first rule out cultural differences. Israeli society is no less anxious or resistant, no more co-operative or collaborative, than most other societies. Anxiety and resistance to programme or policy evaluation are as common in Israel as anywhere else. The Ministries of Education and Absorption have each had their fair share of intra-organizational conflicts. Like groups within all organizations, relations among senior administrators, bureaucrats, school directors, teachers, and other groups with these Ministries have not always been tranquil, collaborative and co-operative. We are thus left with a shorter list of factors that could have influenced the attenuation of anxiety and resistance and promoted co-operation and collaboration. Two remaining factors seem especially important. First, the programmes were nascent. Second, they were born of national emergencies.

Each of the three programme evaluations was formative rather than summative. The programmes were new and emerging, the policies that prompted them were new, and the evaluations reported above were the first conducted of each. The programmes were also unique to their historical circumstances, offering no precedents, standards or criteria for performance. Points of reference for judging policy and programme effectiveness had not yet been established, so no direct historical or social comparisons could be made. As a result, the evaluatees could understand from the context of circumstance that their current efforts would not be compared on an extant metric to their own past performances or to the performances of others.

Still, new and unknown situations generate anxiety, and anxiety creates a need for affiliation with others in similar circumstances. In a classic series of laboratory studies, Schacter (1959) induced feelings of anxiety in volunteer participants and then gave them a choice either to sit alone or with other anxious people. Most chose to sit with the others and were calmed by their company, as we often experience when we call a friend or seek a support group to share our feelings. The company alone may calm, as it does when infants stop crying in the arms of a soothing parent, but the company also gives us a chance to acquire information, which in turn may reduce anxiety. The circumstances of the three case studies gave school administrators ample reason to be anxious and thus a desire to

affiliate with one another. Lacking opportunities to meet, the administrators' next best alternative seemed to be affiliation with the evaluators who were trying to understand their problems and who could shuttle messages of comparison among them. In this respect, the evaluators assisted implementation through their feedback and provision of information. Nevertheless, in the three cases above, the evaluation process was distinct and not part of a team-building strategy.

Anxiety not only originates in unknown aspects of new circumstances. It also is born in reaction to performance evaluation, as evident in the pacing of students before examinations and in stage fright. Historical and social comparisons often stimulate evaluation apprehension or anxiety (Geen, 1983; Leitenberg, 1990; Zajonc, 1965). Many people become anxious when they believe they might not do as well as they had in the past, or when they believe they might not do as well as others. Their anxiety is amplified when they also believe that their performance will be public and when a bad performance will lead to social disapproval or exclusion. There are many strategies to reduce or contain the resulting anxiety, and several are known to be employed by evaluatees. Probably the most common strategy is to exhibit passive resistance by withholding information: declining or missing interviews; becoming too preoccupied or forgetful to return a questionnaire; feigning an illness or emergency to justify a tactical absence from work. Falsifying information to 'look good' is also common, as is organizing fellow evaluatees directly or through unions to co-ordinate more active resistance strategies.

The absence of such resistance in the three case studies suggests that the evaluatees were not anxious about the evaluation. Subjective impressions based on the evaluatees' responses, from participatory to enthusiastic, support the suggestion. The evaluatees may, however, have been affected by other factors. The three programmes were a source of additional work demands, possibly leading some evaluatees, especially the teachers and administrators on the 'front lines' of the new programmes, to feel additional job stress, loss of the time and freedom to pursue normal work activities. Loss of freedom often leads to psychological reactance (Brehm and Brehm, 1981; Dowd et al., 1988), a combination of motivation and action to restore the freedom lost. If the evaluators were viewed as instrumental in 'fixing' the three programmes, their recommendations could be seen as tools to regain lost freedom, so it would make strategic sense for the evaluatees to collaborate.

The emergencies that gave rise to the three programmes created special social psychological conditions likely to reduce anxiety and resistance and to increase collaboration and co-operation. Because of the emergencies, there was great pressure for these programmes to succeed; they could not easily be abandoned if they failed. This pressure was felt not only by the administrators and teachers in these programmes – groups often in conflict over more traditional policies and programmes; it was also felt in the Ministries with the money and the mandate to address three pressing social or safety concerns. As a result, the Ministers who commissioned the programme evaluations, the programme directors and teachers were suddenly in the same boat, all trying to row in the same direction.

They had, in the lexicon of social psychology, a super-ordinate goal, to make the programme work, a goal that could be reached only through joint effort.

Sherif (1966; see also Diab, 1970) demonstrated the powerful influence of super-ordinate goals in developing co-operation among conflicting groups in a boys' summer camp. Often the super-ordinate goal is to survive a common enemy, a goal clearly evident in the development of armies and in Case Study 3. But the super-ordinate goal can also be the public performance of a group effort. Actors and ensemble musicians learn to co-operate and collaborate in search of audience approval or in fear of audience disapproval. All of the actors from the Ministers to the teachers in Case Studies 1 and 2 had the audience of Israeli society watching their performance and thus good reason to ensure they did not fail. With the external pressure to reach their super-ordinate goals came the internal pressure to co-operate in reaching them.

The nascence of the three programmes, the emergencies that gave rise to them, and public pressures for quick solutions constrained the Ministries to improvise many of the policies for programme development and implementation. Perhaps because of this, the Ministries recognized that feedback on the pursuit of goals by means of the programmes was needed to adjust the policies and their implementation. Such recognition was reflected in the questionnaires and interviews, which evaluated policies and related practices rather than the people responsible for implementing them. The success or failure of any programme is, of course, the sum of the quality of its policies and the quality of its personnel. But the psychological difference in evaluating these two foci is profound. Evaluatees often feel justified in being anxious and resisting any attempts to impose evaluations on them, especially if they believe the evaluations will not be favourable. This is the case referred to by Wildavsky (1979), for instance, where the evaluation was viewed as rocking the boat and weakening organizational cohesion. But evaluatees have good cause to be forthcoming, to provide information and suggestions, in any evaluation leading to programme or policy improvements.

This focus-induced difference in an evaluatee's attitude and response toward an evaluation reflects the social psychological processes of causal attribution (Graham and Folkes, 1990; McArthur, 1972; Martinko et al., 1995). Evaluators are not the only players looking for reasons behind a programme's successes and failures. Every evaluatee is likely to have his/her own explanation; it is a natural aspect of human cognition (Heider, 1958) and a primary topic of speculation and gossip. An evaluator may have no vested interest in assessing the relative contributions of policies and people to the success or failure of a programme, but the evaluatees do. What evaluators call cause, evaluatees call credit or blame, with all the attendant connotations of pride or guilt, honour or shame. This is especially true when the people asked to provide the information for an evaluation are themselves the people being assessed.

Although evaluatees may hope to be credited with the successes of a programme, they also may fear being blamed for its failures. The reasons behind this apprehension are motivational and emotional. Almost everyone wants to maintain a positive self-image, and in most cultures, including Israeli culture, this image

includes the belief in one's own competence, honesty and sincerity (Bandura, 1986; Miller and Ross, 1975; Mullen and Riordan, 1988). Competent, honest, sincere people are expected to make good things happen; they are not expected to cause negative results such as a programme failure. Any attempt to evaluate a programme's successes and failures, and to assess the causes of each, will therefore stimulate the hope of credit and the fear of blame. If the evaluatees have reason to believe that the evaluation will link them to success or absolve them from failure, they will probably not feel evaluation anxiety and will probably not resist it. But if the evaluatees believe the evaluation might pin the blame for failure on them, they will almost certainly feel anxiety and resist the evaluation.

By focusing on the assessment of the programme policies, the three case studies provided the evaluatees with neither the hope of credit for the programme successes nor the threat of blame for the programme failures. So, there was little reason for the evaluatees to protect their self-esteem by strategies of defensive attribution (Crocker and Major, 1989; Mullen and Riordan, 1988). Questionnaire and interview respondents were asked to provide their own evaluations of a programme and its policies as well as criteria for feasible, successful or unsuccessful implementation attempts. This focus offered a more distant and neutral area of discourse than the praise-and-blame alternative, and effectively asked each evaluatee to play the role of a co-evaluator, a common prescription in participatory evaluation research (e.g. Patton, 1997; Whitmore, 1998). Thus, the evaluator and the evaluatees could discuss a programme and its policies more as colleagues than as inquisitor and accused. Add to this the flattery of attention (Derber, 1979) paid by the evaluator to the evaluatee, and the stage was set for co-operative and collaborative discussion.

Discussion

The extraordinary, emergency-like situations that gave rise to the three policies and the educational programmes discussed above almost certainly contributed to a norm of co-operation often missing in projects born under more placid conditions. So too did the nascent character of the three programmes which constrained their evaluations to be formative rather than summative. The relative contributions of these two factors cannot be determined. Even if they could, the resulting knowledge would seem to have no direct practical implications. One cannot, after all, recreate the extraordinary emergency conditions hoping to reduce evaluation anxiety and resistance, nor can one necessarily abandon the hope that summative evaluations might be conducted in an environment of reduced anxiety and resistance.

Yet in all three case studies we are struck by the spontaneous nature of the collaboration between evaluators and stakeholders. Much of the literature on participatory and collaborative evaluation stresses the self-conscious nature of the activity (Whitmore, 1998). An evaluator who wants to undertake a participatory or collaborative evaluation is encouraged to make this fact known to the stakeholders and to invest time and effort in soliciting and developing their active involvement in the evaluation. In contrast, the three evaluations reported here

began with no special commitment to participatory or collaborative evaluation; the collaboration came less from the evaluator's goals than from the evaluatees' interest. The interest may be attributable to emergency and incipient conditions. Whatever the source of strong level of involvement, such interest suggests that evaluators not assume that collaboration is the result of their efforts alone.

There remains the possibility that some attenuation of anxiety and resistance was the result of the personal characteristics of the evaluators. Perhaps they were more socially skilled or less threatening people than the evaluators of projects encountering more anxiety and resistance. More likely, they assumed a different role than the one expected of a traditional evaluator. Though no evaluation of the evaluators was undertaken, it seems that some of their most important contributions emerged from their willingness and readiness to listen to evaluatees and to incorporate their suggestions in developing the interview schedules and questionnaires. In addition, the evaluators became conveyers of information, complaints and replies, carrying out the role of courier and mediator between the two stakeholders – the commissioning 'client' and the evaluatees. The evaluatees' feedback during the evaluation process affected policy-related decisions made by the client. In this respect the evaluation became an ongoing communication tool rather than a mere decision-making instrument.

Listening, feedback and communication are activities advocated in grounded theory and participatory research to promote collaboration between the researcher and the participant and ultimately reduce anxiety and resistance (e.g. Charmaz, 2000; Kemmis and McTaggart, 2000; Patton, 1997; Whitmore, 1998). It is, of course, beyond the traditional role of an evaluator to become a confessor, courier or advocate for evaluatees. Indeed, if s/he assumes such a role, the evaluator is open to the criticism of losing objectivity or neutrality. Yet there may be no other way for the evaluator to overcome the evaluatees' anxiety and resistance (Patton, 1997).

At least four practical suggestions come from the three case studies and their analyses. They point to ways an evaluator can construct an evaluation process to enhance collaboration between the evaluatees and the evaluator. The suggestions are not new – they are included in the literature of evaluation use and participatory evaluation, cited above, and elsewhere. But they now have the added support our three case studies supply.

1. Evaluations of nascent programmes and policies do not require full-fledged evaluation criteria in advance. Anxiety can be reduced, and important evaluation criteria can be articulated and refined, by early collaboration with the evaluatees.
2. More co-operation is likely to emerge when policy makers, administrators and front-line workers see themselves as having common goals.
3. Co-operation increases and anxiety declines when the goal of the evaluation is depicted as assessing policies rather than people, and when the evaluation gives all stakeholders, including the evaluatees, a voice in policy improvements.
4. Evaluatees will be more co-operative if they believe they will get something

in return, such as useful and frequent feedback, for their participation. Evaluators who serve as conveyors of information, mediators, supporters and advocates are more likely to induce feelings of reliance, interest and trust in the evaluatees.

We believe the four suggestions address the criticisms of evaluation, discussed by Rog and Fournier (1997) and Whitmore (1998), as being narrow and unrealistic by setting high criteria, by not answering practical and immediate interests, and by serving political games. Evaluation can be innovative and creative, as much of an art as a craft. Finding creative ways of developing evaluations that are collaborative rather than conflictive is a worthwhile challenge.

Notes

1. At the other end of the spectrum, the most resistant to evaluation are government functions. They are defined as a service that transcends organizations and policies. After all, while organizations have multiple policy objectives, policies have fewer allies in the policy arena and are easier to evaluate, isolate and eventually even terminate (Daniels, 1997; deLeon, 1978). Organizations are always at the core of the functions-to-programs spectrum and of any decision making. Thus, progress in the explanation of the nature of organizations and their mechanisms can shed light on the types of obstacles that stand in the way of evaluation. These, in turn, can determine the evaluation plan and control and can assist in creating or taking advantage of opportunity windows as they occur.
2. We prefer to use the term 'evaluatees' rather than 'stakeholders' for this particular population. While the evaluatees may also be stakeholders, a stakeholder in an evaluation process is not necessarily an evaluatee. In this context, stakeholders are regarded as individuals who have a 'stake' in a policy or program. The evaluatees may indeed be regarded as 'secondary stakeholders' having an interest in the program that they try out; yet, the primary stakeholders are the initiators of the policy, who in the cases investigated are also those who commissioned the evaluation. They are the 'clients' interested in knowing whether their policies are feasible and implementable and to what degree.
3. We thank the anonymous reviewer for comments on these issues.
4. According to the findings of the Bank of Israel, analysed by Flug and Kasir (1993, 1995) and Flug et al. (1992), only 35 percent of the immigrants were employed in their profession or in related professions.
5. Of the 174 scientist retrainees involved in the study, 60 percent found jobs as teachers and 98 percent of them continued teaching in the second and third year after the retraining (Geva-May, 1991a).
6. Eventually the immigrant scientist teachers were willing to teach in these regions and classes, and they reinforced science teaching in those areas (Geva-May, 1998).
7. Technion, Israel Institute of Science and Technology, Haifa, Israel.
8. In our study (Geva-May, 1991a) we found that professional adaptation and socialization were personality- and motivation-related, and age was unimportant.

References

- Alkin, M. C. and F. Ellett (1984) 'Evaluation Models' in *International Encyclopedia of Education*. New York: Pergamon.

- Bandura, A. (1986) 'Self-efficacy: Mechanism in Human Agency', *American Psychologist* 37(2): 122-47.
- Bardach, E. (1977) *The Implementation Game: What Happens After a Bill Becomes a Law*. Cambridge, MA: MIT Press.
- Behn, R. D. (1978) 'Closing a Government Facility', *Public Administration Review* 38(3): 332-8.
- Bernstein, I. and H. Freeman (1975) *Academic and Entrepreneurial Research*. New York: Russell Sage.
- Book of Laws* (1952) Jerusalem: Government Publication.
- Braybrook, D. and C. E. Lindblom (1970) *A Strategy of Decision, Evaluation, Policy Evaluation as a Social Process*. New York: Free Press.
- Brehm, S. and J. Brehm (1981) *Psychological Reactance: A Theory of Freedom and Control*. New York: Academic Press.
- Brewer, G. D. (1978) 'Termination: Hard Choices - Harder Questions', *Public Administration Review* 38(4): 338-44.
- Cameron, J. M. (1978) 'Ideology and Policy Termination: Restructuring California's Mental Health System', in J. May and W. Wildavsky (eds) *The Policy Cycle*, pp. 301-29. Beverly Hills, CA: Sage.
- Charmaz, K. (2000) 'Grounded Theory: Objectivist and Constructivist Methods', in N. Denzin and Y. Lincoln (eds) *Handbook of Qualitative Research*, 2nd edn, pp. 509-36. Thousand Oaks, CA: Sage.
- Chelimsky, E., ed. (1985) *Program Evaluation: Patterns and Directions*. Washington, DC: APSA.
- Chelimsky, E. (1986) 'What Have We Learnt about the Politics of Program Evaluation?', Plenary address to the American Evaluation Association conference, November: Kansas, Missouri.
- Cousins, J. B. and K. A. Leithwood (1986) 'Current Empirical Research on Evaluation Utilization', *Review of Educational Research* 56(3): 331-64.
- Crocker, J. and B. Major (1989) 'Social Stigma and Self-esteem: The Self-protective Properties of Stigma', *Psychological Review* 96(4): 608-30.
- Cronbach, L. J. (1963) 'Course Improvement through Evaluation', *Teachers' College Record* 64(4): 672-83.
- Cronbach, L. J., S. R. Ambron, S. M. Dornbusch, R. D. Hess, R. C. Hornik, C. Philips, D. E. Walker and S. Weiner (1980) *Toward Reform of Program Evaluation*. San Francisco, CA: Jossey-Bass.
- Daniels, M. (1997) *Terminating Public Programs: An American Political Paradox*. New York: M. E. Sharpe.
- deLeon, P. (1978) 'A Theory of Policy Termination', in J. May and A. Wildavsky (eds) *The Policy Cycle*, pp. 279-300. Beverly Hills, CA: Sage.
- deLeon, P. (1983) 'Policy Evaluation and Program Termination', *Policy Studies Review* 2(4): 631-47.
- Derber, C. (1979) *The Pursuit of Attention: Power and Individualism in Everyday Life*. Cambridge, MA: Schenkman Press.
- Diab, L. (1970) 'A Study of Intragroup and Intergroup Relations among Experimentally Produced Small Groups', *Genetic Psychology Monographs* 82(1): 49-82.
- Dowd, E., S. Hughes, L. Brockbank, D. Halpain, C. Siebel and P. Siebel (1988) 'Compliance-based and Defiance-based Intervention Strategies and Psychological Reactance in the Treatment of Free and Unfree Behavior', *Journal of Counselling Psychology* 35(4): 370-6.
- Eisner, E. W. (1979) *The Educational Imagination*. New York: Macmillan.

Geva-May and Thorngate: Reducing Anxiety and Resistance in Evaluations

- Flug, K. and N. (Kaliner) Kasir (1993) *The Absorption in the Labor Market of Immigrants from the CIS 97 in the Short Run*. Discussion Paper no 93.09. Jerusalem: Research Department, Bank of Israel.
- Flug, K. and N. (Kaliner) Kasir (1995) *Short-Run Absorption of the Ex-USSR Immigrants in Israel's Labor Market*. Discussion Paper no 95.07. Jerusalem: Research Department, Bank of Israel.
- Flug, K., N. (Kaliner) Kasir and O. Gur (1992) *The Absorption of Soviet Immigrants into the Labor Market from 1990 Onwards: Aspects of Occupational Substitution and Retention*. Discussion Paper No. 92.13. Jerusalem: Research Department, Bank of Israel.
- Geen, R. G. (1983) 'Evaluation Apprehension and the Social Facilitation/Inhibition of Learning', *Motivation and Emotion* 7(2): 203–12.
- Geva-May, I. (1991a) *Policy Implementation Evaluation of the Re-organization of the Israeli Education System during the Gulf War (the 'Flexible School' Policy)*, pp. 1–44. Jerusalem: Ministry of Education. In Hebrew.
- Geva-May, I. (1991b) *Teacher Retraining Courses for New Immigrant Teachers: An Evaluation Research*, pp. 1–68. Jerusalem: Ministry of Education. In Hebrew.
- Geva-May, I. (1991c) *Evaluation Report: Organizational and Pedagogical Parameters (School Re-Organization Policy) for the Absorption of New-immigrant Earners*, pp. 1–48. Jerusalem: Ministry of Education. In Hebrew.
- Geva-May, I. (1998) 'Policy Feasibility and Immigrant Absorption in a Mass Immigration Context: The Case of Scientist Retraining in Israel', *Policy Studies Review* 15(2): 226–64.
- Geva-May, I. (2001) 'When the Motto is "Till Death Do Us Part": The Conceptualization and the Craft of Policy Termination', *Journal of Public Administration* 24(3): 1–26.
- Geva-May, I. and L. Pal (1999) 'Good Fences Make Good Neighbours: Policy Evaluation and Policy Analysis', *Evaluation* 5(3): 259–77.
- Glass, G. V. (1969) *The Growth of Evaluation Methodology*, Research Paper 27. Mimeo. Boulder, CO: University of Colorado.
- Graham, S. and V. Folkes, eds (1990) *Attribution Theory: Applications to Achievement, Mental Health and Interpersonal Conflict*. Hillsdale, NJ: Lawrence Erlbaum Press.
- Guba, E. G. and Y. S. Lincoln (1981) *Effective Evaluation*. San Francisco, CA: Jossey-Bass.
- Heider, F. (1958) *The Psychology of Interpersonal Relations*. New York: Wiley & Sons.
- House, E. R. (1980) *Evaluating with Validity*. Beverly Hills, CA: Sage.
- Ingram, H. (1990) 'Implementation: A Review and Suggested Framework', in N. Lynn and A. Wildavsky (eds) *Public Administration: The State of the Discipline*, pp. 462–80. Chatham, NJ: Chatham House.
- Joint Committee on Standards for Educational Evaluation (1981) *Standards for Evaluation of Educational Programs, Projects and Materials*. New York: McGraw-Hill.
- Kaufman, H. (1976) *Are Government Organizations Immortal?* Washington, DC: Brookings.
- Kaufman, H. (1987) *Time, Chance and Organizations: Natural Selection in a Perilous Environment*. Chatham, NJ: Chatham House.
- Kemmis, S. and R. McTaggart (2000) 'Participatory Action Research', in N. Denzin and Y. Lincoln (eds) *Handbook of Qualitative Research*, 2nd edn, pp. 567–606. Thousand Oaks, CA: Sage.
- Leitenberg, H., ed. (1990) *Handbook of Social and Evaluation Anxiety*. New York: Plenum Press.
- Lynn, L. E., Jr (1987) *Managing Public Policy*. Boston, MA: Little, Brown and Co.
- McArthur, L. (1972) 'The How and What of Why: Some Determinants and Consequences of Causal Attribution', *Journal of Personality and Social Psychology* 22(2): 171–93.

Evaluation 9(2)

- Martinko, M., B. Weiner and R. Lord, eds (1995) *Attribution Theory: An Organizational Perspective*. Delray Beach, FL: St Lucia Press.
- May, J. and W. Wildavsky (1978) *The Policy Cycle*. Beverly Hills, CA: Sage.
- Mazmanian, D. and P. Sabatier (1989) *Implementation and Public Policy*. Lanham, MD: University Press of America.
- Miller, D. and M. Ross (1975) 'Self-serving Biases in the Attribution of Causality: Fact or Fiction?', *Psychological Bulletin* 82(2): 213–25.
- Ministry of Immigrant Absorption (1996) *Immigrant Absorption – Situation, Challenges and Goals*. Jerusalem: Ministry of Immigrant Absorption Publications.
- Mullen, B. and C. Riordan (1988) 'Self-serving Attributions for Performance in Natural Settings: A Meta-analytic Review', *Journal of Applied Social Psychology* 18(1): 3–22.
- Nevo, D. (1983) 'The Conceptualization of Educational Evaluation: An Analytical Review of Literature', *Review of Educational Research* 53(1): 117–28.
- Palumbo, D., ed. (1987) *The Politics of Program Evaluation*. Newbury Park, CA: Sage.
- Patton, M. Q. (1987) 'Evaluation's Political Inherency: Design and Use', in D. Palumbo (ed.) *The Politics of Program Evaluation*. Newbury Park, CA: Sage.
- Patton, M. Q. (1997) *Utilization-Focused Evaluation*, 3rd edn. London: Sage.
- Pressman, J. L. and A. B. Wildavsky, eds (1984) *Implementation*, 3rd edn. Berkeley, CA: UC Press.
- Rog, D. and D. Fournier, eds (1997) *Progress and Future Directions in Evaluation: Perspectives on Theory, Practice, and Methods* 76. San Francisco, CA: Jossey-Bass.
- Rossi, P. H., H. E. Freeman and S. R. Wright (1979) *Evaluation: A Systematic Approach*. Newbury Park, CA: Sage.
- Schachter, S. (1959) *The Psychology of Affiliation: Experimental Studies of the Sources of Gregariousness*. Stanford, CA: Stanford University Press.
- Scriven, M. (1967) 'The Methodology of Evaluation', in R. E. Stake (ed.) *AERA Monograph Series in Curriculum Evaluation*, 1. Chicago, IL: Rand McNally.
- Sherif, M. (1966) *In Common Predicament: Social Psychology of Intergroup Conflict and Cooperation*. Boston, MA: Houghton Mifflin.
- Stake, R. L. (1967) 'The Countenance of Educational Evaluation', *Teachers College Record* 6(8): 523–40.
- Stufflebeam, L. D. (1971) 'Decision-Management Strategies', in L. D. Stufflebeam, W. J. Foley, W. J. Gephart, E. G. Guba, R. L. Hammon, H. O. Merriman and M. M. Provus (eds) *Educational Evaluation and Decision Making*. Itasca, IL: Peacock Publishers, Inc.
- Stufflebeam, L. D. (1972) 'The Relevance of the CIPP Evaluation Model for Educational Accountability', *SRIS Quarterly* 5(1): 3–6.
- Stufflebeam, L. D. (1974) *Meta-Evaluation*, Occasional paper, 3. Kalamazoo, MI: Western Michigan University.
- Tyler, R. W. (1950) *Basic Principles of Curriculum and Instruction*. Chicago, IL: University of Chicago Press.
- Weiss, C. H. (1975) 'Evaluation Research in the Political Context', in E. Struening and M. Guttentag (eds) *Handbook of Evaluation Research*, Vol. 1, pp. 13–26. Beverly Hills, CA: Sage.
- Weiss, C. H. (1983) 'The Stakeholder Approach to Evaluation: Origins and Promise', in A. S. Byrk (ed.) *Stakeholder Based Evaluation*, pp. 3–14. San Francisco, CA: Jossey-Bass.
- Weiss, C. H. (1985) 'The Many Meanings of Research Utilization', in E. Chelimsky (ed.) *Program Evaluation: Patterns and Directions*. Washington, DC: APSA.
- Weiss, C. H. (1998) *Evaluation*. Upper Saddle River, NJ: Prentice Hall.
- Whitmore, E., ed. (1998) *Understanding and Practicing Participatory Evaluation*, New Directions for Evaluation 80. San Francisco, CA: Jossey-Bass.

Geva-May and Thorngate: Reducing Anxiety and Resistance in Evaluations

- Wildavsky, A. (1979) *Speaking Truth to Power: The Art and Craft of Policy Analysis*. Boston, MA: Little, Brown & Co.
- Zajonc, R. B. (1965) 'Social Facilitation', *Science* 149(16 July): 269-74.

IRIS GEVA-MAY is Full Professor of Policy Studies at Simon Fraser University, Canada and at Haifa University, Haifa, Israel. She is editor-in-chief of the *Journal of Comparative Policy Analysis*. Her fields of study are policy analysis, policy evaluation and comparative policies. Please address correspondence to: Department of Political Science, Simon Fraser University, Vancouver, Burnaby, British Columbia V5A 1S6, Canada. [email: igevamay@sfu.ca]

WARREN THORNGATE is a professor of social psychology at Carleton University, Canada. His main research interests are in human decision making and group dynamics. Please address correspondence to: Psychology Department, Carleton University, 1125 Colonel By Drive, Ottawa, Ontario K1S 5B6, Canada. [email: warrant@ccs.carleton.ca]