



Edmond de Rothschild Foundation (Israel)

The Edmond de Rothschild Research Series

A collection of studies in the area of:

Measurement and Evaluation

2018

The Edmond de Rothschild Research Series

A collection of studies in the area of:

Measurement and Evaluation

2018

Dear Partners,

The Edmond de Rothschild Foundation (Israel) is spearheading philanthropic dedication to building an inclusive society by promoting excellence, diversity and leadership through higher education. Catalyzing true change and developing a cohesive society through dozens of innovative projects across the country, the Foundation provides growth and empowerment opportunities to the many communities in Israel. We develop and support novel solutions and creative partnerships, while evaluating result-driven programs with true social impact.

In keeping with its philosophy of strategic philanthropy, the Foundation established the **Edmond de Rothschild Research Series**, to promote excellence in research and expand the knowledge in the Foundation's areas of interest. The booklet before you centers on **Measurement and Evaluation**, as part of the first research series which focused on three main areas:

- 1. Access to and Success in Higher Education:** As part of its efforts to reduce social gaps, the Foundation strives to insure access to and success in higher education for periphery populations. It supports programs aimed at improving access to higher education options through preparation and guidance, reducing academic student dropout rates, and translating graduates' education into commensurate employment.
- 2. Measurement and Evaluation:** The Foundation seeks to constantly enhance its social impact and therefore, emphasizes measurement and evaluation of the projects it supports according to predefined, coherent criteria. The Foundation encourages evaluation as a continuous process that follows the course of programs from the planning stage and throughout their operation, and promotes the development of systematic data collection and analysis tools for all program aspects.
- 3. Baron de Rothschild's ("Hanadiv's") Legacy:** The Foundation is the expression of the Rothschild family's long-term commitment to the pioneering spirit of Israel. Harnessing a philanthropic legacy of more than 130 years, we are investing in groundbreaking agents of change and fostering modern-day pioneering in Israel.

A call for proposals was sent out to Israel's higher education institutions; academic steering committees were established in all three areas; and a total of 13 research proposals were approved. The researchers created new knowledge and distributed it, between the years 2014-2016, through various academic and non-academic channels. With the conclusion of the project, we are presenting summaries of all the completed studies and their main findings. The full publications of these studies can be found on the Foundation's website: www.rcf.org.il.

We would like to thank all the researchers from Ben-Gurion University of the Negev, The Hebrew University of Jerusalem, Tel Aviv University, the University of Haifa, Bar-Ilan University, the Open University, Myers-JDC-Brookdale Institute, and the Center for Educational Technology, who participated in this research series.

Enjoy your reading,

Elli Booch

Director of Philanthropy

Vardit Gilor

Program Officer

Table of Contents:

The Impact of High School on Access to Higher Education

*Moshe Justman, Ben-Gurion University of the Negev and Ruppin Academic Center
Naomi Friedman-Sokuler, Bar-Ilan University*

8-13

How Decision-Makers Cope with Evaluation Results: Predicting Responses of School Principals to Evaluations of their School's Climate

*Dr. Noga Sverdlik, Ben-Gurion University of the Negev
Prof. Shaul Oreg, The Hebrew University of Jerusalem
Prof. Yair Berson, Bar-Ilan University*

14-18

The Effectiveness of Evaluation in the Social Arena in Israel: Perceptions and Uses of Evaluation Research by Agencies Funding Social Programs in Israel

*Dr. Lior Rosenberg, Tirzah Margolin, Anat Kedem, and Dr. Tali Freund,
Center for Educational Technology*

20-24

Shared Measurement in Social Programs

Prof. Jack Habib and Yehonatan Almog, Myers-JDC-Brookdale Institute, Jerusalem

26-32

The Impact of High School on Access to Higher Education

Moshe Justman, Ben-Gurion University of the Negev and Ruppin Academic Center
Naomi Friedman-Sokuler, Bar-Ilan University

Purpose and contribution of the study

Israeli society is characterized by patterns of economic inequality in three main axes: gender, ethnic/religious group, and socio-economic status. In Israel, education—and especially access to higher education—is crucial for successful integration into the workforce. Moreover, equality of opportunity in studies towards a matriculation certificate may reduce socio-economic gaps in the labor market and should therefore be a central criterion in assessing local education systems and schools.

To assess schools' influence on their students' achievement and patterns of choice, it is necessary to isolate it and to differentiate it from background factors not controlled by educators or policy makers, such as the differences in initial conditions among the various population groups and the influence of the home and environment outside the school. The result is an undervaluation of the effectiveness of schools serving disadvantaged populations and over-evaluation of schools serving populations from a strong socioeconomic background.

A positive and significant correlation is evident in Israel between student socioeconomic background and matriculation achievement. The assessment of education institutions and educational systems according to matriculation eligibility rates (or the quality of matriculation as measured by the grades) ignores significant differences between the population of students attending various schools.

The added-value measurement method, the application of which is growing steadily in recent years in various measurements, successfully copes with the difficulty of isolating the effects of educational interventions on patterns of social mobility and widening / narrowing disparities of different types, by measuring the contribution of a curriculum or school to student progress. The main characteristic of this approach is the use of early student achievement as a basis for measurement, to offset the effect of background variables. The current study implements this approach by using the results of the MEITZAV standardized tests as a measure of the starting point, and the results of the matriculation exams as a measure of success of the intervention, using parametric and semi-parametric methods. This analysis enables us to identify the overall effect of the secondary school system on achievement gaps between different population groups.

The MEITZAV tests examine student command of content and skills according to the curriculum and Ministry of Education policy in four areas: mathematics, science, mother-

tongue, and English language. As such, the tests provide a uniform ranking of all Israeli students on a single scale in each subject. Matriculation is a multidimensional index that includes both compulsory and elective components, with variation in the level of the test in each discipline.

This study examines the impact of high schools in Israel on equality of opportunity in access to higher education, as reflected in achievements in the matriculation examinations – eligibility, subjects, and grades – in general, and particularly in scientific-technological (STEM) subjects. We focused on equality of opportunity in three main dimensions: between men and women, between Jews and Arabs, and according to the student's socioeconomic background, as reflected in parents' education and income.

The study provides a comprehensive description of gender and ethnic gaps in scholastic achievement in the MEITZAV tests in eighth grade, in the choice of STEM matriculation electives, and in matriculation achievement at the end of high school, and links them longitudinally to answer the following questions:

1. What is the extent of the gap between eighth-grade students in Hebrew-language and Arabic-language schools, beyond that which can be explained in terms of the differences among parental socioeconomic characteristics?
2. What are the differences between students in Hebrew-language and Arabic-language schools, as reflected in matriculation exams? To what extent can these gaps be attributed to gaps in achievements that were opened at earlier stages, which are evident in the eighth-grade MEITZAV tests?
3. What happens to gaps between population groups during high school?
4. In what way are the patterns of gender achievement and tracking dependent on the ethnic-religious affiliation of students? To what extent do gaps in eighth-grade achievement explain gender stereotyping in the choice of scientific-mathematic subjects in matriculation?
5. Does income and parental education affect matriculation achievement, after controlling for past achievements that represent parental investment in earlier stages of education? Is the effect of parental education and income similar for boys and girls? Is the impact of parental education dependent on the parent's or child's gender?
6. Can controlling for school variables provide an explanation for the existence of ethnic, gender, and socioeconomic gaps? Is there a difference in the influence of schools on student choices for matriculation, after controlling for individual and school characteristics?

The research method

The database specifically built for this study, which combines eighth-grade MEITZAV test data with twelfth-grade matriculation data, allows us to examine how secondary school influences these patterns of participation and differentiation, both in terms of patterns of participation in prestigious matriculation exams subjects and in terms of student achievement. The database contains data on eighth graders in Israel that took the MEITZAV tests in 2002 and 2003. During these two years, all schools in Israel (excluding independent education and special-education schools) participated in four MEITZAV tests: mother tongue, mathematics, English, and science. The study population includes 78,071 students who took the Meitzav test, 21.1% of them in the Arab education sector. For each student, the results of the tests, along with geographic and socioeconomic data, and detailed matriculation exam results (main subjects and grades) were available to us.

This database enables us to isolate the contribution of high school by controlling for eighth-grade MEITZAV test grades, reflecting the level of each student upon entrance to high school, because of school and parental influence up to this stage. It also enables us to control for characteristics of parents, whose influence on student achievements continues into high school.

The research population was parsed according to religious affiliation (as defined by the Ministry of the Interior) and sectoral affiliation of the specific educational sector to which the student's school belongs. With few exceptions, Jews and non-Arab Christians attend Hebrew-language schools, while Muslims, Christian Arabs, Druze, and Bedouins attend Arabic-language schools.

The percentage of girls in each of the groups, in each type of school, ranged from 48% to 50%. The exceptions are Druze schools and Christians in Jewish education, where girls account for 13% and 30% of students, respectively.

The research population's socioeconomic characteristics according to a religious-ethnic division show significant differences between the groups. One of the most striking differences between Jewish and Arab education sectors relates to the distribution of income: In Arab education, more than half of the students are in the two lowest income quintiles, whereas in the Jewish education, more than half of the are in the two highest income quintiles.

Main findings and their significance

The starting point for assessing the contribution of the secondary education system to accessibility to higher education is the MEITZAV test in the eighth grade. We use these MEITZAV test scores to isolate the influence of the high school on student achievements and on gender, sectoral, and socioeconomic differences. At this stage, significant differences in the levels of achievement are already apparent between the various population groups – differences that can be traced back to early educational experiences (prior to eighth grade) and even prior to entering the education system. The observed patterns correspond to patterns of socioeconomic differences.

Already in the eighth grade, there are achievement gaps, with higher achievements in Hebrew-language schools and among students of higher socioeconomic status as indicated

by parents' education or income levels. These gaps are more significant among boys. In Hebrew-language schools, student achievement is highest in single-sex state-religious schools and lowest in coeducational state-religious schools. In Arabic-language schools, Arab-Christian students exhibit the highest achievements, and Bedouin students the lowest. The achievement gap between the language sectors, favoring Hebrew-language schools, is largest in the language subjects, especially in English (Arab students are required to master three languages – Arabic, Hebrew, and English – but are not tested in Hebrew, while Jewish students study only Hebrew and English). The smallest gap favoring Hebrew-language schools is in science and technology, where the relative status of the Arab students is strongest.

Achievement gaps in favor of girls are evident in all subjects, both in Hebrew- and Arabic-language schools. The gap in language skills is greater than the gap in mathematics, and the gap in the Arab sector is considerably greater than the gap in the Jewish sector. This last point is salient particularly in sciences and mathematics: while gender differences favoring girls are negligible in Hebrew-language schools, they are large and significant in Arabic-language schools.

A significant portion of the gaps between the various groups stems from the gaps in family income and parental education between the groups. After controlling for students' socioeconomic background, in Hebrew-language schools, gaps between the state, state-religious single-sex, and state-religious co-educational schools narrow to non-significance, with the exception of English, in which gaps remain significant. In mother tongue and in English, gaps between Arabic- and Hebrew-language schools are reduced by more than half, and in other subjects, lose their statistical significance.

At the end of high school, it is can generally be said that the achievements of students in the Hebrew-language schools are higher than those of students in Arabic-language schools, and that the achievements of girls are higher than those of boys in each parameter and in each population group, except for average matriculation exam scores among Christian Arabs. However, gender gaps are not uniform across sectors: the female advantage is considerably greater in Arabic-language schools.

Matriculation eligibility rates, and specifically, rates of matriculation eligibility that meets university entrance requirements, are significantly higher in the state-education and single-gender religious schools than in other groups, and the gaps in favor of girls are relatively small. In the other groups, boys' eligibility rates are significantly lower and the gap in favor of girls is twice as large. The matriculation eligibility rates of Bedouin students are conspicuously low, among both boys and girls. Beyond simply attaining a matriculation certificate, access to higher education depends, to a great extent, on its quality. In this respect, girls have an advantage in all groups except for Christian Arabs, as measured by the average scores among those eligible for matriculation.

Achievement gaps in mathematics between the Hebrew- and Arabic-language schools are reduced during high school, while achievement gaps in English expand. Socioeconomic background continues to contribute to students' success even during high school, beyond its contribution in previous stages.

Patterns of gender-related choice and tracking in scientific subjects at the end of high school differ considerably between Hebrew- and Arabic-language schools. In Hebrew-language

schools, boys choose advanced physics, computer science, and mathematics electives at higher rates than girls, while girls choose biology and chemistry at higher rates. This gap does not change even when conditioning on prior achievement and socioeconomic characteristics: the low rates of girls choosing physics, computer science, and mathematics cannot be attributed to differences in mathematic abilities.

Girls in Arabic-language schools choose advanced physics, computer science, and mathematics elective at similar rates as girls in Hebrew-language schools; and in biology, at considerably higher rates. Boys in the Arab education system choose scientific subjects at much lower rates than their counterparts in Hebrew-language schools. However, when choice is conditioned on social background and previous academic achievement, Arab girls choose advanced science and mathematics matriculation electives at substantially higher rates than girls in Hebrew-language schools and the gaps favoring boys in physics and computer science are significantly larger in Hebrew-language schools. This finding is somewhat surprising given that Arab society in Israel is characterized by greater gender inequality than the Jewish majority group by any socioeconomic criterion.

Based on these findings, we have formulated the following policy guidelines:

The disadvantage among boys in Arab education in the scientific-mathematical subjects is evident already in the eighth grade and requires intervention early in the educational pipeline. Advanced intervention is further needed to reduce socioeconomic gaps at an early stage.

This is not the case among girls in Arabic-language schools, who attain similar levels of achievement as boys in Hebrew-language schools, conditional on socioeconomic characteristics, and even choose to specialize in scientific subjects at a higher rate. However, these high rates are not translated into increased enrollment in engineering and information-technology fields in higher education. These findings point to the need to focus intervention among girls in Arab education on the transition from high school to higher education.

The under-representation of girls in Hebrew-language schools in pre-engineering courses is evident already in high school, and it shapes similar patterns in subsequent academic choices. This disadvantage does not reflect an early deficiency in mathematics, but rather, should be attributed to patterns directing girls to certain subjects. Moreover, girls who choose to study and matriculate in these subjects achieve, on average, higher grades than boys. These findings point to the need to encourage girls with relevant aptitudes to be exposed to these subjects in middle school, and to inform them about the high success rates of girls in these subjects.

A policy in this spirit may help these groups of students to realize their abilities in the education system and later in the labor market, thereby reducing economic and social gaps in society.

How Decision-Makers Cope with Evaluation Results: Predicting Responses of School Principals to Evaluations of their School's Climate

Dr. Noga Sverdlik, Ben-Gurion University of the Negev

Prof. Shaul Oreg, The Hebrew University of Jerusalem

Prof. Yair Berson, Bar-Ilan University

Purpose and contribution of the study

This study focuses on the responses of school principals to the annual assessment of the school (MEITZAV tests – school climate component), addressing two main questions: What is the relationship between principals' traits and their response to the assessment's findings; and how are these responses, along with the principal's leadership style, related to the likelihood that the principal will introduce changes in the school in response to the assessment. We developed and empirically tested a model to explain how decision makers use the accumulated information, focusing on leaders' dispositional inclinations, motivations, leadership style, and more.

This study is the first to address the psychological factors that promote or inhibit the effective use of feedback pertaining to schools' climate and, as such, may contribute to understanding how leaders use feedback to benefit their organizations.

To test our model, we focused on Israeli schools' responses to the national annual assessment of the school climate (the MEITZAV test, a Hebrew acronym that stands for Measure of Scholastic Efficiency and Growth). An integral part of the MEITZAV test is an assessment of the school climate and pedagogical environment. The test provides school principals and the Ministry of Education with a comprehensive depiction of the atmosphere in the school in general, and students' feelings in particular.

Consensus has yet to be reached with respect to the effectiveness of the MEITZAV tests, and the extent and manner to which principals use their results toward improving their schools has yet to be examined. Studies of evaluation tests in other countries (e.g., Brown, et al., 2011; Remesal, 2011) indicate that teachers' perceptions of tests affect their use of test results. We therefore took into account principals' perceptions of the assessment tests as one of the predictors of their responses.

Based on existing literature in the fields of education, organizational behavior, and psychology, we formulated the following hypotheses about school principals' responses to the findings of the school climate component of the MEITZAV tests:

1. Principals' personal values, perceptions of assessment tests, and style of coping with challenges and stress predict their cognitive, emotional, and behavioral responses to the MEITZAV results. Among personal values, our focus was on the importance principals ascribe to values of conformity, which are likely related to their acceptance of the MEITZAV tests as a legitimate tool for examining the school atmosphere.
2. Principals' responses to the MEITZAV findings predict the extent to which they will make use of the findings towards the improvement of the school climate.
3. Principals' leadership style, in particular, their transformational leadership (Bass, 1985), predict the strength of the relationship between principals' responses to MEITZAV findings and the degree to which they actively engage in dealing with the MEITZAV findings. Transformational leadership refers to a form of leadership whereby leaders inspire their followers, provide intellectual stimulation, and are responsive to their professional needs (Bass, Avolio, Jung, & Berson, 2003).

The research method

We distributed questionnaires among school principals and teachers, prior to the publication of the MEITZAV results, in which we asked about their values, attitudes towards the use of the MEITZAV, and their style of coping with stress. A second questionnaire, distributed to the principals immediately following the publication of the MEITZAV findings, included questions about their emotional, cognitive, and behavioral responses to the MEITZAV findings. In the third stage, approximately six months after the publication of the findings, we distributed questionnaires among school principals and teachers, in which we asked about the extent to which principals had initiated changes following the test results. MEITZAV school climate findings were retrieved from the Ministry of Education's website.

We collected data from 86 schools: 55 elementary state schools, 14 junior-high state schools, and 17 junior-high state-religious schools. In 44 of the schools, the MEITZAV test was conducted in the 2012-2013 school year, and in 42 schools – in 2013-2014. Each round of data collection included the principal and an average of six teachers. In total, we collected data from 1223 teachers (663 in the first stage and 560 in the second stage), and from 86 principals. Of the 86 principals, 65 principals provided data in the second data collection stage. Unless otherwise indicated, respondents were asked to indicate the extent to which they agree or disagree with the questionnaire items, with response options ranging from 1 (totally disagree) to 6 (totally agree).

Principals' personal values were measured using an abbreviated 34-item version of Schwartz's 5X-PVQ questionnaire (Schwartz et al., 2012). Scale items were statements about hypothetical individuals, and participants were asked to rate the extent to which they perceive themselves as similar to the characters described. Principals' perceptions of assessment tests were evaluated using an abbreviated 12-item version of the Conceptions of Assessment Inventory in Hebrew (Brown, 2006). Coping strategies were evaluated using the Hebrew-language version of the abbreviated COPE questionnaire (Carver, 1997), which includes 12 items about

various approaches to coping with a particular stress situation (questionnaire items were adjusted to relate to stress situations in general and item wording adapted to the context of the MEITZAV climate tests). Principals' emotional, cognitive, and behavioral responses to MEITZAV findings were assessed using questions based on Oreg's (2006) scale of attitude toward change. Given the key role of this variable in our research model, our analyses include data only from principals who filled out this scale.

To gauge changes introduced by the principals into the school in response to the MEITZAV findings, we constructed a scale specifically for the purposes of the present study. Scale items consisted of questions about the extent to which the principal introduced such changes. Principals' transformational leadership style was evaluated using teachers' responses to Bass and Avolio's Multifactor Leadership Questionnaire (1993), which includes 20 items.

MEITZAV school climate findings were obtained for the 2012-2013 and 2013-2014 tests (in accordance with the academic year during which our questionnaires were distributed in each school). To gauge the school atmosphere, we focused on pupils' reports of school violence, averaging the level of verbal and other types of violence in the five relevant grades of each school (grades 5-9).

Main findings and their significance

We focused on three personal characteristics of the principals: conformity values, perceptions of the role of assessment tests, and style of coping with stress as predictors of their emotional, cognitive, and behavioral responses to the MEITZAV findings. Our analysis reveals the importance of distinguishing between the emotional, cognitive, and behavioral components of responses to the MEITZAV. We found that the personal attributes are differentially linked to the various attitude components, and that there are indirect relationships between principals' attributes and components of principals' position toward the MEITZAV. Specifically, principals' emotional response predicted their cognitive response, which, in turn, predicted their behavioral response.

An interaction between principals' conformity values and students' reports of violence in school predicted principal's emotional response to the findings. The higher in conformity values, the more negative was the principal's emotional response, particularly in schools with a higher level of violence. Moreover, as expected, the level of school violence also had a direct effect on principals' emotional responses – the higher the level of violence, the more negative was principals' reaction to the findings. In turn, principals' emotional response was positively associated with their cognitive response.

The cognitive response was also predicted by principals' perceptions of the role of assessment tests and their coping style. The more positive their perceptions of the role of the MEITZAV, the more positive was their cognitive response to the findings. And the more a principal's coping style was oriented toward problem-solving, the lower was his/her cognitive response. This is surprising, as we would expect principals who deal with stress through problem-solving to have more positive attitudes towards the MEITZAV, given the feedback for improvement that it provides. A possible explanation of this finding is the distinction between the personal tendency to focus on action and the tendency to focus on the cognitive assessment of the situation and of the gap between it and the ideal situation (Higgins et al., 2013). Although combining both an evaluation of the situation and a problem-

solving orientation is optimal for promoting goals, research suggests that people tend to focus on one or the other. In the present context, it is possible that principals with a problem-solving coping style tend to focus on action and less on a cognitive appraisal of the situation. Contrarily, because behavioral responses to the MEITZAV findings indicate an action-oriented approach, a coping style of focusing on problem-solving is likely to be positively correlated with these responses. The more principals hold a problem-focused coping style, the greater their tendency to report their behavioral intentions of doing something about the findings. In addition, their behavioral response was also predicted by the cognitive response; the more positive the cognitive response, the more positive was their cognitive reaction.

In turn, principals' behavioral response interacted with their leadership style in predicting the initiation of changes in the school following the MEITZAV. Among principals with a transformational leadership style, the relationship between their behavioral response and introducing changes tended to be stronger. Furthermore, transformational leadership also had a direct effect on principals' reports of their initiation of changes.

Overall, our findings present a complex depiction of the role that the MEITZAV response components (emotional, cognitive, behavioral) have in linking principals' personal attributes and their initiation of changes in their schools. Principals' conformity values (together with the students' reports of school violence) predicted principals' emotional responses to the MEITZAV findings. A problem-solving coping style predicted both the cognitive and behavioral response. The findings demonstrate the importance of the emotional and cognitive components in generating the behavioral response, which is the only response component directly linked with the initiation of changes in the schools. As expected, principals' perceptions of the MEITZAV were associated with their cognitive assessment of their school results. Furthermore, in line with previous findings, transformational leadership was positively associated with the initiation of changes. Specifically, principals' behavioral responses were linked with the initiation of changes only when the principal's leadership style is transformational.

The findings have both theoretical and practical implications. With respect to theory, the findings demonstrate the unique role of each component of the principals' response to assessment findings, and the interaction among these components in driving principals' actions. Similarly, the findings demonstrate how different individual traits (values, attitudes, and coping styles) are linked with different aspects of decision-makers' responses to feedback. The findings also bolster previous findings about the role of transformational leadership in initiating changes in organizations.

With respect to practical implications, the findings shed light on principals' use of the information obtained through external evaluations towards the initiation of changes for the improvement of their organizations. The findings call attention to those personal attributes that can predict principals' responses. The evidence for the roles of principals' values, attitudes, and coping styles in leaders' use of organizational feedback suggests different work practices that different leaders may wish to adopt for optimizing their use of feedback. For example, the findings suggest that to stimulate positive behavioral responses, principals may be encouraged to undertake a problem-focused coping style rather than focusing on the emotions that often emerge in response to the evaluation findings. Similarly, to encourage positive cognitive responses, education administrators should focus on increasing principals' trust in the assessment tests. Moreover, the findings indicate that the key component in

principals' responses is the behavioral intention. It is therefore important to consider not only what principals think or feel about the findings, but also to trigger their intentions to act upon the findings and initiate changes toward the improvement of their schools.

In addition, the findings about the positive implications of transformational leadership on the initiation of changes following the MEITZAV highlight the value that training principals toward the greater manifestation of this style of leadership may have. Previous studies of transformational leadership (Dvir, Eden, Avolio and Shamir, 2002) show that such training is possible and furthermore, leads to positive results both for the individual and the organization.

The Effectiveness of Evaluation in the Social Arena in Israel: Perceptions and Uses of Evaluation Research by Agencies Funding Social Programs in Israel

Dr. Lior Rosenberg, Tirzah Margolin, Anat Kedem, and Dr. Tali Freund, Center for Educational Technology

Purpose and contribution of the study

Over the past two decades, measurement and evaluation have become an integral part of the activities of organizations operating in the social field. Studies in the United States point to the expansion of measurement and evaluation practices and an increasing recognition of the need to examine the outcomes, performance, and effectiveness of the activities of governmental and non-governmental organizations, philanthropic organizations, and foundations. A survey of more than 1,000 NGOs in the United States found that 85% of them included some aspect of evaluation in their work. In a similar study conducted in Canada among funders (both from the non-profit and the public sectors), 77% reported an assessment conducted in the past year. For most (64%) of the organizations that evaluated their activity, this was not a one-time measurement; most of them (73%) reported conducting an assessment as part of their work routine, and a similar percentage noted that the decision to conduct evaluation was an internal decision of the organization and not because of pressure from the funding bodies. In Israel, too, the ability to assess the extent to which the funding goals are achieved and efficiency of use is employed by funding entities in the public sector (government ministries, central authorities, local authorities, etc.), as well as NGOs (mainly funds and federations).

However, it was found that despite the high rate of non-profit sector organizations that do conduct some kind of assessment, the rate of professional evaluation is still low. For example, among non-profit organizations in the United States, it was found that although 85% measure some aspects of their work, only 21% employed professional evaluators and only 13% employed at least one full-time evaluation worker. The evaluation was defined as the area of responsibility of managers or the board of directors of the organizations, rather than of evaluation professionals.

This research examines the manner and scope of influence of evaluation studies on the decisions of funding agencies in the social field in Israel. Four central questions form the basis of the research:

1. To what extent is evaluation data used by funding agencies in making their decisions and shaping their strategies about intervention programs in the social arena?
2. Can the various types of funding agencies be characterized by the manner and scope of their use of evaluation studies?
3. Is it possible to characterize aspects of the evaluation process that influence the manner and scope to which they would be used by funding agencies?
4. What characterizes the products of evaluation studies that are more widely used by the funding agencies?

The perceptions and uses of evaluation findings among funding agencies in Israel has so far been the subject of very little research. This report seeks to contribute to the expansion of knowledge on this subject.

The study both sheds light on the uses of funding by evaluators and examines local authorities' evaluation. To the best of our knowledge, this is the first study in Israel that systematically deals with these issues.

The research method

The research population consisted of representatives of government ministries (including governmental authorities and services), local authorities, and NGOs (foundations and federations) that fund social initiatives in Israel. The study was conducted in two stages. In the first phase, 26 in-depth interviews were conducted between January and May 2013. The findings of the interviews provided an infrastructure for the development of an online survey questionnaire used in the second stage of the study. The questionnaire was delivered to 250 representatives of funding organizations between July and September 2013. In total, 79 respondents – of which an absolute majority (98%) were at the highest management levels – provided feedback. Of those, 16 (20%) were government representatives; 25 (32%) were representatives of local authorities; 11 (14%) were representatives of federations, and 27 (34%) were fund representatives.

Main findings and their significance

Most of the survey respondents report having initiated evaluation studies. Their high rate among the research population may indicate that the response to the survey was greater among those who favor evaluation and may accordingly be viewed as an over-estimate. Their responses indicate that government ministries and local authorities tend to employ staff in the field of evaluation, whereas for NGOs, the tendency is to use external evaluation services. The findings show that funding agencies that use evaluation studies are highly involved in most stages of the study, but their involvement is relatively limited in designing the call for research and in the stage of consolidating its conclusions and recommendations. The findings indicate that funders who are more involved in the evaluation process also tend to use evaluation studies to examine a larger scope of topics and make greater use of their findings and conclusions.

Findings indicate that funding agencies perceive evaluation as useful and beneficial mainly for monitoring the implementation of programs and for measuring its outcomes and impact. A large proportion of the funders believe that evaluation studies are an effective means of leading to improvement, and to a significantly lesser extent perceive them as contributing to the expansion of partnerships and mobilization of resources. Evaluation studies serve the funders mainly for the purpose of learning and improving the evaluated programs, and less to influence policy-making authorities or as a means of public relations or for raising funds. They also use them less to examine alternatives to the actual program, or to examine the extent and continuation of investment. It appears that funding organizations hold mainly “traditional” perceptions of evaluation studies and see them less as a resource in strategic processes and in making key decisions about the program.

The funders perceive two main evaluation weaknesses:

1. The evaluation reports do not present a comparison of the program to similar programs in Israel or around the world and do not give reference points.
2. The reports contain ambiguous conclusions and they lack clear recommendations.

In addition to these two main weaknesses, a quarter or more of the survey respondents mentioned other weaknesses: The timing of the research findings is not synchronized with the schedule of decision regarding the future of the program; research reports are too long and contain too much statistical data; and they are not accompanied by a useful summary.

As for the development of the evaluation field, most of the funders thought it would be useful to establish a website providing comprehensive information about evaluation studies in Israel, providing training to members of the organization and/or forming professional communities of similar organizations on the use of evaluation studies.

When non-traditional evaluation services were offered to the funders, a high rate of interest was indicated for ROI studies, facilitation services to promote the effective use of evaluation research and their findings, and in studies that will measure the impact of the intervention beyond the target population and specific program objectives. These findings suggest that there is a great potential to develop the field of evaluation in the social arena to a more “strategic” and systemic level.

The perceptions and patterns of use of evaluation, as reported by funders from the three sectors, indicate a difference between them, which stems, in our opinion, from their position in the social field – the government, which dictates policy and allocates resources and is responsible for supervision and accountability; local authorities, who are dependent on partnerships and the mobilization of resources for the operation of services to residents; and foundations and federations, who provide support to program-operating organizations and are obligated to report to their donors.

In addition to the differences between the sectors, the study highlights common barriers that impair evaluation use, including a lack of budget and the desire to invest resources directly on programs’ beneficiaries and the inherent tension between the funding body and the program operator, which is enhanced by the evaluation process.

The study points to a widespread agreement among the funding agencies that evaluation studies are an effective means of generating improvement processes, and that most of the study participants do use them. However, the prevailing concepts and uses are “traditional”

and express an approach according to which assessment is primarily concerned with the study of implementation and measurement of outputs and outcomes, and it serves less as a “strategic resource” for shaping policies and decisions. These findings point to the limited impact of evaluation studies on the social field.

The challenge of making evaluation studies more meaningful for funding and strengthening their contribution to social activity in Israel emerges as a central issue for discussion. The finding that funders are interested in more “strategic” assessment services, when offered directly to them, indicates a potential that has not yet been actualized.

We believe that the concept of evaluation as a strategic process of change is the main leverage in enhancing the usability and impact of evaluation studies in the social field. To promote this concept, we believe it is necessary to:

1. **Strengthen the evaluator-funder dialogue:** A deep and ongoing dialogue is expected to promote familiarity with the needs and limitations of each of the partners and serve as a catalyst for change in the perceptions and implementation of the evaluation process. Such a dialogue will contribute to the relevance of the evaluation study and its potential for change. Structuring the dialogue into the evaluation research program will ensure its sustainability and serve as an essential mechanism for initiating and managing the change. Both the funder and the evaluator share responsibility to strengthen the dialogue.
2. **Create platforms for management of change:** The responsibility for managing the change is the funders', and evaluators should assist them in this task. It is very important that the evaluator and funder will discuss this aspect of the evaluation work and will design management mechanisms to enable and support the process of change from the first step.
3. **Change role perception of the evaluator:** The idea of the evaluating agencies' work continuing beyond submission of the findings, into assisting in management of findings-derived change, necessitates introducing a change in the perception of the evaluator's role. This change requires the evaluator to develop a new understanding of its role, encompassing dialogue with the funder and acquiring new knowledge and skills, beside research capabilities. It may also involve the work of a multidisciplinary evaluation team.
4. **Expand the knowledge of those ordering the evaluation and create a common language between them and the evaluator:** Without the funding organizations' thorough understanding of the potential and options of the evaluation process, it cannot become a tool for change. The evaluation is a strategic management resource, and funders need to know how to utilize and leverage it to advance the organization's goals – and work to expand this knowledge. At the same time, it is important that the evaluator present to the funder in as much detail as possible the available services and assist the funder in identifying the specific services that are relevant and useful. Moreover, the evaluator must also develop a language that will promote clarity and understanding even among those not versed in professional jargon.

5. **Use comparative data and tools already successfully tested:** The evaluator is expected to be familiar with the field, should strive to implement the tools / methodologies / methods of analysis that have been successfully tested in previous studies, and should refer to comparative data available in open resources. These resources will enable a broader and comparative perspective and will advance learning from one program to another.
6. **Promote inter-organizational cooperation:** It is important to establish inter-organizational platforms that will enable sharing knowledge about programs, including insights from evaluation studies. Moreover, the research points to an emerging shift in thinking, from evaluating a program to evaluating a whole field or area. Such an assessment will examine all interventions undertaken in the same field with common goals or target audiences. The importance of such an evaluation is that it can provide a compass for all operators in the same field to test their impact. It seems that the social field is ready for such cooperation.

In light of the findings and their discussion, we recommend implementing the insights and ideas of this research, for example by publishing our report among broad audiences; conducting a seminar on the findings and insights from it; building a website that will inform the evaluation of funding institutions; establishing a team of experts that will propose an updated concept for the role of the evaluator and ways to disseminate it; as well as advancing the idea of evaluating an entire field through inter-organizational cooperation.

Some possible ideas for further research, which provide new insight about the role of evaluation and contribute to the further development of this field in Israel: Following up on the indicators that were examined in the present study to examine their change over time; creating systematic documentation of cases in which evaluation studies were used in the "learning from success" methodology; and developing methodologies for presenting questions and evaluation data in a strategic context and in managerial – rather than research – language.

Shared Measurement in Social Programs

*Prof. Jack Habib and Yehonatan Almog,
Myers-JDC-Brookdale Institute, Jerusalem*

The following is an abstracted and translated version of "A Handbook for Implementing Shared Measurement" by the authors. The original document (in Hebrew) can be found at: http://www.kshalem.org.il/uploads/pdf/article_6952_1390513936.pdf

Result-oriented thinking and measurement are increasingly becoming a part of work routines in social organizations, reflecting stakeholders' rising expectations of being able to examine the organizations' effectiveness and their evidence-based management practices. The importance of measuring results is further reinforced by increasing demands for transparency and accountability on the part of social services and organizations.

The evaluation processes in each organization are typically disconnected from similar processes in other organizations. This is a situation in which each organization defines its goals and methods of evaluation and conducts separate measurements. A different concept is shared measurement, in which measurement is conducted jointly by several organizations engaged in a defined social field and working towards similar or shared goals, or multiple units in a complex organization. Shared measurement includes a common definition of goals and metrics, agreement on the measurement process, data analysis, sharing the findings of measurement, and possibly a shared learning process, enabling organizations to strengthen aspects of mutual learning and to look together at the larger picture, beyond the work and influence of each organization.

When joint measurement is conducted by organizations that collaborate in additional aspects beyond the scope of measurement, it can be a major means of creating a common language, strengthening and maintaining cooperation and effective targeting of coordinated action. As a result, in the emerging literature on collective impact initiatives, joint measurement is defined as one of the main levers for promoting and strengthening effective cooperation.

Despite these developments, so far, Israel has little experience in the development and implementation of these practices, and this review is one of the first attempts to promote the conceptual and practical aspects of this field. The review deals with the development of concepts, the main benefits and challenges posed by each of them, and the description of several prominent examples from the world.

A distinction should be made between the ability to indicate an improvement in the measured results and the ability to demonstrate the extent of the program's influence on the improvement in the results. Demonstrating the impact of a program requires a measurement system that allows attribution of the improvement in the results obtained to the intervention

program. In addition, when measuring partnerships, it is necessary to distinguish between the ability to indicate the results of the joint activity and the ability to demonstrate the added value of the partnership in relation to the situation in which each organization would operate separately.

Along with the many benefits of measuring results, it also involves challenges and risks, including – beyond the risk of measurement, analysis, and interpretation errors – also unrealistic expectations from stakeholders, a focus on what is easy to measure rather than what is important, and the simple fact that not all results can be measured quantitatively and at a reasonable cost.

Two meta-strategies can be distinguished for measuring results: evaluation and ongoing measurement/monitoring. Evaluation is concerned with measuring results during a limited time, usually by an external body. Ongoing measurement is the measurement and regular follow-up of results, usually performed by a team from within the organization. The two strategies vary in scope and depth of the measurement, among others. Deciding how to measure should reflect an overall information strategy for the evaluated program, which takes into account needs and resources. An overall information strategy enables the development of a results measurement system that best addresses the program's needs and maximizes the use of existing resources.

In deciding what to measure, it should be noted that for organizations that do similar things to achieve similar goals, it is relevant to measure outputs, intermediate results, and final results. For organizations doing different things to achieve similar goals, it is mainly relevant to measure final results; and among organizations dealing with different aspects of a problem to achieve a common over-arching goal, it is mainly relevant to measure the extent to which that goal was achieved, and the contribution of various sub-goals to its achievement.

Steps in the design and implementation of shared measurement

A preliminary stage for the planning and implementation of joint measurement is the foundation-laying stage. The essential starting point for this stage is the various organizations' recognition of the added value of shared measurement as a means to improve their practices and achievements. Laying the foundations starts with the establishment of a joint steering committee as a mechanism ensuring cooperation. Once established, the steering committee should continue developing a shared understanding of the goals of the process and of its value to each organization separately and to the partnership as a whole. In addition, the committee should also arrive at an agreement with regards to the level of inter-organizational cooperation.

Agreement should also be reached with regards to the principles and method of working together, including a division of responsibilities and roles, mechanisms of responsibility, financing strategy, implementation strategy, and, importantly, selecting an entity to integrate all joint measurement activity. This entity will consult in planning and will play a role in measurement implementation, data analysis, and application of the data according to the selected strategy.

An important infrastructure for any shared measurement is deciding on a set of common goals. This process is often challenging, as it embodies a move from a general discourse of cooperation to a discussion of the concrete desired outcomes. While this more concrete discussion ultimately leads to deeper, more substantive agreements between the partners, it may also sharpen differences. A number of steps may be taken to enhance the chances of forming a broad consensus: to set common goals only after the partners are acquainted, have shared some learning, and built mutual trust; to devote sufficient time to the process, enabling inclusive and participatory discourse; to include funding bodies in the discussion as a means of encouraging partners to reach an agreement; and to enlist external experts (e.g., organizational consultants and group facilitators) in planning the process and in managing the discourse.

Next, a measurement plan should be formulated, defining what will be measured – and how. Measurement can relate to inputs, outputs, and outcomes. The measurement of inputs and outputs will enable the formulation of a broad picture of the activity, helping to understand the results and devise improvement processes. But the main issue is which results should be measured. The partners must decide on a limited number of interim and final results they intend to measure. One of the dilemmas at this stage is whether to include the quality of inter-organizational cooperation in the measurement of interim results – a measurement that may contribute not only to understanding common outcomes but also to strengthening the partnership.

A result indicator should be adapted for each measured outcome, defining the outcome operationally and enabling to evaluate the extent to which it was achieved (for example, a measure of the outcome of integration into work can be the percentage of placement of unemployed workers). In selecting the indicators, organizations can decide to adopt standardized common metrics from the literature or develop their own metrics for use in a given context.

To complete the measurement program, the source of information, method of data collection, and timing and frequency of data collection should be defined for each indicator. Additional details can be added as needed. These definitions will be reflected in the development of the appropriate measurement tools. Each indicator can also have a numerical value expressing the desired level or expected change during a set period. If the partners decide to set a goal, it can be shared by all organizations (partnership level) or differential (organization level).

The two main data collection strategies in a shared measurement framework are:

1. Developing a central shared information system managed by the integrating entity;
2. Using separate systems – each partner collects the information and feeds it into its information system using shared measuring tools, methodologies, and protocols. These strategies differ in the depth of the partnership and in the role of the integrating entity versus the role of the partner organizations in data collection. Each strategy allows for the choice between two alternatives:
 - 1a. The integrating entity collects all the information for the partners and feeds it into a central system.
 - 1b. Each of the partners gathers the information and feeds the data into a central system and can use the system to analyze its own data, and sometimes also to examine data of the other partners.

- 2a. The information is transferred to the integrating entity, for integration and analysis.
- 2b. Each of the partners analyzes the information and reports its findings to the other partners.

When a more integrative analysis is desired, a shared central information system should be developed (both options under 1) and if strengthening the independent measurement capabilities of the partner organizations is also a desired outcome, then 1b is the preferable option. For partners with good measurement capabilities who do not attach the utmost importance to maximal integrative analysis, either option under 2 may be appropriate, and if the organizations wish to maintain greater independence, with minimal integration of the findings, then option 2b is preferable. To address the challenge of uniformity in data collection and analysis, if the partner organizations opt for greater independence in data collection and analysis, appropriate quality control and assurance mechanisms should be installed.

Once the strategy is formulated, a concrete plan must be prepared for its implementation. The program should address the preparation of measurement tools, information gathering protocols, the information systems, and the practices of collecting information and feeding it into the system. The plan will be based on principles for joint work, including division of responsibilities and roles, formulation of responsibility mechanisms, financing strategy, and clearly defined timetables. Attention should also be given to training and consultation, and to providing clear definition of the relevant processes. In addition, the control processes and reporting procedures (to the steering committee) of issues, dilemmas, and difficulties arising in the measurement should be defined. It is highly advisable to set a pilot stage, in which the measurement system will be calibrated.

Using the collected data and applying the findings are key to achieving the goals of the measurement process. Thus, the stage of data analysis and publication of the findings requires careful attention and preparation. Effective information sharing is based primarily on the extent to which effective cooperation had been built in the previous stages, but also on additional issues such as the differences, among the various partners, in experience and previous knowledge in applying the findings. For this reason, a prior investment in preparing organizations for information use is required, including, for example, joint training for the partners to create a more uniform starting point and a common language. It is also important to work towards creating a supportive climate, oriented towards learning and improvement and bringing to light the complexity and limitations of a comparison between partners, while avoiding hasty or erroneous conclusions.

To answer the predefined main questions and identify important issues arising from the data, the data should be analyzed according to an agreed format. Presenting the information clearly and concisely is necessary for its effective use. The answers to questions emerging from the findings are generally not conclusive – because of this, the partners' discussion of the findings and their interpretation is very important. The discussion may highlight the need to perform additional processing in order to clarify certain findings and bring them to further discussion, or even raise questions about the reliability of some of the data – a situation that requires further testing and analysis.

The implications of the findings for planning the continuation of the activity and for drawing conclusions and making decisions (shared by all partner organizations or independently)

on ways to introduce improvements must also be discussed in depth. This step is essential for creating a continuous, data-based improvement process. The type of data analyzed and discussed changes as the activity progresses, from information relating to participants' population and their situation (prior to the activity); through information about the implementation; and, in more advanced stages, information about the emerging results.

At every stage, it is significant to conduct a joint discussion of how the activity should be advanced, while creating the processes necessary to ensure the maximum shared outlook on the findings and their implications. It is also highly important to document how the information contributed to advancing the activity at each stage.

Reporting to stakeholders and to the public: Beyond the findings as a basis for improvement processes, it is also a tool for maintaining the support of funders and backers. Joint reporting of the findings is related to the information collection system; a central information system most likely requires a joint report, along with independent reporting by each organization to its unique audiences, taking into account these audiences' different reporting needs.

The following are important points for the success of the measurement process:

Ensuring appropriate organizational infrastructure

Although joint measurement may be more effective than a separate measurement, it still requires considerable resources to develop and apply it on an ongoing basis. It also involves special costs related to the cooperative aspects of the process. To ensure its successful implementation, multi-year funding must be secured in realistic terms. The fact that funding is distributed among several organizations and the possibility of raising funds from third parties, because of the desire to promote joint and broad processes, can facilitate the guarantee of financing.

Shared measurement, in which each partner organization takes upon itself some of the responsibilities and tasks required, performing them as part of its ongoing activity is one viable option. However, an integrating entity with a dedicated team significantly enhances the partnership's potential to succeed. This body deals both with the measurement itself and with the management and leadership of inter-organizational discourse and coordination.

When implementing a shared measurement, suitable training for members of the various organizations will usually be required – even when data collection and analysis are the responsibility of an integrating organization, in such topics as information systems, reports, and use. Therefore, an appropriate training and capacity-building infrastructure should be created. Such infrastructure, which is particularly important, considering possible changes in the partner organizations and their teams over time, can be coordinated by the integrating organization, if such exists.

The use of advanced information technologies greatly strengthens the ability of a group of users from different organizations to gather, analyze, and share information in more efficient and convenient ways than before. Web- and cloud-based solutions are increasingly an integral part of effective partnerships.

Developing effective mutual relations

A strong leadership, committed to the measurement process, is key both to triggering the joint measurement process and assuring its implementation over time. The leadership will

help bring the relevant partners to the table, generate and sustain their commitment, and ensure that the process progresses according to the goals defined.

To strengthen the commitment of the partners and to guarantee that the joint measurement meets the needs of each of them, it is of particular importance to ensure their broad and significant involvement in the planning and design of the measurement.

To secure coordination throughout the measurement stages and avert unnecessary conflict, a clear division of responsibilities and roles of the partner organizations and the integrating organization should be determined in advance. It is important to document such agreements in writing, whether in an informal document or a formal memorandum of understanding (MOU), to follow up on obligations, updating them as necessary, and ensuring accountability in the event of non-compliance.

For each of the partners to agree to share information with the others, it is important to form agreed arrangements on its use and to ensure partners' mutual trust. To alleviate concerns and avoid future conflicts, it is advisable to use a data-sharing agreement regarding the permitted collection and use of information, which may be a part of a broader MOU concerning labor rules and the division of roles and responsibilities between partners.

In some cases, concerns arise among partners regarding the publication of comparative organizational performance data relating to publishing findings to audiences beyond the partnership's boundaries or even within the partnership itself. In addressing such concerns, it should be noted that there are different ways of analyzing and using data without identifying specific organizations, for example, as a comparison between unnamed organizations or by presenting the data in an aggregative manner, with each organization receiving its data and comparing itself to the general average. In addition, it is possible, of course, to distinguish between the level of detail of information available to organizations that are members of the partnership and the level known to external factors.

Establishing effective processes for learning and continuous improvement

Creating opportunities for learning and continuous improvement: Discussion of the findings at different levels, within and between partner organizations, is essential for meaningful learning and for advancing processes of continuous improvement. To enable the ongoing discussion, it is valuable to hold periodic meetings where the partners meet directly and in an open and safe atmosphere. Of course, between meetings it is desirable to maintain regular contact.

Ensuring the benefits from using the findings for each individual partner and for the partnerships as a whole: It is important to enable information use at both levels; to promote the continuity of the shared measurement over time, a balance must be achieved between the emphasis placed on the use of information for the specific needs of each partner and the emphasis on information use for the purposes of the entire partnership.

Ensuring the quality and relevance of measurement over time

Developing processes to assure data collection and analysis quality: In every measurement process there are concerns about missing, inconsistent, or inaccurate information, which is even greater when working in an inter-organizational framework. Partners must therefore set common quality goals and agree to quality assurance processes, such as sampling and periodic training.

Continuous update and improvement of the measurement process: Periodic examination of the measurement strategy and its application processes and obtaining partner feedback are important for their improvement and adaptation to evolving needs. In this context, it is important to document the processes for implementing the measurement and findings both for each partner and at the level of the entire partnership. A periodic evaluation study on the joint measurement process can greatly assist in a systematic examination of the measurement mode of implementation and quality, the way the findings are used, and the implementation of the improvement processes.



Edmond de Rothschild Foundation (Israel)

104 Rothschild Blvd. Tel Aviv, Israel

www.rcf.org.il | foundation@caesarea.com