

Moshe Zeidner
University of Haifa

School-Based Assessment Research in Israel: Current state and future directions

Abstract

The major aim of the present paper is twofold: (1) to survey and evaluate the present "state of the art" in Israeli school-based assessment research; and (2) to delineate potential future directions, foreseeable developments, and needed research. The paper focuses primarily on the assessment of aptitude, achievement, and personality in the schools and highlights a number of key issues engaging the interests and concerns of the Israeli school psychology and measurement community (i.e., group differences in test performance; test bias and equity; and testing the selection policy in the schools). There is currently in Israel a paucity of well developed, soundly normed, and psychometrically validated (or adequately adapted) instruments available for assessment purposes – particularly in the area of non-cognitive assessment. Furthermore, there is a need for the development of criterion-referenced measures in a variety of school subjects for student evaluation and achievement tracking purposes in both the Jewish and Arab sectors. This paper reveals the need for a major concerted effort to improve both the quantity as well as the quality of available school-based assessment instruments, procedures, and practices.

Résumé

L'objectif principal de cet article est double: 1) sonder et évaluer la situation actuelle des recherches sur l'évaluation en milieu scolaire en Israël; et 2) délimiter les orientations futures, les développements prévisibles et les recherches nécessaires. L'article porte essentiellement sur l'évaluation des aptitudes, des résultats et du caractère dans les écoles et il fait ressortir un certain nombre de problèmes clés qui présentent de l'intérêt pour les psychologues scolaires israéliens (différences de groupe dans l'exécution des tests; partialité et équité des tests; et évaluation de la politique de

sélection dans les écoles). On constate actuellement en Israël une pénurie d'instruments d'évaluation intelligemment conçus, aux assises solides et validés selon des moyens psychométriques, notamment dans le domaine de l'évaluation non cognitive. En outre, on constate le besoin d'établir des mesures à base de critères dans diverses matières enseignées à l'école pour évaluer les élèves et contrôler leurs résultats dans les communautés juive et arabe. Cet article fait ressortir le besoin d'un effort concerté pour améliorer la quantité et la qualité des instruments, des modalités et des pratiques d'évaluation en milieu scolaire.

Israel, much like the United States, may be aptly depicted as a test-oriented and test-consuming society, with tests widely employed by the educational system, military, industry, and government as an aid to making decisions about people. Tests have a particularly formidable presence in the school system and are employed for a variety of functions, including: (a) research and program evaluation; (b) vocational guidance and counselling; (c) student selection, classification, and placement (e.g., entrance, streaming, homeroom class-formation, within-class academic tracking); (d) screening and diagnosis; and (e) identification and selection of children for special programs (e.g., gifted children's program, special education classes).

General historical background information

Prior to the establishment of the State of Israel in 1948, student populations were fairly homogeneous, composed largely of children from Jewish middle-class families of European origin (about 90%). Since the inception of the State of Israel, the student body has changed radically in content and scope, with the student population growing seven-fold since 1943. Mass immigration from over 70 countries, including post-war Europe and the Middle East, doubled the population over the first three years of statehood and tripled it over the first twelve (Adler, 1984).

During the early years of statehood, the educational system in Israel was charged with accommodating an increasing percentage of Jewish children from Eastern cultures (e.g., Morocco, Algeria, Yemen, Egypt, Persia), many of whom presented the classical symptoms of cultural deprivation, including inadequate mothering and bonding; low degree of mediated learning experiences; and reported sensory, linguistic, and conceptual deficits (Feuerstein, 1980; cf. Raviv, 1989). As the State absorbed more and more immigrants, the school system faced the problem of greater heterogeneity in its student body. This variety manifested itself in both family structure and cultural-educational orientation, traditional customs, and behaviour patterns, as well as students' achievement profiles and motivation. Whereas immigrant

students of European extraction adapted quite readily to modern Israeli society and its Western-oriented school system, students of Eastern background evidenced considerable difficulty in the school acculturation process; they tend to regard attainment of education in a more instrumental way and were prepared to invest less in education as a means of acquiring status (Kfir, 1988). Since students of Eastern origin had lower levels of achievement aspiration compared to their Western counterparts, they were also destined to be less mobile within a society that holds modern status-attainment norms (Adler, 1984; Kfir, 1988).

It was the educational system that assumed the main burden and responsibility for equalizing opportunities and integrating culturally different groups into Israeli society, no mean task (Adler, 1984). Furthermore, it was largely school psychologists – whose original function in Israel was to assess and diagnose failing students – who were charged with assessing and placing the masses of immigrant children. Unfortunately, few standardized measures were available at the time for assessment functions; the exigencies of the times dictated the use of measures that were often not adequately translated, adapted, normed, or validated for their culturally different target populations. Only in the past two decades or so have serious efforts been made to construct standardized culturally-indigenous instruments appropriate for the Israeli scene or in adequately norming conventional measures on representative school populations in Israel.

Major goals and scope of this paper

School-based assessment and evaluation is currently one of the key research areas in Israeli school psychology (Raviv, 1988). Indeed, the history of school psychology, both in the United States and in Israel, is in many ways synonymous with the history of efforts to assess and evaluate students' cognitive, affective, and psychomotor behaviours. The bulk of school-based assessment research is carried out in academic settings by university faculty rather than by school practitioners. Unfortunately, the development of school-based assessment in Israel has not been systematically documented over the years and little integrative work has been devoted to historical, current, or future perspectives in school-based assessment research (Gina Ortar, personal communication, June, 1985). Consequently, this paper is intended to fill a needed gap by evaluating the present state of the art of Israeli school-based assessment research and by suggesting some potential future trends and possible directions.

The author of this paper relies primarily on three sources: (a) relevant papers identified through conventional bibliographic research as well as a

computerized search of the Israeli periodical literature for assessment-related material published over the past 15 years; (b) documents, papers, and tests provided to the author by a number of key Israeli scholars who were personally approached for relevant material; and (c) data bearing on school-based assessment in Israel, gathered from knowledgeable Israeli informants, as part of a cross-national study of testing and assessment practices among children and youth (Thomas Oakland, personal communication, December, 1988).

The paper is restricted in its scope to assessment measures (i.e., cognitive, achievement, personality) as well as assessment issues and procedures relevant to school-aged populations in Israel; preschool assessment or college-age and adult assessment themes will not be addressed. The choice and selection of material to survey was clearly subjective and by no means purports to cover the entire domain of Israeli school-based assessment research in the past or present. Due to the disparate and fragmentary sources for garnering information about school-based assessment research in Israel, a full and complete description would have been an insurmountable task. Moreover, given the author's own research interests in the areas of ability testing and test bias, coupled with the fact that some of the better assessment instruments, procedures, and research endeavours in Israel have been associated with ability and intelligence measures, there is a distinct "positive bias" in the overrepresentation of the ability-test domain in this paper.

In addition, it is clearly preferable to present a more detailed and in-depth coverage of cardinal instruments and assessment issues and high quality research and evaluation studies, as opposed to aiming for an exhaustive coverage of all relevant issues.

Cognitive Assessment

Notwithstanding the growing trend in Israeli school psychology over the past years to assist children through working with parents and their children's teacher – reflected in the decline in individual diagnosis and increase in prevention services – a good portion of the school psychologist's time is still devoted to assessment (Raviv, 1989). Individual diagnosis, including diagnosis of pupils suspected of retardation or learning and behavioural problems, is generally performed with the aid of a conventional test battery, of which the *Wechsler Intelligence Scale for Children-Revised* (WISC-R) constitutes a primary component (Raviv, 1988). In addition, a variety of group aptitude tests are used in the schools for research, evaluation, vocational guidance, and screening purposes.

Individual intelligence measures

The Israeli versions of the WISC-R. The standardization of the Israeli version of the WISC-R was begun in 1972 and was carried out on a nationwide representative sample of both Jewish and Arab populations. A total of 2700 children from 99 schools were individually tested in their native languages; the sample was found to be representative of the Israeli population with respect to sociocultural parameters. The Hebrew and Arabic versions of the WISC-R are used for diagnostic, guidance, and school classification purposes among normal, gifted, and retarded Israeli populations aged 6–16.

The Hebrew version of the WISC-R (Lieblich, Ben-Shachar, & Ninio, 1976) is currently the most widely used and perhaps best constructed diagnostic tool in Israel today for assessing Hebrew-speaking Israeli children's intelligence and cognitive difficulties. The guiding principle followed by the authors in the development of the Hebrew version of the WISC-R was to stay as close as possible to the original, unless there were strong reasons otherwise. Therefore, only a few changes were made in the general framework and structure of the tests, or the scoring system. The WISC-R consists of a Verbal (Information, $k=30$; Similarities, $k=16$; Arithmetic, $k=18$; Vocabulary, $k=24$; Comprehension, $k=17$; Digit Span, $k=14$) and Performance part (Picture Completion, $k=26$; Picture Arrangement, $k=12$; Block Design, $k=10$; Mazes, $k=7$; Digit Span, $k=93$; [Digit Span for children under eight, $k=45$]; and Object Assembly, $k=3$).

The Hebrew version of the WISC-R was normed on a representative sample of "normal" Israeli children with 100 children chosen from each of 11 age groups (6 to 16). About 8 to 16 children were tested per school (in grades 1 to 8), and children aged 12–16 were sampled from a list of elementary school graduates. The sample was about evenly distributed by sex. About 65% of the children were of Eastern extraction, 33% were of Western extraction, and about 2% third generation Israelis. Within each of the age groups, the Full Scale, Verbal, and Performance raw scores were standardized (following normalization) to a mean of 100 and SD of 15. The effective range of scores for Full Scale IQ was from 40 to 160 (about 4 SD s on both sides of the mean). The Digit Span and Object Assembly tests were not used in the calculation of scores.

As reported in the manual (Lieblich, Ben-Shachar, & Ninio, 1976), average internal consistency estimates for the WISC-R across 11 age groups

show coefficients as high as .96 for the Full Scale, .92 for the Performance, and .95 for the Verbal IQs. Subtest reliabilities are also satisfactory, ranging from .74 to .93. Stability estimates based on a small scale study among fifth grade children (for a 15-day lag) was over .90 for both the Full Scale score and the Verbal and Performance sections.

With respect to the test's construct validity, a number of studies have found the SSA space for the WISC-R to be three-dimensional and to conform to the lawfulness of a "cylinder" (Guttman & Levy, 1980; Lieblich, Ben-Shachar, & Ninio, 1976). As Guttman (1970) pointed out, if a battery of tests is constructed or selected according to three basic facets – language of presentation (verbal, numerical, geometrical); mental operation (rule inference, rule application, rule learning); and modality of expression (oral expression, manual manipulation, and use of paper and pencil) – the expected correspondence between the tri-faceted definitional system of intelligence and the empirical aspect of the correlation matrix is that of a cylinder (Levy, 1985). As reported by Levy (1985), this lawfulness has been replicated for WISC-R scores of children in Israel and the United States. Furthermore, a number of studies provide evidence for the diagnostic validity of the WISC-R profile in differentiating learning disabled from normal children aged 8 to 12 (e.g., Raviv, Margalith, Raviv, & Sade, 1981).

Group ability tests

The *Milta Group Verbal Intelligence Test* (Ortar, 1966) was for many years the most popular group verbal ability test administered in the school system. The test is modeled after the *Lorge-Thorndike Intelligence Test* and was designed as a group-administered verbal scholastic aptitude test for Hebrew-speaking students aged 9-18. It appears in parallel forms for three different grade categories (grades 4-6; 7-9; 10-12). The Milta subtest composition and number of items vary from form to form. The Milta total score is reported to be of acceptable reliability (about $r_{xx} = .90$) and has satisfactory criterion validity (r_{xy} of about .60) (Ortar, 1966).

Despite its lack of reliable Israeli norms, the *Raven Progressive Matrices* (1960) is probably the most widespread group measure of nonverbal ability used among Israeli children aged 8 to 13. The adequacy of the test's validity is generally deemed modest compared to verbal tests such as the Milta (cf. Zeidner, 1988b).

Vocational aptitude measures

In order to optimize the probability of a student's success in a particular track and to supply the school system with reliable information about

students' vocational aptitudes and interests, a good percentage of eighth and ninth grade Israeli students undergo vocational testing either at the Hadassah Guidance Institute, in Jerusalem, or at a number of other privately owned companies offering such services.

The *Hadassah Vocational Aptitude Battery* administered to eighth graders is typically composed of 12 tests assessing the following four domains: Verbal ($k=3$), Numerical ($k=3$), Figural Perception ($k=3$), and Psychomotor ($k=3$). The test battery given to ninth graders is composed of 14 tests along with measures of differential aptitudes (e.g., technical-mechanical, comprehension). The specific tests included in the battery are usually specifically tailored to the needs of the student or the school making the referral. The psychological narrative report, based on representative norms for each age group in the Jewish population, is generally relayed to the school career counsellor who interprets the results to the student and/or the student's parents. A number of large-scale prospective studies reported in the literature have shown that the aptitude test battery, when taken in grades 8-9, does a reasonably good job of predicting twelfth grade matriculation exam scores (Meier & Adler, 1985).

Perceptual tests

Based on the reports of expert educational psychologists (primary results of an international survey on testing) and the personal impression of the author, the *Bender Visual Motor Gestalt Test* (Bender, 1948) is often used for diagnostic, guidance, and placement purposes of Israeli children aged 4 to 6. The test is both group and individually administered and used for the full range of populations by schools, psychological clinics, and medical centres. Little systematic information is available regarding the test's psychometric properties or normative performance. Some psychologists also use the *Frostig Developmental Test of Visual Perception* (Frostig, Maslow, Lefevre, & Whittlesey, 1963) to help in the diagnosis of reading problems.

Aptitude tests in the works

A number of cognitive tests are currently in the process of being adapted and normed for psychoeducational diagnostic purposes in the Israeli school system. The *Kaufman Assessment Battery for Children* (K-ABC) (Kaufman & Kaufman, 1983) is in its early stages of adaptation and norming (Nevo, personal communication, June, 1989) as is the *Luria-Nebraska-Manitoba Neuropsychological Test Battery for Children* (cf. Rieck, 1988) and the *McCarthy Scales of Children's Abilities* (Kaufman & Kaufman,

1977; Miriam Rieck, personal communication, January, 1989). There are also plans for norming the Stanford-Binet. (The Hebrew translation of the Binet that does exist dates back to the 1960s and is devoid of reliable norms.)

Learning Potential Assessment Device (LPAD)

It would seem deficient to end this section without at least mentioning Feuerstein's interesting work in the development of dynamic learning potential assessment procedures. The LPAD procedure is based on a test-instruct-retest paradigm and has been used for assessing the abilities of culturally disadvantaged and learning disabled children in particular. This procedure aids at transforming the test situation from a static to dynamic one and from a product- to process-oriented one. The LPAD is offered as a viable alternative to conventional testing practices, designed to assess the educability and modifiability of an examinee's cognitive ability through active learning of thinking and test-taking strategies. Rather than focus on total scores alone, this procedure uses the peaks in the pattern of results as indicators of cognitive potential.

Based on the author's subjective impressions and knowledge of the Israeli testing context, the LPAD procedures are generally less well known than Feuerstein's internationally disseminated instrumental enrichment program (Feuerstein, 1980), and have yet to gain currency as accepted assessment procedures in Israeli psychoeducational assessment.

Scholastic Achievement Assessment

The largest national testing program ever implemented in the Israeli school system for any extended time period was the *National Achievement Test* (or *Seker*), used for guidance, resource allocation, and research purposes among eighth grade students; it was initiated in 1955 and discontinued in 1972. At present, standardized criterion- or norm-referenced achievement tests are not available, at the national level, to gauge Israeli students' attainments in various mastery domains. However, various achievement tests (e.g., Hebrew, reading comprehension, arithmetic) are administered on a partisan basis, particularly at the elementary school level. In addition, specific tests have been constructed for selection purposes by "absorbing schools," or prepared by municipal authorities for program evaluation and comparative purposes. Clearly, implementation of a large scale standardized achievement testing program at regular intervals in the schools would help improve the quality of student evaluation and allow systematic monitoring and tracking of students' academic performance.

As will be clearly evidenced in this section, a good deal of the evaluation research on student achievement has centred on sociocultural group

differences in performance. This concern surfaced in the early 1950s, when students of Eastern background were observed to fall below their Western counterparts in scholastic attainment at the very outset of their school experience, particularly in basic reading and language skills. Despite a wide array of social and educational programs designed to bridge the gap, the results have generally been disappointing and the differences have been very persistent over the years. For example, based on an index of "disadvantage" adopted by the Ministry of Education in 1974, about 44% of all Israeli students were defined as disadvantaged and, of these, 95% are of Eastern origin (Adler, 1984). It should be mentioned that a number of important internal publications of the Curriculum Division of the Ministry of Education on student achievement in various subjects (e.g., history, math, geography, Bible, writing proficiency, etc.), were unavailable to the author (at his sabbatical site at Stanford University) and will not be included in the survey.

Elementary school level

In the early 1970s a large scale study, modeled after the Coleman report in the United States (Coleman *et al.*, 1966), was launched to evaluate the scholastic achievements of Israeli elementary school children and identify student background and school variables that meaningfully contribute to success and failure in learning. The sample was composed of about 17,000 elementary school students – a representative sampling of 98 elementary schools in Israel – with particular focus on grades 1, 2, 4, and 6. Criterion-referenced achievement tests were composed in line with the demands of the specific curriculum for various subjects in each grade level. The items reflected both optimal and minimal demands of the curriculum, along with anchor items to allow comparison across grade levels. Data were collected on achievement in language, mathematics, science, geography, and Bible.

The highest mean achievement level observed, across grades, was in language; the lowest achievement level was in geography; and moderate degrees of success were reported for math, science, and Bible. Depending on the subject, low levels of student achievement were variously attributed to deficiencies in teaching methods (e.g., geography); lack of adequate facilities for experiments (science); complexity of material (e.g., math), and excessive amounts of material to be covered (e.g., Bible). A consistent rate of progress was discernible on all achievement tests, though significant differences existed between subjects, grade levels, and student groups.

Compared to the serious reading failures reported among disadvantaged students in the 1950s, the results of the language tests administered show that this particular problem has been more or less resolved. Further-

more, in contrast to the one *sigma* unit ethnic group differences found on the *Seker* qualifying exams in the 1950s and 1960s (Ortar, 1967), this report shows differences between Eastern and Western students ranging between .70 and .80 *SD*, implying a small decrease in the ethnic group disparity in achievement. Furthermore, no evidence emerged from this study to support the notion of a cumulative deficit in the achievement gap between Western and Eastern students or children of varying SES groups. Although cultural background and generation in Israel both had a significant relationship with student performance, the influence of cultural background was greater than that of the generational effect.

Comparable to the American results (Coleman *et al.*, 1966), more than one-third of the between-class variance in achievement is attributable to pupil background and less than one-tenth to schools and classroom variables. The data suggest that although the schools have contributed to the advancement of underprivileged students, the contribution is less than what the educational system could achieve and needs to achieve (Minkowitch, Davis, & Bashi, 1982).

The foregoing data are consistent with those collected as part of an international study on school achievement sponsored by the **International Association for the Evaluation of Educational Achievement** (Lewy, Rapaport, & Rimor, 1978). Meaningful differences were found between students aged 10 to 11 of Eastern and Western background – in favour of Western students – in reading comprehension, civics, and English. However, the absolute levels of achievement of Eastern students were closer to the norms of students in industrially developed than developing countries in the survey. (Meaningful ethnic group differences in student achievement in the same direction were also observed for students aged 14-15 and high school seniors). Furthermore, based on a stratified random sample of 69 elementary schools including about 3600 students, fourth to sixth grade students of Eastern extraction were reported to learn at a slower rate than their Western counterparts (Lewy & Chen, 1977), with a one- to two-year difference in achievement among the groups. Students of European origin in all three grades (4-6) consistently scored higher than their Eastern counterparts in the school subjects assessed (e.g., math, reading, English, geography), with close to a standard deviation discrepancy among the groups. However, the ethnic gap in educational achievement was not found to be cumulative and did not increase as students progressed from grades four to six.

Junior high school level

Similar to what was reported for elementary school students, the analysis of the qualifying *Seker* test results for 27,000 pupils in 1962-63 showed that the scores of European students were significantly higher than

that of Eastern students in the range of about one *SD* on all subtests (cf. Smilansky & Yam, 1969; Ortar, 1967). About 50% of Eastern students did not even reach the test's minimal requirement (set at 60) and only 15% passed the test at a score of 80. By comparison, only 10% of Western students received a grade lower than 60 while 40% managed to pass the test with a score of 80 or above.

The *Mechkar Ha-Chativot* junior high school evaluation study (Chen *et al.*, 1978) set out to evaluate the structural reform in the Israeli educational system in its transition from the traditional eight-year elementary school (8:4) to the modern six-year elementary school and three-year junior high and high school (6:3:3). The study was based on a sample of 3200 students drawn from 14 junior high schools, with 12 achievement tests in key subjects administered in each grade level. The test reliabilities ranged from .82 to .93.

On the whole no substantial differences were found in student progress for junior high school and traditional eight-year elementary schools, nor was the junior high school particularly advantageous to the scholastic performance of disadvantaged students. Levels of student attainment in a variety of subjects were disappointing, with only about 40 to 70 percent of the students reaching some reasonable level of mastery (set at 84%). Similar to the data generated by the *Seker*, about a two-year difference in achievement of students of European and Eastern origin was reported. Accordingly, students of Eastern background at the end of the ninth grade achieved at about the same level as Western students at the end of the seventh grade. Furthermore, about 50% of Western students were found to be in the highest tracks and ability groupings in key subjects compared to only about 28 – 30% of Eastern students.

With respect to gender differences in scholastic achievement, additional analysis based on the *Mechkar Ha-Chativot* junior high school evaluation data base (Kfir, 1988) showed that girls have no serious advantage over boys in objective achievement at the beginning of junior high school (grades 7-8). However, girls in each ethnic group succeeded more than boys when judged by other criteria, such as teacher evaluation in grades 7-9, higher ability group placements in junior high, and so on.

These results are partially consistent with those of another study (Dar, Resh, & Erhard, 1989) based on a national sample of about 9000 eighth and ninth graders and designed to explore gender and SES group inequalities in science and reading comprehension achievements. Adolescent boys revealed a small advantage over girls particularly in science – anywhere from .20 to .25 *SD*. The gender group differences in favour of boys was greater in grade

9 than in grade 8 and in understanding than in informational knowledge. The results bearing on ethnic group differences are consistent with previous data (though somewhat less extreme) showing an achievement gap of about .67 *SD* between students of Western and Eastern descent (Dar, Resh, & Erhard, 1989). Comparable to what was found in the Minkowitch report (Minkowitch, Davis, & Bashi, 1982), the achievements of mixed-origin students were closer to those of western origin, with the student's generation in Israel having a positive effect on achievement. SES was shown to have a stronger relationship with achievement than either ethnic origin or generation in Israel, with a gap of about one *SD* between upper and lower SES students. Multivariate analysis showed that SES mediates a considerable part of the ethnic and generational effects on achievement (Dar, Resh, & Erhard, 1989).

High school level

Matriculation exams have become the primary basis for evaluating achievement in Israeli high schools. They are required for high school credentials and university admissions in Israel and are also used for monitoring of curricula and teacher accountability as well as for research purposes (Nevo & Ben-Shakhar, 1985). Tests are given in a variety of school subjects (English, math, Hebrew, Bible, history), with each test assessed by two independent judges. The interjudge reliability of the matriculation tests varies by test content and has been reported to range anywhere from about .50 to .90 (Nevo & Ben-Shakhar, 1985). Matriculation exams are often preferred to teacher grades as achievement measures since the latter are typically less reliable, are difficult to interpret, and need to be calibrated on a common scale across schools.

Although a substantial ethnic gap still persists in matriculation exam success, it appears to be narrowing. In 1956-57 the ratio of Eastern to Western students who obtained matriculation diplomas was about one in three; by 1976-77 it was about 3 in 4. In the 10-year period from 1966-77, Eastern students who successfully passed the matriculation exam (from 17-year cohort) doubled in these years from 3.7% to 7.4%. In 1972-73, 17.4% of Eastern and 31.7% of Western student 17-year-olds obtained matriculation diplomas (as reported by Adler, 1984). On the whole, a survey of the studies available (Adler, 1984) suggests that the educational gap between sociocultural groups in Israel, despite its persistence, is in flux and heading toward a substantial decrease in the future.

A large scale evaluation study (Kfir, 1988) reports that Western students at ages 17-18 perform better than their Eastern student counterparts, both with respect to objective achievement measures, as well as number of

years of schooling and teacher evaluations. Furthermore, with respect to gender group differences, Kfir (1988) found that 18-year-old adolescent girls enjoyed an advantage over boys in the number of years of schooling, curricular track, ability group, school grades, and teacher evaluations, whereas boys had an edge on objective achievement tests.

Assessing achievement in the Arab sector

At present, there are over 1,350,000 children in the Israeli school system – about 200,000 of whom are in the Arab sector (Raviv, 1989). Israeli Arabs are a minority group in the true sense of the term, comprising over 18% of the total population and generally socially disadvantaged relative to their Jewish counterparts (Smootha & Hofman, 1976/77). Jews and Arabs differ on a wide array of sociocultural parameters, including mother tongue, religion, nationality, norms, customs, and family structure (Smootha & Hofman, 1976/77). Due to the wish of both Jews and Arabs in Israel to preserve their own cultural symbols, values, and norms, the two educational systems are kept apart. Arab schools, as a rule, have suffered from a wide array of deficiencies, including understaffed personnel, poor facilities, ill-trained teachers, and poor parental support – all interacting to contribute to the generally low levels of student achievement in the Arab sector (Bashi, Kahn, & Davis, 1981).

A comprehensive survey conducted in 1974 showed that in several areas of science and math, the average achievement of Arab elementary school students was substantially lower than that of Jewish students (Bashi, Cahan, & Davis, 1981). Interestingly, the average attainment of Arab students is remarkably similar to that of lower class first-generation Eastern Jewish students, particularly in geography and science. Although the formal curriculum is basically the same in a number of subjects (e.g., math, English, sciences), due to major differences between the Jewish and Arab school systems in language of instruction and certain curricular elements, serious limitations arise in attempting to compare student achievements in the two sectors.

Furthermore, only 35% of the Arab students who sat for matriculation exams in 1973 passed, compared with 50% of the Jewish candidates. However, in recent years, the success rate of Arab candidates in matriculation exams has been steadily going up. For example, in 1983 most Arab students who took higher level math in matriculation exams (4-5 units) passed (Saad, 1984). Comparably, a recent comparison of matriculation scores for 1778 Jewish and 1017 Arab student candidates in Israel showed an average difference of only about .25 *SD* across the subjects: math, English, and Hebrew languages (Zeidner, 1987). Also, results of the biology matriculation exams for 1983 and 1984 imply that while Jewish students obtain significantly

higher scores on the exams, the gap is narrowing, especially in tasks which require lower cognitive abilities or applications of inquiry skills (Tamir, 1986).

Educational integration and bridging the gap

Any extensive survey of the vast Israeli literature dealing with the effects of ethnic integration reforms and experimental projects on students' achievements and morale is beyond the scope of this study. (The interested reader is referred to Adler [1984].) However, it should be mentioned that, as in the United States, virtually all the Israeli summative evaluations of desegregation lead to the conclusion that at the microlevel of analysis, desegregation in itself yields little, if any, positive psychological, social, or academic results for minority pupils in the first few years following implementation (Chen, Levy, & Adler, 1978). By the same token, studies in Israel (Gutman, 1972) have shown that ability grouping does not produce a major beneficial effect for the achievement of disadvantaged children. However, Dar & Resh (1986) recently reported that classroom intellectual composition positively affects students' academic achievement, with low-resource students more sensitive than high-resource students to class compositional quality.

Personality Testing

A variety of projective measures have traditionally been part and parcel of the psychodiagnostic arsenal of Israeli school psychologists, and various standardized inventories have been employed to assess students' basic personality constructs for research and evaluation purposes. However, based on face value, many of these instruments have been indiscriminately adapted from abroad and would fail to meet systematic checks for: (a) indigenous Israeli cultural context; (b) tests for differential item-functioning; (c) tests for satisfactory validity and reliability on Israeli samples; and (d) adequacy and representativeness of available norms.

Commonly used projective measures

School psychologists in Israel administer a variety of individual projective measures for diagnostic, screening, and clinical purposes among both normal and clinical school-aged populations. Two of the more popular measures are the *Thematic Apperception Test* (TAT) (Murray *et al.*, 1938) and the *Children's Apperception Test* (CAT) (Bellak, 1975). Whereas the TAT is typically used for diagnostic and guidance purposes in Israel among both normal and disturbed children aged 8 to 16, the CAT is used primarily for younger children aged 5 to 10 for placement, guidance, and diagnostic purposes. There is little systematic data bearing on either test's psychometric

properties (e.g., reliability, generalizability, and construct or predictive validity), and both tests lack representative national norms.

The *Rorschach* (1942) is an individually administered projective measure used by school and clinical psychologists among children aged 5 to 16. The test is used for placement, guidance, and diagnostic purposes among both normal and disturbed children in the schools as well as in psychological centres and medical clinics. Here again, there is little evidence for the test's reliability or validity among national representative groups of school-aged children. In addition, various versions of the *Draw-a-Man-Test* and *Sentence Completion* projective measures are also used.

The *Bar-Ilan Picture Test for Children* (Itskowitz & Strauss, 1982) is one of the few examples of a semi-projective personality test both constructed and validated with Israeli children in mind, and was designed to elicit children's attitudes about roles in class and peer group. Information based on both test-retest and internal-consistency reliability indices show satisfactory reliability coefficients and a number of construct validity studies support the test's inferences in various areas including school anxiety, preparation for school entry, and family communications (Itskowitz & Strauss, 1982).

Standardized paper and pencil measures

The past decade has witnessed a flurry of Hebrew versions of standardized paper and pencil inventories for assessing key personality constructs adapted to the Israeli scene (e.g., *State-Trait Anxiety Inventory*, *Test Anxiety Inventory*, *16PF*). Unfortunately, a detailed survey of the various measures is beyond the scope of this study.

Sociocultural and Gender Group Differences in Cognitive Abilities

Overall, research in Israel shows that the social and educational environment throughout the school years appears to be more beneficial to the intellectual development of Western, middle-class, and male students compared to their respective Eastern, lower-class, and female student counterparts. The concern of Israeli educationists and psychologists surrounding group differences in ability and achievement is understandable in view of the acknowledged importance of cognitive performance for future academic success and social mobility in modern Israeli society, coupled with the national goal of fully integrating culturally different groups in Israel by providing them with equal educational and social opportunities. Indeed, the

documented differences in sociocultural and gender group attainment are in direct contrast to the basic values of Israeli society – first and foremost those of equality and national identity (Adler, 1984) and its common ideology which views all children as having equal potential and deserving equal treatment (Blass & Amir, 1984).

Sociocultural differences in the Jewish sector

Major sociocultural group disparities have been documented in mean levels of aptitude test performance from the very inception of intelligence testing in Israel in the late 1940s (cf. Ortar, 1953). Numerous studies over the past 40 years continued to report marked differences between Jewish examiners of Western and Eastern (Asian/African) background in intelligence test performance. The group disparities range anywhere from about one-half to one-and-half standard deviations and are documented at both the preschool (Lieblich, 1983), school (Lieblich, 1983; Minkowitch, Davis, & Bashi, 1982; Ortar, 1963; Zeidner, 1985), and university level (Zeidner, 1987).

Although there is little evidence for meaningful variations in the profiles of multiple-ability measures, by cultural group (Lieblich, 1983), Eastern examinees generally score further below their Western counterparts on verbal relative to nonverbal measures of ability (Lieblich, 1983; Minkowitch, Davis, & Bashi, 1982; Zeidner, 1985). Furthermore, middle-class children are reported to outperform their lower-class counterparts in total WISC-R performance by about .5 *sigma* units – in each ethnic group (Lieblich, 1983). A similar trend has been reported for group verbal ability tests (cf. Zeidner, 1985, 1988a). The IQ difference among SES groups starts out at about a .7 *SD* discrepancy at ages 6-7 and increases to about a 1.1 *SD* discrepancy at ages 15-16. Thus, there appears to be a cumulative increase of IQ difference on the WISC-R with age for middle- and lower-class students. By the same token, there appears to be a cumulative increase in ethnic group difference with age, rising from about .75 *SD* at ages 6-7 to a discrepancy of about two *SDs* at age 15-16. Ethnic group is reported to have a more sizable effect than SES on students' WISC-R performance, though there is a tendency for the ethnic gap to narrow among second generation Israelis.

The observed group differences have been attributed mainly to socio-cultural group differences in socialization patterns, home language, and differential schooling experiences. Furthermore, in view of the relatively more marked cultural group disparities commonly found on verbal relative to nonverbal tests, a number of Israeli scholars have argued that verbal deficit may be the main locus of the poor showing of minority groups on psychometric ability tests (Feuerstein, 1980; Minkowitch, Davis, & Bashi, 1982).

Sociocultural differences in the Arab sector

Partly because of difficulties involved in adapting conventional tools to the Arabic language and culture (Lieblich, 1983), and partly due to political considerations (Saami Mari, personal communication, May 1983), research on cognitive abilities and achievements in the Arab sector has been relatively sparse over the years. For many years psychoeducational assessment tools were simply not available for purposes of diagnosis, selection, or educational placement and intervention in the Arab sector.

In the 1970s a number of pioneering attempts were made to adapt and develop instruments for studying the cognitive development of the Arab child. In one study (Bashi, 1976; cf. Bashi, Cahan, & Davis, 1981), a subtest of *Raven Matrices* as well as the test of "associative ability" constituted part of the test battery administered to a representative group of Arab school children. Analysis of group differences at various grade levels showed that religion has a consistent effect on test performance, with Christian students performing meaningfully higher than their Moslem or Druze counterparts. Also, boys scored higher than girls on both tests and higher SES groups generally performed better than their lower SES counterparts. Religious affiliation was a better predictor of achievement than SES or sex in the Arab sector.

Overall, the few studies examining the effects of religious subgroup affiliation on intelligence test scores within the Arab population (Bashi, 1976; Lieblich, 1983; Lieblich & Kugelmass, 1981) tend to concur that Christian Arabs outscore their Moslem and Druze counterparts, whereas the latter two groups are not reliably differentiated in mean test performance. However, much like in the Jewish sector, subgroup affiliation is correlated with social background, with most of the Moslems belonging to the lower class and the Christians to the upper class. Indeed, the rank order of groups by religion conforms to what would be expected on the basis of the respective groups' socioeconomic status. In a large-scale study (Lieblich & Kugelmass, 1981) in which the WISC-R was specifically adapted to the Arab culture and standardized on a representative group of Arab children, Arab students evidenced significantly lower WISC-R total scores (about one *sigma* unit) compared to a Jewish comparison group; cultural group mean differences were greater on the performance than verbal section. In addition, several other studies conducted over the past 20 years have shown that Israeli Jewish groups outscore their Arab counterparts by about one standard deviation on standardized tests of intelligence and scholastic ability at the preschool (Kugelmass, Lieblich, & Bossik, 1974; Lieblich, 1983) and college entrance level (Zeidner, 1988b) as well.

Gender group differences

Israeli boys and girls appear to enter elementary school similarly equipped with necessary intellectual functions, as evidenced by nonsignificant gender group differences in intelligence test scores at the preschool level (Lieblich, 1983; 1985; Safir, 1986). However, boys evidence consistent superiority on intelligence tests from about the third grade, with boys beginning to outscore girls at age 9. The gender group differences grow through age 16, culminating in a 12-point difference in IQ. Similar trends are observed for both the Verbal and Performance sections. Remarkably parallel findings are reported for the Arab sector (Lieblich, 1985), with differences of about a third *SD* in favour of adolescent boys in total IQ. At age 16, the differences between gender groups are maximal – at about 10 IQ points (Lieblich, 1985).

As Safir (1986) has pointed out, gender group differences in IQ in Israel are considerably greater than in the United States from an early age on. Significant gender group differences are evident on both verbal and analytical tests among both Jewish and Arab college candidates as well (Zeidner, 1986a, 1986b). In fact, sex is the third demographic factor in magnitude of effect on intelligence in both the Jewish (ethnicity > SES > sex) and Arab (religion > SES > sex) population. These data are in contrast to much previous research in the West showing that adolescent boys are usually better in spatial and numerical abilities, while girls are superior in verbal abilities. These reported gender differences in ability are generally attributed to gender differences in life experiences, including: unequal intellectual opportunities; differential expectations of parents and teachers for boys and girls; differential peer group pressures; and exposures to different role models (Lieblich, 1985; Safir, 1986).

Test Bias and Equity

The widespread reliance on measures of scholastic ability for assessment and placement of students from diverse cultural backgrounds in the Israeli school system has raised serious concern regarding cultural test bias and fairness (Zeidner, 1987; 1988a; 1988b). The recent Israeli “antitest movement” and public outcry is noteworthy, since the chronology of litigious events closely parallels those leading up to the enactment of PL 94-142 in the United States. Indeed, elements in the anti- and pro-testing debate are nearly identical to those heard in the American educational system over the last 20 years. Among the striking cross-cultural parallels in the Israeli and American antitest movement are:

(a) *Antecedents of public concern with standardized testing.* These are: sizable sociocultural differences in aptitude and achievement test scores; popularity of the cultural difference position in the mid-1970s (Stahl, 1977); contentions of test bias, and adverse impact for cultural minority groups in the population; severe overrepresentation of lower-class Eastern children in special education programs and low-status tracks; and the sequence of Israeli attempts to deal with ethnic and SES divergence in abilities and school performance at a structural level (Miller, 1984).

(b) *Manifestations of antitest sentiments.* These are: vehement attack on scholastic aptitude tests in the media; appeal to the courts by concerned parents to ban the usage of aptitude tests in elementary schools for student selection and classification purposes (Cahan, 1986); an Israeli version of the "antitest syndrome" (cf. Jensen, 1980) surrounding the testing melodrama, characterized by an intense emotional involvement in the test debate on the part of both test protagonists and antagonists, contamination of reason by affect and ideology, and armchair speculation concerning the degree of bias and equity in testing (cf. *Israeli Psychological Association Bulletin*, 1986).

(c) *Consequences of the antitest campaign.* These include: a professional committee mandated to examine the goals, usages, problems, costs, and ethical implications of group ability test usage in the schools (Ministry of Education, 1985); ban of widespread and massive administration of group ability tests for selection and placement purposes in elementary schools by the Ministry of Education as of January 1985.

(d) *Results of test-bias research.* Research in the Israeli scene setting out to empirically test for cultural fairness in the predictive validity of both verbal and nonverbal tests for majority and minority group (Zeidner, 1986b) lend generalizability to much previous research conducted in the United States, providing some evidence of intercept bias (scores were overpredictive of minority group students) and little evidence in predictive slope bias by sociocultural group. Furthermore, a series of tests of the "situational bias" hypothesis conducted by this author among school children revealed that sociocultural group membership does not interact with test atmosphere (Zeidner, 1985) or with examinees' test attitudes and dispositions (Zeidner, 1988a) in affecting ability test performance. In view of the Israeli evidence, the cultural bias hypothesis is not vindicated.

Rather than acknowledge the persistence of sociocultural group differences in test performance and focus additional efforts on raising the level of minority group performance, many simply preferred to "kill the messenger" by attacking the validity and equity of the tests themselves (cf. Lerner, 1989).

Future Trends and Directions

Future developments

This review has shown that notwithstanding remarkable advances and developments in the field of school-based assessment in Israel over the years, there is much left to be desired. The current state clearly reveals the need for a major and concerted effort on the part of Israeli psychologists and measurement experts in improving both the quantity and, even more important, the quality of school-based measurement instruments and procedures.

Additional school-based assessment instruments are needed for diagnostic and evaluation purposes for a host of special student populations characterized by learning disabilities, attention disorders, mental retardation, emotional disturbance, physical disorders, and intellectual giftedness. Among the school-based measures that appear to be most desperately needed are individual and group tests of intellectual, perceptual, and emotional development, not to speak of standardized norm- and criterion-referenced measures of achievement.

The Israeli school system needs to direct systematic effort towards the development of criterion-referenced measures of achievement for the various school subjects taught in the Israeli school system — all the way up from elementary to high school. These measures would preferably include items indicative of both minimal and optimal levels of mastery (for formative and summative evaluation purposes) as well as anchor items for future tracking of student achievement across time (cf. Minkowitch, Davis, & Bashi, 1982). In particular, there is a crying need for developing standardized achievement tests in the various school subjects, at all grade levels, for use in the Arab sector.

Among the foreseeable developments and trends in assessment practices as a consequence of more global and general developments in school-based assessment are: decreased emphasis on norm-referenced measurements and increased emphasis on criterion-referenced achievement tests for student assessment and placement decisions; multifaceted assessment and use of a greater variety of assessment tools — supplementing standardized tests with other measures (interviews, direct observations); increased use of tests for program evaluation and decision-making purposes and explicit consideration of the costs of various types of errors in the decision-making process based on tests (e.g., banning ability tests for selection purposes may decrease the number of false negatives, but increase the number of false positives); and increased use of behavioural assessment technologies to

assess learning problems and to monitor progress in educational programs and school settings.

In addition, more emphasis will probably be placed on establishing and implementing elements of sound and ethical testing practices – a relatively neglected issue in Israeli testing-practice in the past. In specific, considerable attention will most likely be given to the development of ethical testing standards; specification of technical guidelines for test development; guidelines for test administration and usage; credentials and licensure; implementing informed consent procedures; securing test data and storing test material; and confidentiality of test or assessment information. Although very little *a priori* thought has been given to systematically establishing testing standards and ethical codes for test construction, usage, interpretation, and decision making, this situation is fortunately changing. The “Nevo Committee” (Nevo, 1988), mandated by the Ministry of Education, has recently suggested some very important, timely, and practical suggestions towards improving testing and selection practices in the schools.

Furthermore, school-based assessment in Israel would most likely profit from the application of recent advances in testing technologies, including item-banking of ability and achievement test items; computer adaptive testing programs; computerized test scoring and interpretation (with some headway already being made in the area of mathematics testing); and the use of modern test theory in the construction and scaling of items (e.g., logistic models and item-response theory and techniques).

Needed research

Future developments in testing and assessment in Israel will most likely be accompanied by systematic research in a wide variety of areas. Some of the potential areas for much needed school-based assessment research are (not necessarily in order of importance): (a) the feasibility of item-banking and computerized testing, scoring, and interpretation for school-based aptitude, achievement, vocational, and personality testing purposes; (b) examination of student trajectories and growth curves in achievement, abilities, and school-related personality variables over time; (c) psychometric properties of school-based instruments administered conventionally and by electronic means; (d) test bias for special subgroups in the population (e.g., racial, gender, handicapped), including tests for invariance of tests' structural components, differential item functioning, and differential predictive validity; (e) examinee test attitudes and dispositions; (f) effects of coaching on standardized aptitude tests; (g) contextual and personal variables in the assessment situation (test anxiety, test motivation, self-efficacy);

(h) test-taking strategies and skills of various subgroups in the population; (i) cognitive skill analysis of test performance; (j) specific strengths and weaknesses in test-taking and performance in various subgroups in the population; and (k) specific loci of low test performance of ethnic minority groups.

Concluding remarks

After an initial spate of enthusiasm and widespread adoption of a variety of tests in Israel during the 1960s and 1970s, the antithetical reaction to testing appeared in full blown force in the 1980s. According to the Hegelian triad, we are now bound to reach some sort of resolution and arrive at a consensus with respect to valid test usage or alternative assessment practices in the schools in order to secure a satisfactory synthesis between the opposing test and antitest forces at play.

In the first few decades of its existence, the Israeli educational system has almost miraculously managed to come to grips with pressing problems of both a quantitative nature (i.e., accommodating the waves of immigrant children and helping them assimilate to Israeli society) and of a qualitative nature (raising minimum levels of achievement of disadvantaged groups). The future demands a concerted effort at raising the level of functioning and standards of the educational system as a whole. A major effort directed at upgrading and improving school-based assessment procedures may play a key role in helping to achieve this important national goal.

Assessment tools and procedures used for educational decision-making purposes have important pedagogical, social, and economic implications for social systems (cf. Wainer & Braun, 1988). However, given the uniquely vulnerable and particularly difficult economic, social, political, and military situation of modern-day Israel, the importance of valid educational assessment, facilitating the optimal utilization of Israeli society's human resources, indirectly affects the welfare of the nation. Thus, the importance of valid assessment tools goes beyond their importance in sheer maximization of utility for the decision-maker; it may have critical consequences for the future social, economic, and military development and very survival of Israel in the years to come.

NOTE

This paper was completed while the author was on sabbatical at the Program for Psychological Studies, School of Education, Stanford University. The help of the host university in facilitating the completion of this manuscript is gratefully acknowledged. I would like to thank Thomas Oakland of the University of Texas at Austin for making available to me relevant Israeli research data gathered as part of a cross-national study of assessment practices and procedures. In addition, I would like to express my thanks to the following individuals for providing me with relevant documents and materials in the course of preparing this manuscript at my sabbatical site abroad: Sarel Cahan and Uriel Last of the Hebrew University of Jerusalem; Baruch Nevo and Miriam Rieck of the University of Haifa; Rivkah Itskowitz of Bar Ilan University; and Elchanan Meier and Amiram Raviv of Tel-Aviv University. Thanks are also due to Dan Inbar of the Hebrew University of Jerusalem, Aryeh Lewy of Tel-Aviv University, Ozer Schild of Haifa University, and Yoav Cohen of the National Institute of Testing and Evaluation in Jerusalem for their useful comments and suggestions on an earlier draft of this manuscript. The responsibility for the content of this paper, however, lies with the author alone.

Correspondence regarding this manuscript may be addressed to: Moshe Zeidner, Ph.D., School of Education, University of Haifa, Mt. Carmel, 31999, Israel.

This paper was presented at the 97th Annual Convention of the American Psychological Association at New Orleans, Louisiana, August, 1989.

REFERENCES

- Adler, Ch. (1984). School integration in the context of the development of Israel's educational system. In Y. Amir & S. Sharan (Eds.), *Social desegregation: Cross-cultural perspectives* (pp. 21-46). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Bashi, Y. (1976). *Verbal and nonverbal abilities of students in grades four, six and eight in the Arab sector*. Jerusalem: School of Education, Hebrew University.
- Bashi, Y., Cahan, S., & Davis, D. (1981). *Achievement of Arab pupils in the elementary schools in Israel*. Jerusalem: School of Education, Hebrew University.
- Bellak, L. (1975). *The TAT, CAT, and SAT in clinical use* (3rd ed.). New York: Grune & Stratton.
- Bender, L. (1948). *Manual for instruction and test cards for Visual and Motor Gestalt Test*. New York: American Orthopsychiatric Association.
- Blass, N., & Amir, B. (1984). Integration in education: The development of a policy. In Y. Amir & S. Sharan (Eds.), *Social desegregation: Cross-cultural perspectives* (pp. 63-98). Hillsdale, NJ: Lawrence Erlbaum Associates.

- Cahan, S. (1986). *The role of ability tests in educational decision-making: The case of tracking and assignment to homogeneous homeroom classes in the Israeli junior high school*. Paper presented at the 21st International Congress of Applied Psychology. Jerusalem, Israel, July 1986.
- Chen, M., Lewy, A., Adler, C., Inbar, D., Kfir, D., & Resh, N. (1978). *Process and result in education: Evaluating the contribution of junior high school to the educational system*. Tel-Aviv & Jerusalem: Tel-Aviv University and Hebrew University (in Hebrew).
- Coleman, J.S., et. al. (1966). *Equality of educational opportunities*. Department of Health, Education and Welfare, Office of Education, Government Printing Office, Washington, DC.
- Dar, Y., & Resh, N. (1986). Classroom intellectual composition and academic achievement. *American Educational Research Journal*, 23, 357-374.
- Dar, N., Resh, N., & Erhard, R. (1989). *Learning achievements in junior high schools in reading comprehension and science*. Jerusalem: School of Education (in Hebrew).
- Frostig, M., Maslow, P., Lefevre, D.W., & Whittlesey, J.R.B. (1984). The Marianne Frostig Development Test of Visual Perception. 1963 Standardization. *Perceptual and Motor Skills*, 19, 463-499.
- Feuerstein, R. (1979). *The dynamic assessment of retarded performers*. Baltimore: University Park.
- Feuerstein, R. (1980). *Instrumental enrichment*. Baltimore: University Park.
- Gutman, I. (1972). *An experimental investigation of the effects of various grouping methods on the cognitive and psychosocial development of elementary school pupils*. Jerusalem: Szold (in Hebrew).
- Guttman, L. (1970). Integration of test design and analysis. In *Proceedings of the 1969 Invitational Conference on Testing Problems*. Princeton, NJ: Educational Testing Service.
- Guttman, L., & Levy, S. (1980). Two structural laws for intelligence. *Megamot*, 25, 421-438 (in Hebrew).
- Hadassah Vocational Guidance Institute. (1982). *Bulletin of the Guidance Department* (in Hebrew).
- Itskowitz, R., & Strauss, H. (1982). *The Bar-Ilan Picture Test for Children* (Revised Ed.). Copenhagen, Denmark: Dansk psykologisk Forlag.
- Israeli Psychological Association Bulletin*. (1986). February issue (in Hebrew).
- Jensen, A. (1980). *Bias in mental testing*. New York: Free Press.
- Kaufman, A.S., & Kaufman, N.L. (1977). *K-ABC: Kaufman Assessment Battery for Children*. Circle Pines, MN: American Guidance Service.
- Kfir, P. (1988). Achievements and aspirations among boys and girls in high school: A comparison of two Israeli ethnic groups. *American Educational Research Journal*, 25, 213-236.
- Kugelmass, S., Lieblich, A., & Bossik, D. (1974). Patterns of intellectual ability in Jewish and Arab children in Israel. *Journal of Cross-Cultural Psychology*, 5, 184-198.
- Lerner, B. (1989). Intelligence and law. In R.I. Linn (Ed.), *Intelligence: Measurement, theory, and public policy*. (pp. 172-192). University of Illinois Press.

- Levy, S. (1985). Lawful roles of facets in social theories. In D. Canter (Ed.), *Facet theory: Approaches to social research* (pp. 59-96). New-York: Springer-Verlag.
- Lewy, A., & Chen, M. (1977). Differences in achievement: A comparison over time of ethnic group achievement in the Israeli elementary school. *Evaluation in Education, 1*, 3-72.
- Lewy, A., Rapaport, Ch., & Rimor, M. (1978). *Educational achievement in the Israeli school system*. Tel Aviv: Ramot (in Hebrew).
- Lieblich, A. (1983). Intelligence patterns among ethnic and minority groups in Israel. In M. Nisan & U. Last (Eds.), *Between education and psychology* (pp. 335-357). Jerusalem: Magnes Press (in Hebrew).
- Lieblich, A. (1985). Sex differences in intelligence test performance of Jewish and Arab school children in Israel. In M. Safir, M.T. Mednick, D. Israeli, and J. Bernard (Eds.), *Women's worlds: From the new scholarship*. New York: Praeger.
- Lieblich, A., & Kugelmass, S. (1981). Patterns of intellectual ability of Arab school children in Israel. *Intelligence, 5*, 311-320.
- Lieblich, A., Ben-Shachar, S., & Ninio, A., (Eds.). (1976). *Manual WISC-R: Wechsler Intelligence Scale for Children*. Revised. Institute for Psychological Development of the Hebrew University and Psychological Advisory Service of Ministry of Education and Culture, Jerusalem.
- Meier, E., & Adler, N. (1985). Prediction of success in matriculation exams. *Studies in Education, 41*, 35-50 (in Hebrew).
- Miller, N. (1984). Israel and the United States comparisons and commonalities in school desegregation. In Y. Amir & S. Sharan (Eds.), *Social desegregation: Cross-cultural perspectives* (pp. 237-253). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Ministry of Education. (1985). *Report of the Rashi Committee*. Jerusalem: Government Printing Office (in Hebrew).
- Minkowitch, A., Davis, D., & Bashi, Y. (1982). *Success and failure in Israeli elementary education*. New-Brunswick, NJ: Translation Books.
- Murray, H.C. et al. (1938). *Explorations in personality*. Cambridge, MA: Harvard University Press.
- Nevo, B. (1987). *Report of the Committee for Formulating Recommendations for Classifying Students in High School*. Jerusalem (in Hebrew).
- Nevo, B. (1988). Tests in Israel and the US: The state of the art. In B. Nevo & Y. Cohen (Eds.), *Selected problems in evaluation and measurement*. Jerusalem: National Institute of Measurement and Evaluation (in Hebrew).
- Nevo, B., & Ben-Shakhar, G. (1985). Interjudge reliability of matriculation exams. *Megamot, 29*, 42-56 (in Hebrew).
- Ortar, G. (1953). The diagnostics of the Wechsler exam among individuals with different intelligence levels. *Megamot, 3*, 199-216 (in Hebrew).
- Ortar, G. (1963). *Is a verbal test cross-cultural? Scripta Hierosolymitana* (pp. 219-235). Jerusalem: Hebrew University.
- Ortar, G.R. (1966). *Milta Intelligence Test*. Jerusalem: Hebrew University

- Ortar, G. (1967). Thirteen years of "Seker": Student achievements. *Megamot*, 15, 220-230.
- Raven, J.C. (1960). *Guide to using the Standard Progressive Matrices*. London: Lewis.
- Raviv, A. (1988). *School psychology research in Israel*. Unpublished Draft Manuscript.
- Raviv, A. (1989). School psychology in Israel. In P. A. Saigh & T. Oakland (Eds.), *International perspectives on psychology in the schools* (pp. 11-124). Hillsdale, NJ: LEA.
- Raviv, A., Margalith, M., Raviv, A., & Sade, F. (1981). The cognitive patterns of Israeli learning disabled children as reflected in the Hebrew version of the WISC-R. *Journal of Learning Disabilities*, 14, 411-415.
- Rieck, M. (1988). Evaluation of children with learning disabilities on Luria's Neuropsychological Test Battery. *School Psychology International*, 9, 285-289.
- Rorschach, H. (1942). *Psychodiagnostics: A diagnostic test based on perception*. Berne: Huber (1st German Edition).
- Saad, S. (1984). The Arab educational system: Trends and problems. In A. Ramsberg & A. Moskovitz (Eds.), *The Arab citizens of Israel*. Jerusalem: Van Leer (in Hebrew).
- Safir, M.P. (1986). The effects of nature or of nurture on sex differences in intellectual functioning. *Sex Roles*, 14, 581-590.
- Smilansky, M., & Yam, Y. (1969). The relationship between family size, ethnic origin, fathers' education, and students' achievement. *Megamot*, 16, 248-273 (in Hebrew).
- Smootha, S., & Hofman, J.E. (1976/77). Some problems of Arab-Jewish coexistence in Israel. *Middle East Review*, 9, 5-15.
- Stahl, A. (1977). *Language and thought of culturally deprived children in Israel*. Otzar Ha-Morch: Tel-Aviv (in Hebrew).
- Tamir, P. (1986). Achievement of Jewish and Arab students who studied inquiry oriented curriculum for several years. *Studies in Educational Evaluation*, 12, 191-195.
- Wainer, H., & Braun, H.L. (1988). *Test validity*. Hillsdale, NJ: Lawrence Erlbaum.
- Zeidner, M. (1985). A cross-cultural test of the situational bias hypothesis – the Israeli scene. *Evaluation and Program Planning*, 8, 367-376.
- Zeidner, M. (1986a). Sex differences in scholastic aptitude: The Israeli scene. *Personality and Individual Differences*, 7, 847-852.
- Zeidner, M. (1986b). Sex differences in scholastic ability in Jewish and Arab college students in Israel. *Journal of Social Psychology*, 7, 847-852.
- Zeidner, M. (1987). A test of the cultural bias hypothesis: Some Israeli findings. *Journal of Applied Psychology*, 72, 38-48.
- Zeidner, M. (1988a). Sociocultural differences in examinees' attitudes towards scholastic ability exams. *Journal of Educational Measurement*, 25, 67-76.
- Zeidner, M. (1988b). Cultural fairness in aptitude testing revisited: A cross-cultural parallel. *Professional Psychology: Research and Practice*, 19, 257-262.

Moshe Zeidner is a faculty member at the School of Education, University of Haifa. Research areas in which he has published include personality and individual differences, with particular concern for anxiety; cultural, sex, and age bias in psychological tests; sex, sociocultural, and age differences in ability; test attitudes and motivations. He serves as co-editor and methodological consultant for the new international journal *Special Education and Rehabilitation: A Multidisciplinary Journal* and methodological consultant and member of the review board for the American publication *Journal of Classroom Interaction*. He is currently a member of the Executive Committee of the International Society for Test Anxiety Research (STAR), chairs the Research Committee of the International School Psychology Association, and is a member of the American Psychology Association.

Moshe Zeidner enseigne à la faculté des sciences de l'éducation de l'université de Haïfa. Ses recherches portent sur les différences de caractère et les différences individuelles, notamment au niveau de l'anxiété, sur la distorsion des tests psychologiques au niveau de la culture, du sexe et de l'âge, sur les différences de compétences liées au sexe, aux éléments socioculturels et à l'âge et enfin sur les attitudes et les motivations des tests. Il a été coéditeur et consultant de la nouvelle revue internationale *Special Education and Rehabilitation: A Multidisciplinary Journal* et consultant et membre du comité de lecture de la revue américaine *Journal of Classroom Interaction*. Il est actuellement membre du comité exécutif de l'International Society for Test Anxiety Research (STAR), il préside le comité de recherche de l'Association internationale des psychologues scolaires et il est membre de l'American Psychology Association.

