

The Initiative for Applied Research in Education

Israel Academy of Sciences and Humanities

Ministry of Education, Culture, and Sport

Rothschild Foundation (Yad Hanadiv)

Committee on Assessment and Evaluation in Education

The Knowledge Base for Assessment and Evaluation in Education

A Framework for Curricula: Academic Studies and Professional Development Programs

August 2005

The **Initiative for Applied Research in Education** was founded in 2003 as a joint project of the **Israel Academy of Sciences and Humanities**, the **Ministry of Education, Culture, and Sport**, and the **Rothschild Foundation (Yad Hanadiv)**.

The aim of the Initiative is to help improve educational achievement in Israel by developing the field of applied research in education—research that answers user-inspired questions, thereby furthering knowledge and research methods. The Initiative is based on examples from the United States and Europe, where science academies work together to advance the education system, making use of both research and experience. In these countries, improved student achievement has, under certain conditions, proven to be related to the systematic accumulation and intelligent use of knowledge and evidence.

Three working assumptions guided the establishment of the Initiative:

- New knowledge in various fields, from brain science to information management, may contribute to research and practice in education. In Israel there are research capabilities—in education and other fields—that can be encouraged to focus on improving educational achievement.
- Asking research questions derived from decision-makers' agenda may encourage education researchers, on their own and in collaboration with scholars in other fields, to expand the creation of knowledge that can benefit education practitioners. In an effort to answer these questions, new tools and theories may be developed, partly in connection with research and international development.
- Decision-makers in education, from teachers to the Education Ministry administration, will want to derive practical benefit from knowledge and rigorous research that is made available to them and to contribute to the development of a growing body of knowledge based on their own professional experience.

Up-to-date information about the activities of the Initiative can be found on the Academy's Website: www.academy.ac.il, Science Policy, Applied Research in Education.

The Israel Academy of Sciences and Humanities was founded in 1959. Its membership currently comprises ninety-one top Israeli scientists and scholars. According to the Israel Academy of Sciences and Humanities Law, 1961, its principal objectives are to bring together outstanding Israeli scholars; to foster and promote scientific activity; to advise the Government on research activities and scientific planning of national importance; to maintain ties with equivalent bodies abroad; to represent the Israeli scientific world in international institutes and conferences; and to publish articles that can further scholarship.

The Ministry of Education, Culture, and Sport was founded in 1948, when Israel became an independent, sovereign state. Under the State Education Law, 1953, the Ministry is responsible for the education of Israeli children from preschool through high school, up to and including the twelfth grade. In addition, the Ministry is in charge of teacher-training in colleges of education. The Ministry deals both with pedagogical policy (e.g., development of curricula, teaching methods, and standards and tests) and with organizational policy (e.g., budgeting for the education system, logistical planning, attention to special population groups, and inspection of educational institutions).

The Rothschild Foundation (Yad Hanadiv) is continuing the Rothschild family's philanthropic activity in Israel, which Baron Edmond de Rothschild began in the late nineteenth century. In the field of education, the Rothschild Foundation works on improving educational achievement, especially by increasing opportunities for all Israeli pupils to have a high-quality education. The Rothschild Foundation makes cutting-edge knowledge and expertise available to education workers, thereby spurring innovation, which can improve vital components of the Israeli education system.

The Knowledge Base for Assessment and Evaluation in Education: A Framework for Curricula

Assessing student achievement and evaluating educational programs, interventions, institutions, and systems are growing fields of activity and research in many countries and are high priorities in Israeli educational policy.

Israel has a highly developed system of state matriculation exams administered at the end of high school. In the past decade, aware that a final assessment at the end of the school educational process is not sufficient, the Ministry of Education has been developing a system of indicators of school effectiveness and growth known as **GEMS** (Growth and Effectiveness Measures for Schools). Data are gathered in tests and questionnaires in grades 5 and 8 in half of all schools each year.

The quality of the educational process contributes to its achievements no less than the do the resources and inputs invested. It is therefore essential to describe, assess, and evaluate the process and its outcomes, based on the assumption that the educational process can and should be managed in such a way as to strive for well-defined objectives.

The Initiative chose to assist the Israeli professional community in this field by addressing the question of *what people engaged in assessment and evaluation in education need to know*. The answers will enable institutions of higher education, professional organizations, and the Ministry of Education to offer programs of study and training that take into account local needs and conditions, as well as the research frontier and developments in the science and practice of assessment and evaluation. These programs will serve the needs of the Israeli professional community and contribute to its further development.

To this end, a committee of experts, scholars, and evaluators from leading Israeli institutions was formed. The committee began by organizing an educational seminar in December 2004 and January 2005 entitled “What Do People Involved in Assessment and Evaluation in Education Need to Know?” The seminar brought together some two hundred researchers from Israel and abroad, as well as education practitioners (teachers, principals, and evaluators). Both groups presented knowledge and experience and discussed basic questions regarding the assessment of achievement and program evaluation in practice and in research.

The Initiative’s steering committee asked the committee of experts to propose a “framework for curricula” for academic studies and professional development programs. In preparing the framework, the committee of experts developed programs designed to address the experience,

knowledge, and needs of different categories of potential candidates and to meet various objectives, ranging from professional advancement of schoolteachers to academic training of evaluation researchers in universities.

In the course of its work, the committee consulted with experts in numerous relevant fields. We would especially like to thank Chani Broderson, director of the Department of Elementary Education in the Municipality of Tel Aviv; Ram Cohen, principal of Ironi Aleph High School in Tel Aviv; Professor Victor Lavy of the Department of Economics at the Hebrew University of Jerusalem;¹ Professor David Nevo, head of the Tel Aviv University School of Education;² and Shlomit Shimron, an organizational consultant for systems of education, society, and the community. Their comments were very helpful to the committee in its discussions and in the drafting of its report.

The proposed “framework for curricula” was sent to experts in Israel and abroad for peer review. The purpose of the review was to obtain critical and objective outside perspectives and to help the committee members produce the best possible report. To ensure objectivity, the committee members were unaware of the referees’ identity until the report was published. We would like to thank the referees for their professional assistance:

Dr. Haggai Kupermintz, Faculty of Education, University of Haifa

Dr. Lauress (Laurie) Wise, President, Human Resources Research Organization (HumRRO)

Professor Alan Lesgold, Dean, School of Education, University of Pittsburg

Dafna Lev, Ministry of Education, Director Central Region

Dr. Miri Levin-Rozalis, Dept. of Behavioral Sciences, Ben Gurion University of the Negev

¹ Victor Lavy is a full professor of economics at the Hebrew University of Jerusalem and director of the Maurice Falk Institute for Economic Research in Israel. His main research interests are the economics of education and the economics of work and development. He studies program evaluation, as well as educational, social, and welfare experiments and policy. He served as advisor to the Evaluation Department of the Ministry of Education and to the director general of the Ministry from 1997 to 2002 and was a member (and sometimes head) of public and professional committees on educational and social issues. He is currently a member of the National Task Force for the Advancement of Education in Israel. He has a doctorate in economics from the University of Chicago.

² David Nevo heads the Tel Aviv University School of Education. He specializes in educational evaluation, with a special interest in theories of evaluation: curriculum evaluation and evaluation of educational projects, school-based evaluation, and the pedagogical and social significance of educational evaluation. He evaluates curriculum units and educational projects in the Israeli education system and works on the development of school-based evaluation in schools in Israel and abroad. His research topics include the significance of combining internal and external evaluations in schools. He is a former chief scientist of the Ministry of Education. His books include *Beneficial Evaluation* (Masada, 1989) (in Hebrew); *Evaluation in Decision Making* (Kluwer, 1988; co-author: Naftaly S. Glasman); *School-Based Evaluation* (Pergamon, 1995); and *School-Based Evaluation: An International Perspective* (Elsevier, 2002). He is also the editor of the international journal *Studies in Educational Evaluation* (SEE).

Committee members

Professor Ruth Butler, Hebrew University of Jerusalem

Professor Menucha Birenbaum, Tel Aviv University

Professor Gershon Ben-Shakhar, Open University (President) and Hebrew University of Jerusalem

Dr. Yoav Cohen, National Institute for Testing and Evaluation

Dr. Fadia Nasser Abu al-Hija, Tel Aviv University

Dr. Tali Freund, Center for Educational Technology (CET)

Ms. Gina Shimborsky, Henrietta Szold Institute

Professor Michal Beller, Educational Testing Service (ETS), Princeton, New Jersey (resigned from the committee before taking over as head of the National Authority for Assessment and Evaluation on May 1, 2005)

Committee coordinators

Ms. Leora Sidi (until April 30)

Ms. Mevasseret Nevo (from May 1)

Contents

1. Introduction	6
1A. Educational Assessment and Evaluation	6
1B. Theoretical Background	7
1C. Academic Degrees and Certification or Licensing	15
1D. The Situation in Israel: Programs and Enrollment	16
2. The People Involved in Educational Assessment and Evaluation and the Knowledge They Need	17
2A. Principles of the Framework for Curricula	17
2B. The “Framework” as a Matrix of Topics and Professions: Knowing and Knowing How	19
2B.1 Itemization and Explanation of the Framework: The Knowledge Axis	21
2B.2 Itemization and Explanation of the Framework: The Axis of Professions	26
3. Sample Curricula for Academic and Professional Development Programs Based on the Framework	33
3A. Teachers: Proposal for Training and Continuing Education in Assessment and Evaluation	33
3B. Principals: Proposed Curriculum in Assessment and Evaluation	34
3C. Assessment and Evaluation Coordinators: Proposed Curriculum in Assessment and Evaluation	36
3D. Master’s Degree: Proposed Curriculum in Assessment and Evaluation	37
4. Recommendations	41
4A. Creation of a Support System for the Development of Researchers and Practitioners	41
4B. Further Development Required	42
4C. Monitoring of Activity Surrounding the Framework	42
5. Appendices	43
5A. Appendix 1: Seminar Schedule	43
5B. Appendix 2: Seminar Speakers	46
5C. Appendix 3: Examples of Assessment and Evaluation Programs at U.S. Universities	49
5D. Appendix 4: Members of the Committee on Assessment and Evaluation	51

1. Introduction

1A. Educational Assessment and Evaluation

Assessment and evaluation in educational contexts are valuable and meaningful so long as they further the achievement of educational goals, contribute to the knowledge and proficiency of educators, and guide them in their work. In this context, the present report reflects several assumptions.

First, we assume that people, organizations, and systems can learn and improve, and that they function best when they work together to achieve common goals. Second, the overarching goal of assessment and evaluation in education is to provide people and systems with information that can help them learn and become more effective in setting goals for the future. achieving their goals, and in making informed decisions regarding policy and practice. Third, instead of approaching assessment and evaluation narrowly as a means of making teachers and principals more accountable for student achievement, we emphasize the shared responsibility of all partners in the education system for striving for worthwhile goals and outcomes that include, but are not limited to, academic achievement. Members of the educational community must be proficient in the language, methods, and instruments used in assessment and evaluation, and must acquire the knowledge and expertise needed to identify, produce, and interpret information relevant to monitoring, evaluating, and improving the methods, organization, and outcomes of learning and instruction.

Here, it is important to emphasize that assessment and evaluation are not ends in themselves. Rather, they should serve and promote the goals and priorities of authorized decision-makers in the education system. It should not be left to assessors and evaluators to determine the goals of education. In this context, it is important to keep in mind that the relationship between assessment and evaluation on the one hand and educational processes and outcomes on the other is reciprocal: assessment and evaluation not only reflect goals and outcomes but also influence them. Furthermore, practices that promote the achievement of some objectives frequently undermine the achievement of other objectives. In this case, informed assessment and evaluation must take into account both the benefits to students, teachers, schools, and the education system and possible costs and detrimental effects and must therefore be based on a deep understanding of the relationship between different educational goals and different methods of assessment and evaluation.

Scholars and academic institutions can also play an important role in such a system of shared responsibility, and can make a substantial contribution to the academic and applied realm of educational assessment and evaluation. In brief, first, they are responsible for producing knowledge

relevant to understanding how educational objectives and achievement of them are related to methods of assessment and to the development of valid, reliable methods and instruments. By so doing, they can provide the educational community with a scientific basis for making intelligent decisions and for using assessment in beneficial ways, while considering and attempting to avoid possible misuses, abuses, and detrimental effects. Second, they are responsible for developing and offering programs that can meet the need for capable, highly trained professionals in the field of assessment and evaluation.

A framework has been designed for conceptualizing these issues, with which all practitioners and scholars engaged in one way or another in educational assessment and evaluation must be familiar, and for determining the levels of knowledge and skill needed by different professionals. Based on the framework, we then propose principles and content for academic degree programs, pre- and in-service training, and professional development programs in the field of assessment and evaluation.

1B. Theoretical Background

First, it should be made clear that the terms *assessment* and *evaluation* are not unequivocal. They refer to at least two subtopics that, despite some overlap, differ fundamentally. In this report we distinguish between **educational assessment**, which focuses on educational units (the individual student, class, teacher, or school) and examines various aspects of them (e.g., academic achievement in a particular subject, school climate, and attitudes) at a single point in time or over a particular period (e.g., assessing improvement in achievement), and **program evaluation**, which attempts to determine whether an educational intervention (e.g., a curriculum unit or violence-reduction program) has achieved its objectives.

In the first type of assessment there are many different approaches and methods (e.g., external assessment, internal assessment, and alternative methods of assessing achievement), and new instruments are continually being developed. All the approaches are geared towards assessing educational units, whether in relative (normative) terms or in absolute terms (compared with criteria). There are also several approaches to the second type of assessment. Although program evaluation also tends to evaluate educational units, it is based on a different methodology and on a set of research methods designed to answer the question “Has the intervention achieved its objectives?” or, in other words, is there a *causal relationship* between a particular intervention (or program) and specific results?

Shared Principles of Assessment

Although the kinds of assessment used in different contexts for different purposes are distinct from one another, they do share certain principles. One important principle is that assessment is always a process of drawing conclusions based on data and evidence. However, like any other kind of measurement, educational assessment inherently has a certain degree of imprecision: the results of an educational assessment are merely an estimate of what a particular person knows and is capable of doing.³

Educational Assessment

A committee of American experts that studied the issue of “knowing what students know”⁴ described the assessment process as follows: Any assessment of students, whatever its purpose, must be based on three elements that together make up an “assessment triangle.” The triangle represents the three key components underlying any educational assessment: a model of the student’s **cognition** and learning in any field; a set of **perceptions** regarding the kind of **observations** that may provide information about the student’s knowledge and abilities; and a process of **interpretation**, which helps us make sense of the data and evidence. The three are represented as the vertices of a triangle, with each vertex linked to and dependent on the other two. Understanding the data and evidence should be subject in part to criteria pertaining to the quality of the assessment and evaluation, such as reliability and validity (internal and external), and to possible ramifications of the assessment.⁵

The cognition vertex in the triangle has to do with theory, or with a set of ideas about how students demonstrate knowledge and develop expertise in a particular school subject. Whenever

³ National Research Council (NRC), *Knowing What Students Know: The Science and Design of Educational Assessment* (Washington, DC: National Academies Press, 2001), p. 2. For a list of committee and board members, see below.

⁴ Ibid.

Committee on the Foundations of Assessment: James W. Pellegrino (co-chair), Robert Glaser (co-chair), Eva L. Baker, Gail P. Baxter, Paul J. Black, Christopher L. Dede, Kadriye Ercikan, Louis M. Gomez, Earl B. Hunt, David Klahr, Richard Lehrer, Robert J. Mislevy, Willie Pearson, Jr., Edward A. Silver, Richard F. Thompson, Richard K. Wagner, Mark R. Wilson, Naomi Chudowsky, Tina Winters, M. Jane Phillips.

Board on Testing and Assessment: Eva L. Baker (chair), Lorraine McDonnell, Lauress L. Wise, Richard C. Atkinson, Christopher F. Edley, Jr., Ronald Ferguson, Milton D. Hakel, Robert M. Hauser, Paul W. Holland, Daniel M. Koretz, Richard J. Light, Barbara Means, Andrew C. Porter, Loretta A. Shepard, Catherine E. Snow, William L. Taylor, William T. Trent, Guadalupe M. Valdes, Vicki Vandaveer, Kenneth I. Wolpin.

⁵ For details, see Menucha Birenbaum, *Alternatives in Assessment*, chap. 6 (Tel Aviv: Ramot, 1997) (in Hebrew).

achievement is assessed, there must be a theory of learning in that particular subject in order to identify the important knowledge and skills that should be assessed.

The observation vertex represents a description or itemization of instructions for assessment tasks that will produce thoughtful responses from students. Every assessment of this sort is based on a set of perceptions regarding the types of tasks and situations that will prompt the students to say, do, or create something that demonstrates important knowledge and skills. The tasks that the students are asked to perform in the assessment are not arbitrary. They should be planned carefully in order to provide data and evidence related to the cognitive model of learning, and in order to support inferences and decisions based on the results of the assessment.

Every assessment is based on specific assumptions and models for interpreting the data and evidence gathered in the observation process. The interpretation vertex includes all methods and tools used to draw conclusions from observations. It explains how distinctions that emerge from observation of a series of tasks provide data and evidence regarding the skills and knowledge to be assessed.⁶

A distinction is customarily drawn between two main types of educational assessment: large-scale assessment and classroom assessment.

Large-Scale Assessment

Large-scale assessments are designed to provide information about the achievements of a large number of students. Program evaluations offer further information about the relationship between the findings regarding achievement and factors that may explain or contribute to them.

In a workshop on “assessment in support of instruction and learning,” organized by the National Research Council (NRC),⁷ Professor Lorrie Shepard distinguished between three uses of large-scale tests.

The **first use** links the assessment of achievement with program evaluation: Some large-scale assessments make it possible to compare the achievements of a large number of students in order to identify key patterns of strengths and weaknesses and indicate ways in which the curriculum, teaching methods, or organization of schools may be modified and improved. Assessments designed for large-scale use in order to provide data on the achievements of a large number of

⁶ NRC, *Knowing what Students Know*, pp. 44–49.

⁷ *Assessment in Support of Instruction and Learning: Bridging the Gap between Large-Scale and Classroom Assessment: Workshop Report* (Washington, DC: National Academies Press, 2003).

students may, Shepard argues, demonstrate the achievement of the educational goals described in the standards and curriculum.

A **second use** of large-scale assessment is for confirmation at a specific point in time or for screening or placement purposes. For example, when we want to identify students who are not ready for a reading assignment on a particular level, it may be helpful to assess their specific achievements **in order to identify the weaknesses that have to be addressed.**⁸

The **third use** links large-scale assessment with classroom assessment. Large-scale assessment can offer teachers opportunities for effective professional development. Test development, scoring rubrics, curriculum units, and standards-based professional development give them an opportunity to incorporate in the curriculum efforts to improve classroom assessment strategies.⁹

There are two ways to assess changes in students' achievements over time: *Cohort-to-cohort* models compare the performance of one group of students at a particular stage in their schooling with previous groups at the same stage (e.g., fifth-graders in one year with fifth-graders the previous year). In contrast, *value-added measurement* models assess the development of the same students over time (e.g., comparing the achievements of students who are now in fifth grade with their previous achievements in fourth grade).

In Israel, large-scale assessment is carried out by means of two main mechanisms: a highly developed system of state matriculation exams and a system of Growth and Effectiveness Measures for Schools (GEMS). Both mechanisms can, in principle, make possible system-wide comparisons using the cohort-to-cohort method. The matriculation exams assess student achievement in the various subjects at the end of high school. GEMS is a set of academic, social, and educational indicators designed to provide a snapshot of the education system on its various levels. The basic assumption of GEMS is that the school is a system made up of many parts: organization, relationships, curriculum, teaching methods, staff development, and achievement. GEMS looks at the following topics: the school's priorities and staff development based on these priorities; characteristics of the learning environment; academic achievement in the mother tongue, math, science, and English; special help for students who are having difficulty; and the school climate and work environment.

Committee members: J. Myron Atkin (chair), Eva Baker, Jan de Lange, Tom Keller, James Minstrell, Marge Petit, Anthony Scott, Loretta Shepard, Guadalupe Valdes.

⁸ Ibid., p. 11.

⁹ Ibid., pp. 11–12.

GEMS points out the system's strengths and weaknesses. It is primarily a tool to help school principals plan their work. In addition, it serves the central staff of the Education Ministry, the district offices, local authorities, networks of schools, and so on. The GEMS data are gathered from tests and questionnaires administered each year in grades 5 and 8 in half of the schools (alternating schools each year), which makes cohort-to-cohort comparisons possible. In order to avoid too many state exams, each class takes the GEMS tests only once, in fifth or eighth grade; thus, no analysis of the findings can be done by the value-added method.

Although all types of assessments share certain principles, different assessments serve different goals. It should be noted that different assessment methods may be appropriate for different functions. Often, for financial and practical reasons, one assessment instrument is used for a wide variety of goals. It should be borne in mind, however, that the more goals a particular type of assessment claims to serve, the less effectively it can achieve each individual goal. In particular, it is important not to extrapolate the findings of one large-scale assessment to the ability of individual students.¹⁰ Everyone involved in large-scale assessment must maneuver between the desire to improve the ability of schools and education systems to take responsibility for students' achievements and the desire to improve their ability to provide specific guidance to specific students.

Classroom Assessment

Professor Dylan Wiliam has defined classroom assessment as “assessment for learning.”¹¹ In other words, if the goal of an assessment is to improve the students' learning process in some direct manner rather than to rank, measure, or confirm some final aspect of their performance, then it belongs to the realm of classroom assessment.¹² Furthermore, the emphasis in classroom assessment is on the feedback that it provides for the sake of improving both learning and teaching. All three components—teaching, learning, and assessment—must be aligned in order for the objectives of education to be achieved.

Another common way of characterizing various types of assessment is distinguishing between “formative assessment” (assessment *for* learning) and “summative assessment” (assessment *of* learning). Assessment that is intended primarily to give students feedback and influence their learning is often called “formative assessment”; “summative assessment” is designed to assess the

¹⁰ Jay P. Heubert and Robert M. Hauser, eds., *High Stakes: Testing for Tracking, Promotion and Graduation* (Washington, DC: National Academy Press, 1999), p. 12.

¹¹ *Assessment in Support of Instruction and Learning*, p. 13.

¹² *Ibid.*

achievements of students at the end of a particular learning process. Nevertheless, it should be stated that both assessment methods—large-scale and classroom—can serve both formative and summative purposes.¹³ The actual assessment must suit the goal for which it is intended.

William and Black (Assessment and Classroom learning, assessment in education, 1998) conducted a comprehensive review of studies that examined the relationship between formative assessment and improvement of student achievement. Their review showed that formative assessment boosts student achievement, even though the manner in which it is currently conducted leaves much to be desired.¹⁴ Their study, carried out in various school subjects and in various countries among students from the age of five to university undergraduates, showed that broader use of formative assessment yields greater, more significant improvement in achievement.¹⁵

Some, though not all, studies have shown that improved formative assessment is particularly beneficial to low-achieving students and that it contributes to reducing variance and raising the overall level of achievement.¹⁶ As noted above, the researchers also concluded that standard practices of formative assessment are inadequate and showed how formative assessment can be improved. They also demonstrated the relationship between such improvement and improved student achievement.¹⁷

In 2001, the Committee to Review the Format of the Matriculation Exams and Final Exams (the “Ben-Peretz Committee”) submitted its recommendations to Education Minister Limor Livnat. In the wake of the recommendations, an experiment was performed to develop classroom assessments that would improve teaching and learning processes (i.e., formative assessment), and others that would serve screening and certification purposes (i.e., summative assessment). The project, termed “Matriculation 2000,” was implemented in twenty-two high schools in 1994–1999.¹⁸ In each school, school-based assessments were conducted in one to three of the subjects on the list of matriculation-exam subjects. These schools were authorized to give final grades based on “alternative assessment” formats. These grades were equivalent to the matriculation scores based

¹³ Ibid., p. 14.

¹⁴ Paul Black, Christine Harrison, Clare Lee, Bethan Marshall, and Dylan William, *Assessment for Learning: Putting It into Practice* (Maidenhead: Open University Press, 2003), p. 6.

¹⁵ Ibid., p. 9.

¹⁶ Ibid.

¹⁷ Paul Black and Dylan William, *Inside the Black Box: Raising Standards through Classroom Assessment* (London: School of Education, King's College, 1998), p. 2.

¹⁸ Report of the advisory committee on the Matriculation 2000 project in 22 schools, Ministry of Education, Committee of University Heads (Jerusalem, 1999).

on external exams (whereas at present the matriculation score is composed of an average of the score on a national exam and the final grade from school).

The Inter-University Committee on Admissions Policy approved the experiment as a way of examining the feasibility of the recommendations of the Committee to Review the Format of the Matriculation Exams and Final Exams. The approval was made contingent on the appointment of an advisory board.

The report of the advisory board was submitted at the beginning of the 1999/2000 school year. Its main findings were as follows: The project had improved teaching and resulted in a process of constant reexamination of the state curriculum in the subject. Teaching and learning had been diversified, with a variety of instruments being used to assess achievement. Gains were evident in high-order thinking as well as in knowledge and understanding of the subject matter. There was evidence of an improved academic atmosphere in the project subjects and of increased motivation among students. The vast majority of students wanted the project to continue. The validity of assessments—an indicator of the degree to which the assessment accurately represents the overall curriculum and goals in a particular subject—had increased. In general, no grade inflation was found. The collaboration among the teaching staffs was an important component in achieving the objectives and was accompanied by increased motivation among teachers and reinforcement of their professional status.

Alongside the project's achievements, there were some constraints, difficulties, and limitations. Overall, students' reports of a better atmosphere and higher motivation were more lukewarm than those of the faculty. The project entailed a heavy workload for the teachers. The expectation that the project would also influence methods of teaching and assessment in subjects not included in the project in the particular school was only partially realized. As a rule, the range of assessment instruments was not used to the fullest. Although in principle teachers recognized the importance of the validity of the assessments (an indicator of accuracy and objectivity), in practice adherence to it was only partial. Difficulties were also observed in faculty teamwork; in some cases individual teams did not function at the required level.

At our educational seminar, Dr. Rita Sever described the project and the accompanying research study, and Mirit Lavi described her school's activity in the project.

It is important to combine internal assessment in the school (formative and summative) with external assessment (e.g., the GEMS tests) in order to improve teaching and learning processes. One prerequisite for a productive dialogue between the findings of the two types of assessments for the purpose of maximum evaluation utilization is the existence of a culture of organizational learning in the school.

Program Evaluation

Program evaluation is a broad and vibrant field of research, and there is no standard definition that applies to the entire field and its activities.¹⁹ The definition proposed by Carol Weiss may serve as a basis for further definitions. According to Weiss, evaluation means systematically reviewing the implementation or outcomes of a program or project and comparing them with explicit or implicit standards as a means of improving the program or project or making decisions related to it.²⁰

Many social scientists, especially those involved in evaluation research, regard randomized experiments as the best method for this research (this opinion was represented at the seminar by Professor Joshua Angrist, who discussed methods of evaluating educational interventions). Some recommend a hierarchy for selecting the appropriate method when planning an evaluation study: the randomized experiment is at the top and should be preferred when circumstances allow; when they do not, one should use a quasi-experiment, cross-sectional study, panel study, or time-series study. If none of these is possible, the “before and after” method should be used. Guba and Lincoln (1989) believe, in contrast, that qualitative studies best meet the needs of program initiators and implementers.

In all cases, in-depth, responsible evaluation activity should combine diverse methods and base its conclusions on a variety of instruments.

Assessment of Teachers

In recent decades, the assessment of teachers and principals has gained widespread attention in educational policy and in the professional literature on educational research. In this report we discuss the assessment of teachers. The assessment of principals in Israel is currently carried out by administrative echelons and is influenced by assessments of the schools’ success. Perhaps in the future it will be important to develop concepts and tools for assessing principals and to provide appropriate training to the people who will engage in such assessment.

The Ministry of Education, influenced in part by the recommendations of the National Task Force for the Advancement of Education in Israel, currently wants to institutionalize the assessment of teaching and teachers. In general, the approach of the present report to the issue of teacher assessment reflects the assumptions and principles outlined in the introduction. First and foremost,

¹⁹ Based on Itzhak Friedman, *Measurement and Evaluation of Social and Educational Programs* (Jerusalem: Szold Institute, 2005) (in Hebrew), p. 127.

²⁰ Carol H. Weiss, *Evaluation: Methods for Studying Programs and Policies*, 2nd ed. (Upper Saddle River, NJ: Prentice Hall, 1998), p. 4.

teachers and school principals can benefit from reliable, valid, and fair information that can help them evaluate the degree to which they are achieving their objectives, identify weaknesses, and come up with methods for professional development and for improving teaching and the processes and outcomes of learning. One of the committee’s recommendations is thus to train those involved in assessment and evaluation in the education system, and especially school principals and assessment and evaluation coordinators (A&E coordinators), in the theoretical, social, and practical aspects of teacher assessment.

1C. Academic Degrees and Certification or Licensing

The committee members faced a special dilemma: a tradition of academic, mostly theoretical instruction in educational assessment and evaluation as a component of advanced degree programs versus a need for graduates who can perform various assessment-related tasks in the education system almost here and now. The committee therefore addressed two kinds of programs—academic programs and professional development programs—but made a deliberate effort to link them conceptually and structurally.

Even if, as recommended by the Implementation Committee on Incorporating Internal Assessment in the Schools as part of the implementation of the recommendations of the National Task Force,²¹ a master’s degree will eventually be required for “school assessment and evaluation coordinators,” and even if advanced degrees are required for various positions in the regional educational administrations and the National Authority for Assessment and Evaluation, holders of such degrees will need practical knowledge when they near the completion of their studies. Therefore, when drawing up guidelines for new programs, the committee saw fit to combine the content studied with the ability to use it.

Furthermore, the committee saw added value in training senior evaluators to be knowledgeable, at least to some extent, about the educational context and subjects in which achievement is assessed and in training teachers to be knowledgeable, at least to some extent, about system-wide assessment.

²¹ The members of this committee are Dr. Gili Schild (chair), Motti Asulin, Dorit Dimon, Shimon Harel, Dr. Ilana Zeiler, Dr. Miri Levin-Rozalis, Shalom Slonim, Dr. Haggai Kupermintz, Michal Karkoff, Batya Schochen, and Dr. Esti Schleyer.

1D. The Situation in Israel: Programs and Enrollment

1. Professional Development for Teachers and Principals

According to the Department for In-Service Training in the Administration for Teacher Training and In-Service Training and Mentoring in the Ministry of Education, in the 2003/04 school year teachers and principals attended in-service courses in assessment and evaluation in various Pisga (faculty development) centers (Taibe, Eilat, Ashkelon, Beersheva, Bat Yam, Herzliya, Kefar Sava, Netanya, Azata, Acre, Afula, Kiryat Gat, Kiryat Malakhi, Ramat Gan, and Tel Sheva), at the Center for Development of Senior Educators, at Beit Berl College, at the Open University, at the Israel Institute of Technology, at the Mofet Institute, at Oranim, and in schools. This year, seventy-seven such courses were taught in the various institutions (seven for principals and the rest for teachers).

In 2003/04, 120 principals and approximately 1,150 teachers took courses in assessment and evaluation and earned in-service course credit (the average principals' course lasted 30 hours; the teachers' courses generally lasted 28–56 hours).

The Ministry of Education plans to expand opportunities for the professional development of teachers and principals in this in-demand field in the coming years.

2. Advanced Degree Programs

Based on the data obtained by the committee from the universities, it seems that only Tel Aviv University currently offers a master's-degree program in research methods, assessment, and evaluation. Since its inception in 1979, some eighty people have graduated from this track. All the other universities offer courses on research methods and educational assessment, some of them compulsory, as part of their advanced degree programs in education. (Training programs and in-service courses for teachers at universities and colleges of education include courses on educational assessment, but these are not all advanced degree programs.)

Various academic institutions, especially Tel Aviv University and Ben-Gurion University of the Negev, are about to offer a range of expanded programs of study and training programs in assessment and evaluation and to start training evaluators within the education system. These programs are in the final stages of approval.

In view of the system's great need for professional knowledge in the field of assessment in each of the echelons dealing with education, as well as the need for program evaluators for the education system, below we propose a framework for curricula (for academic programs and professional development) to help ensure the quality of activity in this important field.

2. The People Involved in Educational Assessment and Evaluation and the Knowledge They Need

2A. Principles of the Framework for Curricula

The purpose of the framework for curricula (hereafter, the “framework”) is to make possible—and perhaps even to ensure—an agreed-upon knowledge base shared by everyone involved in assessment and evaluation in education. The expected growth in training, in-service courses, and academic studies in this field creates an opportunity to review academic programs and in-service courses in the context of shared concepts and abilities, for the benefit of the education system and research. A common framework of knowledge and skills based on the best research and experience in Israel and abroad—one that will make possible monitoring and assessment of educational activity on its various levels—may help to constantly improve educational achievement and strengthen the Israeli professional community.

The framework is meant to describe the shared concepts and abilities and the boundaries of the field and to guide the planning of curricula and programs in assessment and evaluation. It should facilitate the development of curricula in this field, in accordance with the goals of the professional community.

The framework combines research and academic knowledge with descriptions of the jobs of practitioners, from teachers to program evaluators. The idea of linking theoretical and practical knowledge with the professions needed in the system required a re-examination of the content of assessment and evaluation, of the ways knowledge of the material is expressed, and of the numerous ways in which the content and indicators of knowledge can be combined; the acid test for inclusion of a particular context in the curriculum is the need for it in practice. This effort is preliminary, especially because the various jobs related to assessment and evaluation (see below) have not yet been formally established and introduced in Israel. As a result, there must be ongoing, formative monitoring to improve the framework, in keeping with the roles that alumni of the academic programs and in-service courses will fulfill and the ways in which they will do so.

The proposed framework is not based on established knowledge about the success of specific curricula and programs, evaluated according to the professional success of their alumni. We could not find any such evaluations; nor could we find any studies of criteria for evaluating the success in assessment work of graduates of academic programs or in-service courses for teachers, principals, or other practitioners in the education system. (The appendix contains information about selected programs of study, especially in American universities that offer advanced degree programs in this

field. Some of them are designed for researchers and evaluators of programs or interventions; others are for teachers or principals.)

The framework specifies essential content and topics for all those working in educational assessment and evaluation, by profession. Sample curricula for each profession are detailed in section 3.

The proposed framework is set up as a field with two axes: (1) the knowledge axis (content and topics); (2) the axis of professions (positions and specializations). The committee chose to present the two axes in one summary table (see pp. 19–20).

The knowledge axis presents the subject matter by category (classroom assessment, large-scale assessment, research methods, statistics, assessment, program evaluation, assessment of teaching personnel). Each category includes several topics (e.g., “assessment in the educational context”). In the itemization of subject matter (see pp. 21–26), each of these topics is then divided into subtopics. This itemization represents a mapping of the field, and the committee tried to include in it the material that should be learned and known by the people involved in assessment and evaluation in education.

The axis of professions shows the workers in the system. Naturally, different jobs involve different aspects of assessment and evaluation. For each topic on the list, every professional can be expected to have some level of theoretical and practical knowledge to do his/her job as a consumer of knowledge (P1) or as a producer of knowledge (P2). The “profile” of required knowledge on each topic varies from job to job. The markings in the table denote the type of knowledge required (theoretical/declarative, procedural as a consumer and procedural as a producer). The table is intended to provide a general picture of these profiles, showing both the body of knowledge common to all workers and the specific profile of each individual job.

Due to the wide range of jobs—from teachers to psychometricians—the theoretical knowledge is divided into three levels, from basic knowledge and familiarity with concepts (level 1) to broad, in-depth knowledge (level 3).

**2B. The “Framework” as a Matrix of Topics and Professions:
Knowing and Knowing How**

	Teachers	Principals	A&E coordinators	Program evaluators	Psycho-metricians
<u>Educational assessment</u>					
Classroom assessment					
Assessment in the educational context	D3	D3	D3	D2	D2
Setting objectives	D3 + P2	D3 + P2	D3 + P2	D2	D3
Assessment instruments	D3 + P2	D3 + P1	D3 + P2	D3 + P1	D3 + P1
Assessment quality	D2 + P1	D2 + P1	D3 + P2	D3 + P1	D3 + P2
Summarizing and reporting findings	D3 + P2	D3 + P2	D3 + P2	D3 + P1	D3 + P1
Large-scale assessment					
Assessment in the educational context	D2	D2	D3 + P1	D2	D3
Setting objectives	D2 + P1	D2 + P1	D3 + P1	D2 + P1	D3 + P2
Assessment instruments	D2 + P1	D2 + P1	D3 + P1	D2 + P1	D3 + P2
Assessment quality	D1	D2 + P1	D2 + P1	D3 + P1	D3 + P2
Findings	D1	D1	D2 + P1	D2 + P1	D3 + P2
Reporting results to various audiences	D1	D2 + P1	D3 + P2	D2 + P1	D3 + P2
Assessment-based decision-making	D3 + P2	D3 + P2	D3 + P2	D3 + P2	D1 + P1
<u>Assessment of teaching personnel</u>					
Essence and principles of assessing teachers	D3	D3	D3	D2	D3
Assessment methods and instruments	D3 + P2	D3 + P1	D3 + P2	-	D3 + P2
Possible uses of assessment of teachers	D3 + P2	D3 + P2	D3 + P2	-	D3
<u>Research methods, statistics, and assessment</u>					
Research methods					
Research paradigms	D1	D1	D2	D3	D3
Quantitative research	D1	D2 + P1	D3 + P2	D3 + P2	D3 + P2
Qualitative research ²²	D3 + P2	D2 + P1	D3 + P2	D3 + P2	D2

²² For teachers, action research only.

Combined research methods	D1	D2 + P1	D3 + P2	D3 + P2	D3 + P2
Statistics					
Basic level ²³	D2 + P1	D2 + P1	D3 + P2	D3 + P2	D3 + P2
Medium level	-	-	D2 + P1	D3 + P2	D3 + P2
Advanced level	-	-	D1	D3 + P2	D3 + P2
Assessment					
General background	D1	D1	D2	D3	D3
Theories	-	-	D2 + P1	D2 + P1	D3 + P2
Planning and development of assessment instruments	D1	D1	D2 + P1	D2 + P2	D3 + P2
Scoring	D1	D1	D2 + P1	D2 + P1	D3 + P2
Quality control	D1	D1	D3 + P1	D3 + P1	D3 + P2
<u>Program evaluation</u>					
Curriculum evaluation and evaluation of educational projects	D1	D2	D3	D3	D3
Setting objectives	D1	D2 + P1	D3 + P2	D3 + P2	D2 + P1
Evaluation protocols and evaluation instruments	D1	D1	D2 + P2	D3 + P2	D2 + P1
Evaluation quality	D1	D1	D2 + P2	D3 + P2	D2 + P1
Evaluation-based decision-making	D3 + P1	D3 + P2	D3 + P2	D3 + P2	D2 + P1

Legend:

D—Declarative

The worker should have general knowledge about the subject.

D1—knowledge on a basic level

D2—deeper knowledge

D3—knowledge on the most in-depth level

P—Procedural

The worker should have practical knowledge in the field.

P1—The worker is a consumer, i.e., uses the knowledge.

P2—The worker is a producer, i.e., is personally expected to produce the knowledge.

²³ For teachers and principals, descriptive statistics only.

2B.1 Itemization and Explanation of the Framework: The Knowledge Axis

Educational assessment

A. Classroom assessment

1. Assessment in the educational context
 - Approaches to assessment and the underlying concepts of education
 - Traditional versus constructivist teaching, learning, and evaluation cultures
 - Assessment functions and audiences
 - Ethical considerations in assessment
 - Assessment for learning (formative)
 - Assessment of learning (summative)
 - The relationship between assessment on the one hand and motivation and learning strategies on the other
 - Student involvement in assessment
 - Attitudes of teachers and students towards assessment
 - Effects and ramifications of the use of tests
2. Setting objectives
 - Standards for school subjects
 - Objectives of internal assessment
3. Assessment instruments
 - Familiarity with the various assessment instruments (test, assignment, study journal, portfolio, observation) and the appropriate context for their use
 - Design of assessment instruments, analysis, and interpretation of the results
 - Development of scoring rubrics
 - Assessment of study-related self-direction skills (assessment in the cognitive, metacognitive, resource-management, social, motivational, and emotional realms)
 - Adaptation of assessment instruments for special-needs students
 - Computer uses in classroom assessment
4. Assessment quality
 - Validity
 - Reliability
 - Response bias—test anxiety, social desirability, response sets, test-wiseness
5. Summarizing and reporting

- Grading methods (based on norms, criteria, or individual progress)
- Reporting results to various audiences (students, parents, workers)

B. Large-scale assessment

1. Assessment in the educational context

- Approaches to assessment and the underlying concepts of education
- Assessment functions and audiences
- Ethical considerations in assessment
- Assessment for learning (formative)
- Assessment of learning (summative)
- Attitudes of teachers and students towards evaluation
- Effects and ramifications of the use of tests
- Familiarity with international and official Israeli tests

2. Setting objectives

- Standards for school subjects
- Objectives of external assessment

3. Assessment instruments

- Types of tests (criterion-referenced tests, norm-referenced tests, and diagnostic tests)
- Types of items and test tasks
- Sources of information about tests and other assessment instruments
- Development of assessment instruments (structure, specifications for items, writing of the items)
- Item analysis
- Issues related to adapting assessment instruments to different population groups and languages
- Methods of setting standards and cutoff points
- Computer uses: databases of items, test generators, computerized tests, adaptive tests, automated assessment of essays and answers to open-ended questions

4. Assessment quality

- Validity
- Reliability and assessment error
- Fairness
- Bias and differential functioning of items

- Response bias—test anxiety, social desirability, response sets, test-wiseness
5. Reporting of results
 - Grades, transformations, and comparisons
 - Scaling techniques
 - Development of test norms
 - Handling of information from different forms of assessment instruments (linking and equating)
 - Comparison of groups on national and international tests, long-term comparisons (cohort to cohort, value added)

Assessment of teaching personnel

1. Essence and principles of assessing teachers
 - Assessment functions: formative and summative assessment and assessment for licensing purposes
 - Assessment of teachers in various education systems
 - Ethical considerations
 - Audiences for assessments
2. Assessment methods and instruments
 - Standards and criteria of good teaching and of good teachers
 - Analysis of teaching-learning and teacher-student interactions
 - Methods and instruments: tests, portfolios, teaching journals, lesson plans, observations, products of students' learning, peer assessment, self-assessment, assessment by students
 - Familiarity with the assessment instruments to be used in the education system
3. Possible uses of assessment of teachers
 - Improvement of teaching and education; professional development
 - Certification, licensing, and promotion

Research methods, statistics, and assessment

A. *Research methods*

1. Research paradigms
 - Positivist or post-positivist
 - Interpretive or constructivist
2. Quantitative research

- Articulation of research questions and research hypotheses
 - Quantitative research protocols (causal and correlative studies; random experiments, quasi-experiments, cross-sectional studies, panel studies, “before and after” studies, and surveys)
 - Aspects of the validity of the research plan (e.g., external and internal validity)
 - Research tools: tests, scaling methods, and questionnaires
 - Data-analysis methods (including use of computers)
 - Synthesis of research findings (meta-analysis)
 - Experimental and research ethics
 - Writing research reports
3. Qualitative research
- Trends in qualitative research: ethnographic research, phenomenological research, action research, case studies, narrative research, grounded theory
 - Articulation of research questions
 - Qualitative research tools: observations, interviews, focus groups, documents, and the researcher’s interpretations
 - Data-analysis methods (including use of computers)
 - Ethics in qualitative research
 - Quality standards in qualitative research
 - Writing research reports
4. Combined research methods

B. Statistics

1. Basic level
- Descriptive statistics (types of variables, frequency distribution, measures of central tendency, measures of variability, comparative grades, measures of association, linear regression)
 - Principles of inferential statistics (sampling distribution, hypothesis testing, type I and type II errors, confidence intervals)
 - T-tests, one-way analysis of variance
2. Medium level
- Multiple regression and path analysis
 - Analysis of variance and covariance
 - Multidimensional scaling: factor analysis, cluster analysis, and smallest-space analysis

- Meta-analysis
3. Advanced level
 - Structural equation modeling: confirming factor analysis, latent variables, and types of structural models
 - Categorical data analysis: the log-linear model
 - Multilevel analysis: features of the analysis and its uses in educational research, models of multilevel analysis
 - Simulation techniques

C. Assessment

1. General background
 - Assessment theory as a discipline
 - The history of assessment
2. Theories
 - Classic assessment theory
 - Generalizability theory
 - Item response theory
3. Planning, development, and evaluation of assessment instruments (cognitive and affective)
4. Scoring
 - Scoring methods (norms, standard-setting)
 - Anchoring, calibration, and linking of test forms
 - Models for measuring change
5. Quality control
 - Validity and reliability
 - Test standards
 - Models of fairness and bias

Program evaluation

1. Evaluation of curricula and interventions in the educational context
 - The history of evaluation and the nature of the field (in public policy in general and education in particular)
 - Theories and models of evaluation
 - Possible uses of evaluation

- Identifying interested parties and objects for evaluation
 - The political, social, and cultural context of evaluation
 - Ethics in evaluation research
2. Setting objectives
 - Evaluation of needs
 - Indicators of effectiveness in education
 - Articulation of evaluation questions
 3. Evaluation protocols and evaluation instruments
 - Planning the evaluation protocol
 - Research protocols: experimental protocols, quasi-experimental protocols, use of natural experiments, case studies
 - Quantitative and qualitative research tools
 - Analysis of the evaluation research data in keeping with the research protocol
 4. Evaluation quality
 - Standards in evaluation
 - Bias and subjectivity versus expert judgment
 5. Use of the evaluation findings
 - Reporting the evaluation results to the various audiences
 - Cost-benefit considerations in program evaluation
 - Use of the evaluation findings as a basis for policymaking and educational reforms

2B.2 Itemization and Explanation of the Framework: The Axis of Professions

The purpose of this report is to serve a distinct need of the Israeli education system today, namely, ensuring that workers, old and new, have the knowledge that the education system will need when a comprehensive accountability-based system is instituted. The committee members therefore surveyed the descriptions of the new jobs included in the National Task Force recommendations (school A&E coordinators, assessment officers, officials in the regional education administrations, and evaluators in the National Authority for Assessment and Evaluation) and the new demands made of present workers (teachers and principals). Based on these descriptions and demands, as well as on their own knowledge and experience, the committee members decided what knowledge is necessary to carry out these jobs.

The committee assumes that most of the requirements mentioned below will find their way into the education system in one way or another in the coming years.

National Authority for Assessment and Evaluation

In accordance with the recommendations of the National Task Force, an independent statutory body—the National Authority for Assessment and Evaluation (NAAE)—was established. The NAAE will spearhead assessment and evaluation processes in the education system and will carry out the following tasks:²⁴

- Periodic assessments of the education system (yearly or other, depending on the topic) to review compliance with standards and objectives set by the Education Ministry in the various areas of education, including academics, values, and social issues
- External baseline assessments in every school in Israel, based on standards set by the Education Ministry, with a scope and frequency similar to those of the GEMS tests currently administered in grades 5 and 8. The assessments will look at achievements and other educational indicators (school climate, satisfaction of students and teachers, etc.). The annual report of each school will include data on the school itself and a comparison with the countrywide data and the data of the regional education administration.
- Publication of the findings of the assessment in an annual report, which will then be submitted to the National Education Council, passed on to the Government and the Knesset, and made available to the public
- Responsibility for the development of matriculation exams, professional supervision of the exams, assessment and analysis of their results, and establishment of criteria for authorizing high schools to carry out school-based assessments for matriculation certificates
- Supervision and coordination of participation by the Israeli education system in international research studies and tests
- Professional guidance for the education system in assessment and evaluation
- Development of diverse assessment instruments to be used by educational institutions and regional education administrations
- Evaluation of countrywide projects and programs at the request of the Ministry of Education, in accordance with periodic work plans approved by the NAAE board
- Preparation of guidelines and procedures for routine assessment of programs and projects in the education system

²⁴ National Task Force for the Advancement of Education in Israel, *National Educational Program* (January 2005), p. 169.

Regional Education Administrations

The National Task Force recommended the establishment of regional education administrations (REAs) that would be responsible for all schools and educational activity in their jurisdictions. Once the REAs are established, the schools will be subordinate to them alone. Each REA will oversee, monitor, and assess the resources of the schools in its jurisdiction, the processes that take place in those schools, and the outcomes of the schools' work.²⁵ (At present, these and other functions are the responsibility of the central staff of the Education Ministry, the district offices, local authorities, networks of schools, and so on.)

Each REA will be able to provide professional assessment and evaluation services in accordance with Education Ministry policy, local needs, and the professional guidelines of the NAAE. Each REA will consider establishing an assessment and evaluation unit.²⁶ The functions of the REA will include the following:²⁷

- Participating in setting objectives for the schools in its jurisdiction, approving work plans, and ensuring compliance with the plans (including compliance with the education laws and compulsory guidelines and achievement of academic, behavioral, cultural, and social objectives), while maintaining the principle of school autonomy
- Monitoring and assessing the schools in its jurisdiction
- Overseeing and monitoring the school-based assessment processes designed and carried out in the schools in its jurisdiction (including intensive monitoring of school-based assessments in high schools that count towards matriculation certificates), and providing the schools with assessment and evaluation services

The Teacher's Tasks and Responsibility for Assessment

Teachers "educate by means of knowledge," and most of their work focuses on the classroom setting. Consequently, an understanding of various demonstrations of knowledge in the subjects that they teach and its development is an important component of the teachers' work, and the ability to assess learning and to encourage and assess development and achievement is a professional need. Teachers must therefore be capable of designing and adopting diverse assessment instruments (tests, portfolios, study journals, homework assignments, etc.) with which

²⁵ Ibid., p. 173.

²⁶ Ibid., p. 176.

²⁷ Ibid., pp. 175–176.

they can obtain data for various purposes, such as diagnosing an educational situation in order to map students' baseline knowledge of the topic about to be taught, examining students' proficiency in a topic that has just been taught, or identifying difficulties and misconceptions that have to be addressed on the level of the individual student or the entire class. In addition, teachers must be able to understand reports of their students' external test results and use them to plan and progress in their teaching in class.

In this context, the core teacher of the class must also assemble information received about his/her students from the various subject teachers to form a complete picture of the academic achievement and progress of each and every student.

The Principal's Tasks and Responsibility for Assessment

The school principal, together with the administrative staff and the school's pedagogical council, is responsible for preparing an annual and multiyear list of objectives and indicators, constructing annual and multiyear work plans based on this list, and conducting a regular, transparent process of monitoring and assessing achievement of educational objectives.²⁸

The National Task Force recommends that, at the beginning of each year, the principal inform parents and other interested parties of the objectives, both those required by the Ministry of Education or REA and those set by the school itself. The objectives should be posted, if possible, on the school's Website.²⁹

At least once a year, the principal should report on the school's achievement of the objectives and should publish a work plan for the school; this plan should include measures to improve areas in which the school failed to meet the objectives and should set additional objectives.³⁰ The report should also include the findings of the external assessment provided to the school by the REA, along with the school's comments on the findings and a presentation of the findings of the school's internal assessment.³¹

²⁸ Ibid., p. 120.

²⁹ Ibid., p. 121.

³⁰ Ibid.

³¹ Ibid.

School Information and/or A&E Coordinator

Due to the cardinal importance of in-school assessment, the National Task Force recommended adding a new position in the schools: a school information and/or assessment and evaluation coordinator (A&E coordinator).³²

In order to implement the recommendations of the National Task Force, implementation committees were formed by the Education Ministry administration. The Implementation Committee for Incorporating Internal Assessment in the Schools recommended creating the position of school assessment and evaluation coordinator in a two-stage process—first an “information and feedback coordinator” and then an “assessment and evaluation coordinator.”

The role of the information and feedback coordinator will be to handle the gathering and organization of a variety of data that the schools already have, including the findings of external assessment reports. The coordinator shall produce reports summarizing these data and initiate a discussion of the findings in conjunction with the school administration and relevant interested parties. The coordinator’s reports will be for the use of the school, will be issued in response to routine pedagogical and administrative needs, and will be exclusively internal, unless the administrative staff decides otherwise.

The committee set several professional requirements for information and feedback coordinators, including the ability to do the following:

- Use existing data-management software (school administration software, scheduling software, etc.)
- Incorporate and integrate additional information systems as needed by the school
- Retrieve and update information from national or local databases, as needed by the school or at the request of officials outside the school
- Use information-management software (Excel or Access) to create a basis for gathering and updating information
- Make decisions regarding types of information: what information should be gathered, what is more and less important, how to present the information, what can be done with the information gathered, to whom it is relevant, etc.
- Consistently and systematically use tried-and-true, professional, internal assessment instruments that already exist in the education system³³

³² Ibid., p. 118.

The Implementation Committee believed that the A&E coordinator's job requires skills and knowledge in addition to those described above for the information and feedback coordinator. Therefore, it recommended the establishment of a master's degree program in educational assessment and evaluation. The A&E coordinator shall be responsible for developing and promoting the school's assessment culture and for integrating processes for wise use of the data for decision-making purposes. The A&E coordinator shall be the ultimate professional authority in the school with respect to internal assessment. His/her work shall focus mainly on advising and guiding the school staff in assessing students, processes, and outcomes.

The committee specified the main tasks of the A&E coordinator, on a scale of priorities and feasibility:

1. The coordinator shall be in charge of providing guidance to the school staff and advising it with respect to assessment and evaluation. The coordinator shall be familiar with and skilled in the use of a range of assessment instruments (e.g., portfolios of the individual's work, research papers and lectures, and school tests). The coordinator shall assist the faculty and administration in their regular assessments of classes and subjects. The committee stressed that this recommendation is not meant to detract from the teacher's central role in assessing his/her students, but rather to create a professional, skilled support base.
2. The coordinator shall help with external assessment and with the use of its findings; shall be in touch with authorized external agencies that carry out assessments in the school (regional education center, GEMS, and matriculation exams); and shall read, interpret, and convey information from external reports (GEMS, reports on the five-year plan, immigrant-absorption reports, matriculation-exam reports, etc.). The coordinator shall be acquainted with external-assessment suppliers and external assessment instruments, including the support system to be established, in order to obtain the professional instruments (tests, portfolios, questionnaires, observations, experts, methods, etc.) that the school needs to carry out assessment activities. The coordinator shall be responsible for ordering assessment instruments or commissioning assessment activities, ensuring quality control, and supporting the integration of the instruments and findings.

³³ The instruments currently available to schools include a violence index (Counseling Psychology Service); a database of assignments (Elementary Education Department, together with the Office of the Chief Scientist and Department for Curriculum Planning and Academic Entities); observation aimed at assessment of educational behaviors (Elementary Education Department, together with the Szold Institute); a kit for mapping and evaluating self-direction skills for learning through research (Elementary Education Department, together with Tel Aviv University [Professor Menucha Birenbaum]); a kit for assessing integrative learning processes (Elementary Education Department, together with Tel Aviv University [Professor Aryeh Levy]).

3. The coordinator shall issue reports as needed and as requested by members of the school staff. The coordinator shall also be in charge of issuing periodic reports (including an annual report), either at the request of the administration or at the request of authorized people outside the school. The coordinator shall support the decision-making process using external or internal assessment data, and shall help teachers use the findings of assessments to plan and improve their teaching methods. The coordinator shall help the staffs to come up with plans of action.
4. The coordinator shall give teachers guidance with respect to assessment and shall further the development of the school's assessment practices. He/she will be responsible from a professional standpoint for the quality of the assessments carried out by the school staff and shall help the teachers with the assessment process. In addition, the coordinator, together with the faculty, shall draw up an annual assessment plan for the school. The annual internal assessment plan will be part of the school's activity plan.
5. The coordinator shall perform routine monitoring of important achievements and indicators in the school, such as the level of violence, tardiness and absence, and the school climate, and shall warn of issues in need of attention. The coordinator shall monitor special projects carried out in the school and shall examine their effectiveness and contribution to the extent possible.
6. The coordinator shall initiate and carry out additional assessment activities in the school in conjunction with the school staff and shall bring up matters that should be looked into. In consultation with outside professionals, he/she shall plan the examination, identify and choose appropriate assessment instruments, gather and analyze the data, and prepare reports. The coordinator shall ensure that students' achievements in various educational realms are mapped in the school.
7. The coordinator shall help students carry out assessment activities initiated by them, shall analyze the results together with the students and teachers, and shall explain the assessment activities carried out in the school. When necessary, he/she shall present findings to the students.

3. Sample Curricula for Academic and Professional Development Programs Based on the Framework

3A. Teachers: Proposal for Training and Continuing Education in Assessment and Evaluation

In the present context, teachers are in a unique situation: although assessment is not their main job (as it is for the A&E coordinators and psychometricians), it affects many aspects of their main job. Consequently, it is extremely important to give teachers a decent foundation in the field of assessment, even if it is sometimes difficult to isolate the teacher's job as an evaluator from his/her other functions. The three courses proposed here are intended to provide such a foundation. This foundation can also serve as a basis for more advanced studies for those teachers who wish to train as A&E coordinators.

Introductory Course

- Approaches to assessment and the underlying concepts of education
- Assessment functions and audiences
- Ethical considerations in assessment
- Assessment for learning—formative
- Assessment of learning—summative
- The relationship between assessment on the one hand and motivation and learning strategies on the other
- Assessment of teachers: essence, principles, and functions—formative and summative assessment
- Evaluation of curricula and projects in the educational context

Basic Course in Statistics and Assessment

- Descriptive statistics (types of variables, frequency distribution, measures of central tendency, measures of variability, comparative grades, measures of association, linear regression)
- Assessment quality: validity, reliability and assessment error, response bias (test anxiety, social desirability, test-wiseness), fairness
- Types of tests (criterion-referenced tests, norm-referenced tests, and diagnostic tests)
- Grading methods (based on norms, criteria, or individual progress)
- Item analysis
- Findings of assessments as a basis for planning teaching (including reading external reports)
- Computer uses in assessment

Course in Classroom Assessment

- Standards for school subjects
- Setting objectives for internal assessment
- Familiarity with various assessment instruments (tests, assignments, study journals, portfolios, observations) and the appropriate context for their use (including knowledgeable selection and use of prepared instruments)
- Design of assessment instruments (including sampling of fields of content and skills, rules for constructing and wording assignments and tests, and types of test items), analysis and interpretation of results
- Development of scoring rubrics
- Adaptation of assessment instruments for special-needs students

Notes:

- The assumption is that the courses are suitable for the final years of teacher-training in universities and colleges of education.
- The first two courses are prerequisites for the third.
- The third course, or at least most of it, must refer constantly to the subjects taught by the teachers.
- Practicing teachers who have not trained in assessment can take courses of this sort as in-service courses.

3B. Principals: Proposed Curriculum in Assessment and Evaluation

The proposed curriculum is based on the school principal's job and responsibility for assessment, as described in the report of the National Task Force.³⁴

Below is the list of proposed courses in these subjects for principal-training programs and for professional development programs for principals.³⁵

³⁴ National Task Force for the Advancement of Education in Israel, *National Educational Program* (January 2005), pp. 51-60.

³⁵ The list is divided into required courses and electives. The required courses include core topics in assessment, statistics, and evaluation; the electives include administrative aspects related to assessment. It is proposed that the assessment module in a principal-training program include only the required courses, since the topics of the electives will presumably be covered in other courses in the program. Both the required courses and the electives should be included in an independent program of study in assessment for principals.

Required Courses

1. Introduction to assessment
 - Theory of assessment as a discipline (basic concepts in assessment)
 - Planning and design of assessment instruments (on the theoretical level)
 - Scoring
 - Quality control
 - Assessment ethics
2. Introduction to statistics
 - Descriptive statistics
 - Types of variables, frequency distribution, organization of data and various ways of presenting data, measures of central tendency, measures of variability, comparative grades, measures of association
 - Reading assessment reports
3. Program evaluation
 - Program evaluation in the educational context
 - Theories and models of evaluation
 - Indicators of effectiveness in education
 - Evaluation of needs
 - Setting objectives
 - Articulation of evaluation questions
 - The evaluation system and evaluation instruments
 - Evaluation quality
 - Use of evaluation findings
4. Classroom assessment
 - Assessment in the educational context
 - Setting objectives
 - Assessment instruments
 - Assessment quality
 - Summarizing and reporting
 - Reporting results to various audiences
5. Large-scale assessment
 - Objectives of external assessment

- Evaluation utilization
- Familiarity with state tests (national feedback tests, GEMS, matriculation exams) and international tests (e.g., PISA and TIMSS), including research reports
- Effects and ramifications of standardized tests
- Issues in external assessment
- Expansion of the use of school administration software to generalize assessment findings

Electives

6. Assessment of the faculty
 - Essence and principles of assessing teachers
 - Assessment of teachers in various education systems
 - Standards and criteria of good teaching and of good teachers
 - Methods and instruments for assessing teachers
 - Possible uses of assessment of teachers
7. Making system-wide decisions
 - Setting assessment objectives derived from state, regional, and school goals
 - Drawing conclusions about the degree to which the objectives were achieved, designing work plans, and allocating resources accordingly
8. Assessment mechanisms in an educational organization
 - Foundations of managing assessment data in a way that allows for long-term follow-up
 - Foundations of gathering and disseminating assessment information
 - Settings for discussing aspects of assessment (dialogue between external and internal assessment)
 - Making system-wide decisions on aspects of assessment

3C. Assessment and Evaluation Coordinators: Proposed Curriculum in Assessment and Evaluation

As recommended by the Implementation Committee on Incorporating Internal Assessment in the Schools, A&E coordinators should have master's degrees. A&E coordinators shall have one of the following degrees and qualifications:

1. A master's degree in assessment and evaluation with a specialization in school assessment
2. A master's degree in another field, successful completion of the required courses in the master's degree program in assessment and evaluation, and completion of a field research project. (The required courses are listed in the sample master's degree curriculum in section 3D.)

A&E coordinators will have an extremely demanding job. They will have to contend with complex theoretical material as well as difficult practical problems. The main difficulty concerns planning assessment systems for the organization of which they are a part. In the case of internal assessments, they must gather the data, and in all cases they must interpret data and explain them to other teachers and to the principal. In order to earn the trust of their colleagues and supervisors, they must demonstrate professional knowledge and inspire confidence in their use of the assessment instrument and interpretation of the findings. Another difficulty has to do with the constant need to contend with whatever tension may exist between theory and practice; between professional requirements and the ability to comply with them in practice; and between the need to convey the findings faithfully and fear of harming students and teachers.

It is therefore important that the training of A&E coordinators include a large dose of practical experience. It is proposed that students in this track carry out a supervised assessment project in a school. Because group work is beneficial, the project can take the form of a regional assessment project by a group of students carried out under academic supervision in several schools.

In order to address actual needs, it is recommended that the aforementioned required courses be taught in two successive summers. During the academic year between the two summers, the students in this track can carry out their practical assessment project. Each summer, they will earn ten credits (140 hours per summer), and the supervised project can give them another six credits.

3D. Master's Degree: Proposed Curriculum in Assessment and Evaluation

Introduction and Guiding Principles

There is already a severe shortage of professionals capable of carrying out educational assessment and evaluation functions on the countrywide, district, and school levels. There is also a shortage of academic programs focusing on assessment and evaluation in educational contexts. On the assumption that the need for skilled personnel will increase, appropriate academic programs must be developed. The outline proposed for such a program is based on several assumptions or principles regarding the skills and knowledge required by professionals in order to work thoughtfully and critically in educational assessment and evaluation:

1. Assessment and evaluation issues cannot be studied out of context. Students in all the tracks must take courses on the educational context, in which they will be introduced to the main methods, distinctions, agreements, and disagreements in this field (external and internal assessment, summative and formative assessment, standard and alternative assessment, explicit and implicit goals of methods and of various programs, impacts on outcomes and educational

processes, etc.). In addition, theoretical and applied courses in statistics and research methods (including seminars) will be solidly based in the field of education.

One component of the “educational context” is the knowledge domain (i.e., the school subject); much of educational assessment is based on the subjects taught. It is proposed that all students be introduced to this component so that they can contribute to a discussion between writers of tests in the various subjects and assessment researchers.

2. In addition, it seems that no hard-and-fast distinction can be made between producers and consumers of information: just as the “producers” (evaluators outside the school) must be furnished with knowledge about the “consumers” (the education system), the people who are going to fulfill assessment functions in the school must be given the knowledge they need not only to understand, disseminate, and request information but also to produce and evaluate assessment and evaluation instruments. It is therefore proposed that the required courses on methods of assessment, evaluation, and research be included in all the tracks (with slight modifications).

Another assumption underlying the proposed outline is that different audiences and functions have different needs. Therefore, three specialization tracks are proposed: psychometrics, school-based assessment, and program evaluation. The three share a common core of required courses. Each of them can be completed with or without a thesis.

Required Courses

Total of 280 academic hours = 20 credits

1. Statistics course, including the “medium level” topics (see pp. 24–25)—84 hours (6 credits)
2. Course on qualitative research methods (topics listed on p. 24)—28 hours (2 credits)
3. Introductory course on educational assessment, including general background, assessment theories, planning and development of assessment instruments, scoring, and quality control (see p. 25)—42 hours (3 credits)
4. Course on classroom assessment (topics listed on pp. 22–23)—42 hours (3 credits)
5. Introductory course on program evaluation (topics listed on pp. 25–26)—42 hours (3 credits)
6. Course on assessment and evaluation in the educational context—42 hours (3 credits)

Specialization in Psychometrics

This track is intended mainly to train professionals for the NAAE or district level. These people are expected to deal with external assessment and to develop standardized instruments for assessing achievement and other dimensions relevant to the education system (e.g., school climate). The curriculum therefore consists mainly of advanced courses in psychometrics. However, in order to

familiarize students with concepts in school-based assessment, one advanced course on this subject is proposed. It should be noted that almost all the courses will be either lectures with sections or seminars.

1. Additional required courses: total of 140 academic hours = 10 credits

Advanced statistics, including the “advanced level” topics (see p. 26)—56 hours (4 credits)

Choice of two of the following:

- a. Advanced course in item response theory (IRT)—42 hours (3 credits)
- b. Advanced course in generalizability theory—42 hours (3 credits)
- c. Large-scale assessment—42 hours (3 credits)

2. Electives: total of 112 academic hours = 8 credits

Two seminars with seminar papers (e.g., bias and fairness in the use of tests, cost-benefit considerations in the use of tests, grade calibration problems, differential item functioning (DIF), adaptive tests, assessment in international contexts, development of standardized assessment instruments)

3. Choice of one of the following: total of 42 academic hours = 3 credits

- a. Alternative assessment, including setting goals, setting standards, designing achievement tests (sampling questions and tasks), and assessing portfolios—42 hours (3 credits)
- b. Assessment in the school as an organization—42 hours (3 credits)

Thesis or supervised applied research project

Total of 574 academic hours = 41 credits

Specialization in School-Based Assessment

This track is intended mainly to train school A&E coordinators. These people must be knowledgeable consumers of external assessment data and, in particular, must be able to develop school assessment instruments and advise the school administration and faculty on all aspects of assessment and evaluation.

1. Required courses: total of 280 academic hours = 20 credits

Note: Because this track is especially suited to teachers and A&E coordinators, a basic statistics course may be sufficient.

2. Alternative, internal, and school-based assessment: total of 140 academic hours = 10 credits

- a. Alternative assessment, including setting goals, setting standards, designing achievement tests (sampling questions and tasks), and assessing portfolios—42 hours (3 credits)

- b. Assessment in the school as an organization—42 hours (3 credits)
 - c. Choice of practical seminar: developing instruments and an assessment plan for a subject, for teachers, or for an organization; thoughtful integration of external and internal assessment—56 hours (4 credits)
- 3. Use of assessment for improvement and development in school: total of 140 academic hours = 10 credits**
- a. Motivational ramifications of assessment—28 hours (2 credits)
 - b. Skills in providing feedback in school—28 hours (2 credits)
 - c. School, system, and community (ramifications of assessment for the school’s status, resources, relations with parents, etc.)—28 hours (2 credits)
 - d. Practical seminar: evaluation of the impact of assessment (on achievement, traits of the organization, and relations with various people; assessment as a stimulus for creating a learning organization)—56 hours (4 credits)

Thesis or applied research project³⁶

Total of 560 academic hours = 40 credits

Specialization in Program Evaluation

This track is intended mainly to train professionals who will be evaluating educational programs and interventions on the countrywide or district level and on the school level.

Courses in the track: total of 224 academic hours = 19 credits

- a. Advanced statistics (see previous page)—56 hours (4 credits)
- b. Research protocols: experimental protocols, quasi-experimental protocols, use of natural experiments—56 hours (4 credits)
- c. Formative assessment—42 hours (3 credits)
- d. Practical research seminar in evaluation of educational programs, including fieldwork—56 hours (4 credits)
- e. Seminar in assessment methods, such as time-series analysis, regression discontinuity, and natural experiments—56 hours (4 credits)

Thesis or applied research project

Total of 588 academic hours = 42 credits

³⁶ Students who write two practical seminar papers may be exempt from a final project.

4. Recommendations

Insofar as improvements in educational achievement depend in part on the development of an ongoing assessment culture and routine assessment and evaluation mechanisms, the need for training a large number of professionals is critical. The quality of the programs offered in various settings in the coming years will be of decisive importance in determining the professional skills of education practitioners and researchers.

The committee found a shortage of professionals, a deficiency in professional knowledge among the various workers, and a shortage of assessment researchers and test developers, relative to the needs and opportunities.

The proposed framework is aimed at ensuring the quality of curricula. Nevertheless, it seems that the development of a training system and programs of study is being hindered by teaching and training capabilities and resources. Apparently, there are not enough instructors and professionals in Israel who can train and teach potential students in a reasonable amount of time. Furthermore, there is a shortage of study materials and mechanisms for professional discussion of the accrued experience.

4A. Creation of a Support System for the Development of Researchers and Practitioners

1. The committee recommends offering **intensive courses for teachers and lecturers** from the various universities and colleges so that they can teach in-service (professional development) courses for future professionals in the education system. Renowned foreign experts, such as those who spoke at the seminar, could also be invited to teach some of these courses.³⁷
2. The committee asks the Council for Higher Education to express its opinion on the need for academics and professionals in this field and to propose ways of increasing **the number of researchers in this field**, while maintaining the quality of their training.
3. The committee considers it appropriate to develop programs of study suitable for a **diverse, geographically scattered population**, especially by offering **online courses** in a suitable (academic or Education Ministry) setting, while ensuring supervision and mentoring in the various stages of the programs.

³⁷ For the names of the seminar participants, see Appendix 1.

4. The committee recommends encouraging the **creation of a virtual journal on a high professional level**. This journal would publish assessment-related research, discussions of diagnostic and assessment instruments, and abstracts of Israeli research studies published in foreign journals. The committee proposes encouraging scholars to share data and findings on various intervention programs, so as to encourage further research and discussion among scholars and practitioners.

4B. Further Development Required

1. The committee found a shortage of instruments that can be used by professionals, such as countrywide instruments for assessing instruction and teachers; monitoring instruments for school principals; and “banks” of school assignments, test items, and diagnostic items. The committee recommends that the NAAE, together with the expert institutions, undertake to **develop these instruments**, which are essential for efficient, beneficial work by teachers, principals, A&E coordinators, and officials in the top administrative echelons of the Ministry.
2. The committee found that assessment and the use of its findings is a growing field worldwide and involves important methodological and organizational questions. Whether we are talking about value-added measurements of student achievement versus assessments of cohort-to-cohort improvement, or research instruments for monitoring teaching and learning for the purpose of cost-benefit analyses of specific interventions, there is much room for development and innovation. It would seem that Israel, which has a good reputation for teaching math, statistics, and information science, can contribute to and benefit from active participation in assessment research. *The committee recommends that the universities, led by the Council for Higher Education, encourage promising scholars to specialize in these fields.*

4C. Monitoring of Activity Surrounding the Framework

The attempt to combine the development of academic material with the ability to take action cannot be taken for granted. The committee recommends devising indicators for the success of the framework as a basis for improving it. The success of the alumni of various professional development and advanced degree programs is an important element in evaluating the success of the program.

5. Appendices

5A. Appendix 1: Seminar Schedule

Sunday, Hanukkah, 12 December 2004, Van Leer Jerusalem Institute

8:30–9:00 Registration

9:00–11:15 First Session: The Purview of the Discussion

Chair: *Dr. Yoav Cohen*

Greetings: *Prof. Dan Shechtman*

1. Keynote Address: Measurement and Evaluation in Education Systems Today: Trends and Main Issues
Prof. Richard Shavelson
2. Measurement and Evaluation Needs in the Wake of the National Task Force for the Advancement of the Education System (Dovrat Commission)
Mr. Shmuel Har Noy
3. Measurement and Evaluation in Education: Understandings and Misunderstandings in the Israeli Discourse
Prof. David Nevo
4. Measurement, Evaluation, and the Goals of Teaching and Learning: Conflicting Goals and Hidden Prices
Prof. Ruth Butler
5. Summation

11:15–11:35 Coffee break

11:35–2:00 Second Session: What Do the People Who Define and Evaluate Educational Achievement Have to Know?

Or, Classroom and Large-Scale Assessment as Complementary Tools

Chair: *Prof. Menucha Birenbaum*

1. What Did Teachers, Principals, and Ministry Officials Learn When the GEMS (**G**rowth and **E**ffectiveness **M**easures for **S**chools) System Was Nationally Introduced?
Presenter: *Dr. Gili Schild*
Respondent: *Ms. Flori Vedrovnik*
2. What Did Teachers, Principals, and Ministry Officials Learn When the Matriculation Exams Were Replaced with Alternative Evaluations?
Presenter: *Ms. Mirit Lavi*
Respondent: *Dr. Rita Sever*
3. Developing Professional Monitoring Systems: Improving Education through Feedback to Teachers and Schools
Prof. Peter Tymms
4. Summation

2:00–2:40 Break (light lunch)

2:40–4:20 Third Session: Bridging the Gap between Classroom and Large-Scale Assessment

Chair: *Dr. Alia Sheety*

1. A “Smart” Standardized Test and How It Relates to Teaching-Learning: The Case of the Biology Matriculation Exams
Dr. Bruria Agrest
2. “School-Based Evaluation”: An Exemplary Case of a School-Based Work Plan and How It Relates to State Tests
Ms. Eilana Yarchy
3. Bridging the Gap between Classroom and Large-Scale Assessment: Lessons from a Workshop Held by the National Research Council/National Academies
Prof. Michael Feuer

4:20–5:00 Fourth Session: Tools for Gathering, Processing, and Reporting Data: The Case of the NAEP Data Tools

Prof. Michal Beller

Wednesday, 19 January 2005, Van Leer Jerusalem Institute

9:00–9:30 Registration

9:30–10:15 First Session: Post-Graduate Programs—Educational Evaluation Studies

The rationale for the Tel Aviv University curriculum (a survey of examples from around the world will be prepared and distributed in advance)

Prof. Menucha Birenbaum

10:15–12:30 Second Session: What Is Needed and What Can Be Done to Ensure a High-Quality Measurement and Evaluation System: Learning from Experience and Ramifications for the Future

Chair: *Prof. Dan Koretz*

Introduction: Lessons from the Health System

Prof. Gad Rabinowitz

Brainstorming by a team of experts: Given opportunities to improve practice and research in evaluation and assessment in education, what are the present strengths, what is lacking, and what practical steps should be taken?

Participants:

Prof. Rami Benbenishty

Dr. Yoav Cohen

Prof. Ruhama Even

Dr. Tali Freund

Prof. Baruch Nevo

12:30–1:15 Break (light lunch)

1:15–4:15 Third Session: New Trends in Educational Evaluation Research

Chair: *Prof. Yaacov Katz*

1. Indicators of Teaching-Learning: The Relationship between Resources and Achievement
Prof. David Cohen

2. Measuring Improvement
 - a. Presentation and Methodological Questions
Prof. Dan Koretz
 - b. The Added-Value Methodology
Dr. Haggai Kupermintz
3. American Education Research Changes Tack
Prof. Joshua Angrist

4:15–5:00 Fourth Session: Conclusion

Chair: *Prof. Gershon Ben-Shakhar*

Remarks: *Ms. Ronit Tirosh*, Director General of the Ministry of Education, Culture, and Sport

Summation: Drafting Curricula—Understandings and Challenges

5:00–5:15 Refreshments and informal discussion

5B. Appendix 2: Seminar Speakers

Bruria Agrest is a lecturer in science education and in the training of biology teachers at the Bar-Ilan University School of Education. She has a doctorate in science education from the Hebrew University of Jerusalem.

Rami Benbenishty is a professor and the head of the research group on mental health and well-being in childhood and adolescence in the School of Social work and welfare in The Hebrew University of Jerusalem. He has a Ph.D. in social work and psychology from the University of Michigan.

Ruhama Even is an associate professor at the Weizmann Institute of Science and head of the mathematics group in the Department of Science Teaching. She holds a Ph.D. in mathematics education from Michigan State University.

Shmuel Har Noy heads the Implementation Staff for Educational Reform and was the coordinator of the National Task Force for the Advancement of Education in Israel (“Dovrat Commission”). He has a master’s degree in public health from the University of Maryland and a master’s degree from the Department of Physiology and Pharmacology at Tel Aviv University. He is also a graduate of the Mandel Leadership institute in Jerusalem.

Yaacov Katz is the chair of the Pedagogical Secretariat in the Ministry of Education, Culture, and Sport and a faculty member at the Bar-Ilan University School of Education. He holds a Ph.D. in educational psychology from Witwatersrand University in Johannesburg, South Africa, and is a member of the Initiative’s steering committee.

Haggai Kupermintz is an assistant professor in the Faculty of Education at the University of Haifa. He is a consultant and participant in various projects assessing students and teachers and evaluating educational programs. He holds a Ph.D. in educational psychology and an M.S. in statistics from Stanford University.

Mirit Lavi is a teacher of literature, Hebrew language, and expository writing and currently serves as a subject coordinator at ORT Motzkin Comprehensive High School. She was the school’s Matriculation 2000 project coordinator from September 1994 until the project came to an end. Ms. Lavi has a bachelor’s degree in literature, Hebrew language, and Bible and a master’s degree in education, journalism, and communication.

Victor Lavy is a full professor in the Department of Economics at the Hebrew University of Jerusalem and director of the Falk Institute for Economic Research in Israel. His primary research interests are the economics of education and of work and development. Professor Lavy, who holds a Ph.D. in economics from the University of Chicago, is a member of the National Task Force for the Advancement of Education in Israel.

Baruch Nevo heads the Intelligence Research Laboratory in the Department of Psychology at the University of Haifa and is president of Zefat College. He is a member of the Committee on Learning Disabilities in Higher Education (Planning and Grants Committee), the research committee of the National Institute for Testing and Evaluation, and the academic committee for the GEMS tests. Professor Nevo has a Ph.D. in psychology from the Hebrew University of Jerusalem.

David Nevo heads the Tel Aviv University School of Education. His major interests are evaluation theory, program evaluation, school-based evaluation, student assessment, and the pedagogical and social consequences of educational evaluation.

Gad Rabinowitz is an associate professor in the Department of Industrial Engineering and Management at Ben-Gurion University of the Negev. He earned a B.Sc. and M.Sc. from Ben-Gurion University and a Ph.D. in operations research from Case Western Reserve University.

Gili Schild has headed the Evaluation and Measurement Department in the Bureau of the Director General of the Ministry of Education, Culture, and Sport since 1991. In this capacity, she developed the GEMS system and evaluates the main programs run by the Ministry. She has a doctorate in social psychology from the Hebrew University of Jerusalem.

Rita Sever is the Director of Planning, Evaluation and ICT at the Authority for Education, Culture and Sport of the Municipality of Tel Aviv–Yafo. She has a Ph.D. in cognitive social psychology from the Hebrew University of Jerusalem.

Alia Sheety is the principal of Sisters of Nazareth High School in Haifa and an adjunct lecturer in the Department of Education in Technology and Science at the Technion. Dr. Sheety has a Ph.D. in education from Arizona State University.

Flori Vedrovnik is the vice-principal in charge of ORT Shein Junior High School in Rehovot. She has twenty-four years of teaching experience. Ms. Vedrovnik has a bachelor's degree in education and is completing a master's degree in educational administration.

Eilana Yarchy founded the Uziel State-Religious School in Beit Shemesh and has been its principal for eleven years. The school has won two national awards—the Jerusalem District Principal's Prize and the Environment Prize—and was nominated for the Education Prize. Ms. Yarchy has a master's degree in Jewish studies from Touro College.

Guest Speakers:

Joshua Angrist is a professor of economics at the Massachusetts Institute of Technology and a research associate in programs on children, education, and labor studies at the National Bureau of Economic Research. He previously taught at the Hebrew University of Jerusalem. He is a fellow of the Econometric Society and a co-editor of the *Journal of Labor Economics*. Dr. Angrist has a long-standing interest in public policy. He holds a B.A. from Oberlin College and an M.A. and Ph.D. from Princeton University. He also spent some time at the London School of Economics as an undergraduate and at the Hebrew University of Jerusalem as a master's student.

David K. Cohen is John Dewey Collegiate Professor of Education and Walter H. Annenberg Professor of Education Policy at the University of Michigan. His current research interests include educational policy, the relations between policy and instruction, and the improvement of teaching. With Professors Brian Rowan and Deborah Loewenberg Ball, he co-directs the Study of Instructional Improvement, a large longitudinal study of efforts to improve instruction and learning in reading/language arts and mathematics in high-poverty elementary schools. See <http://www.sii.soe.umich.edu/about.html>.

Michael J. Feuer is the executive director of the Division of Behavioral and Social Sciences and Education (DBASSE) in the National Research Council (NRC) of the National Academies, where he was formerly the director of the Center for Education and director of the Board on Testing and Assessment. He holds a Ph.D. in public policy from the University of Pennsylvania, where his research focused on mathematical modeling and human resource planning in organizations. He also has an M.A. from the Wharton School and has studied at the Hebrew University of Jerusalem and the Sorbonne.

Daniel Koretz teaches educational measurement at Harvard University. His interests include education policy and the social context of education, as well as assessment. His most recent work has focused on the development of psychometric methods for evaluating score inflation and the application of value-added modeling to educational performance. Professor Koretz has a Ph.D. in developmental psychology from Cornell University.

Peter Tymms is director of the CEM Centre, at Durham University which involves thousands of primary and secondary schools across the UK and beyond. It is the largest educational research group in a UK university. Professor Tymms taught in a wide variety of schools from Central Africa to the northeast of England before starting an academic career. He has a Ph.D. from Cambridge University.

5C. Appendix 3: Examples of Assessment and Evaluation Programs at U.S. Universities

Undergraduate Programs:

1. Nebraska Assessment Cohort

Course title: Making Sense of Assessment: Large Scale Assessment & Classroom Based Assessment
For more information, see <http://cehs.unl.edu/hpark/nac/nac03syl.doc>.

2. Nebraska Assessment Cohort

Course title: Leadership for Learning: Leadership for Excellence in Assessment and Classroom Based Assessment
For more information, see <http://cehs.unl.edu/hpark/nac/nl103syl.doc>.

3. Harvard School of Education

Course title: Understanding Today's Educational Testing
The course is meant for students with no statistical training and has no prerequisites. For more information, see <http://www.gse.harvard.edu>.

4. University of Washington, College of Education—Measurement, Statistics and Research Design

Course title: Introduction to Educational Measurement
This course is geared towards school administrators, curriculum specialists, counselors, and other educators who need to understand measurement at a conceptual level and who need skills in evaluating external tests and testing programs. For more information, see http://www.depts.washington.edu/coe/programs/ep/deg_programs/msrd/mstrand.html.

Advanced Degree Programs:

1. University of Washington, College of Education—Measurement, Statistics and Research Design

- a. **Course title:** Advanced Educational Measurement—for Ph.D. students
- b. **Course title:** Classroom Assessment Strategies
- c. **Course title:** Instrument Development—for M.A. and Ph.D. students

For more information, see http://www.depts.washington.edu/coe/programs/ep/deg_programs/msrd/mstrand.html.

2. Harvard Graduate School of Education

- a. **Course title:** Methods of Educational Measurement
- b. **Course title:** Theories and Methods of Child/Adolescent Cognitive and Psychological Assessment

For more information, see <http://www.gse.harvard.edu>.

3. UCLA—Research Methodology Division, Department of Education

For more information, see <http://www.gseis.ucla.edu/division/srm/maflyer.html>.

4. University of Minnesota, College of Education and Human Development, Educational Policy and Administration

For more information about the master's program curriculum, see <http://education.umn.edu/edpa/evaluation/MAcurric.html>. For more information about the Ph.D. program curriculum, see <http://education.umn.edu/edpa/Evaluation/PhDcurric.html>.

5. University of Colorado School of Education

For more information about the Ph.D. program curriculum, see <http://www.colocorado.edu/education/programs/REMphd.html>.

6. Utah State University, Department of Psychology, Research and Evaluation Methods Program

For more information about the Ph.D. program curriculum, see www.coe.usu.edu/psyc/rem/program/phdedtrack.htm.

7. Penn GSE: Degrees and Programs—Policy Research, Evaluation and Measurement

For more information about the Ph.D. program curriculum, see <http://www.gse.upenn.edu/pdf/premoreview.pdf>.

8. West Michigan University

This university has two programs:

- a. An “evaluation, research and measurement” program in the School of Education for Ph.D. students. For the full program, see <http://www.wmich.edu/edstudies/emr/phd.htm>.
- b. An interdisciplinary program in the Evaluation Center. For the full program, see <http://www.evaluation.wmich.edu/phd/courses/index.htm>.

9. Cornell College of Human Ecology, Department of Policy Analysis and Management

For more information about the Ph.D. program curriculum, see <http://www.human.cornell.edu/units/pam/grad/sample.cfm>.

10. Stanford MA Policy Analysis and Evaluation Program

For more information about the curriculum, see <http://www.stanford.edu/~davidf/evaluation.html>.

5D. Appendix 4: Members of the Committee on Assessment and Evaluation

Michal Beller is a principal research scientist at Educational Testing Service (ETS) in Princeton, New Jersey, and a senior research director with responsibility for the Assessment and Evaluation Research Cluster, which includes the Assessment Design and Scoring, New Constructs, and Validity Research centers. The aim of the cluster is to utilize ETS's core competency—measurement and assessment—in the service of learning. Before joining ETS, Professor Beller held a faculty position as associate professor in the Department of Education and Psychology at the Open University of Israel. She has a Ph.D. in psychology from the Hebrew University of Jerusalem.

Gershon Ben-Shakhar is the president of the Open University of Israel. Before joining the Open University, Professor Ben-Shakhar served as the pro-rector and dean of the Faculty of Social Sciences at the Hebrew University of Jerusalem. His main research interests are in the fields of cognitive psychophysiology, psychological testing and personnel decisions, and the confirmation bias in experts' judgments. He holds a Ph.D. in psychology from the Hebrew University.

Menucha Birenbaum is a professor at the Tel Aviv University School of Education, where she heads the research methodology, measurement, and evaluation program. Her main research interests are the analysis of learning interactions (in face-to-face and distance education); instruction-learning-assessment cultures that promote self-regulated learning; and large-scale diagnostic assessment. Professor Birenbaum has a Ph.D. in educational psychology from the University of Illinois at Urbana-Champaign.

Ruth Butler is an associate professor at the School of Education of the Hebrew University of Jerusalem and a former head of the school. Her main research interests are students' motivation for schoolwork; the consequences of various types of feedback and assessments on students' motivation; self-evaluative strategies and judgments; and teachers' motivation. Professor Butler has a Ph.D. from the Hebrew University of Jerusalem.

Yoav Cohen is the executive director of the National Institute for Testing and Evaluation (NITE). His main research interests are computer-based testing, automated essay scoring, the theory and practice of grading, and international testing programs. He has a Ph.D. in cognitive psychology and a master's degree in computer science from the University of Oregon.

Tali Freund is the vice-president for planning, information, and evaluation at the Center for Educational Technology (CET). Before joining CET, Dr. Freund established and directed the Planning and Evaluation Unit in the Education and Culture Administration of the Municipality of Tel Aviv–Yafo. She also played a leading role in the Assessment and Evaluation Unit in the Ministry of Education, where she was an active member of the GEMS development team. She has a Ph.D. in social psychology from Tel Aviv University.

Fadia Nasser Abu al-Hija is a senior lecturer at the Tel Aviv University School of Education. She previously served as a research coordinator for the GRE study conducted by the Educational Testing Service (ETS) in Princeton, New Jersey. Her main research interests are achievement assessment with a focus on comparative studies, teacher and teaching evaluation, and construct validity (methodological studies). Dr. Nasser Abu al-Hija holds a Ph.D. in research, evaluation, measurement and statistics (REMS) from the University of Georgia.

Gina Shimborsky works at the Center for Matriculation Examinations at the Henrietta Szold Institute. She is responsible for the coordination of committees that write the examinations and is involved in test development and analysis and in studies pertaining to test results. She has an M.A. in developmental psychology from Northwestern University.