



The Joseph and Alma Gildenhorn
Institute for Israel Studies
University of Maryland

Learning Assessments in Israel's Schools

Beyond Controversy and Towards Best
Practice

Laurence Wolff
January 9, 2014
Washington, DC

Research Paper 11
January 2014





The Gilddenhorn Institute for Israel Studies (GIIS) is proud to publish Laurence Wolff's analysis of learning assessments in Israeli schools. We will be making it available to scholars of education, academics, teachers, parents, and all those concerned with education in Israel today, and we hope that both its critique and its recommendations will contribute to the current debate over education in Israel as well as to solving some of its serious problems.

The GIIS is one of the largest and most active centers for the study of Israel in the United States today. Headed by Professor Yoram Peri, the Abraham S. and Jack Kay Chair in Israel Studies, the GIIS offers 15 courses each year to 500 students, which focuses on five different aspects of Israel: Israeli History, Society, Politics, Culture, and the Middle East Conflict. The GIIS also sponsors conferences, lectures, cultural events, and community forums.

One of the important aspects of the GIIS is sponsoring, conducting, and publishing research on various aspects of Israel Studies. Our signature research project is **Israel 2023**, a series of three monographs that construct different possible scenarios for Israel as it approaches its 75th anniversary in three broad areas of major concern: Relations between Arab and Jewish Israelis, Israel's Geo-strategic position, and Religion and State in Israel. The Institute also regularly publishes on its website both original research and translations of important articles that have already appeared in Hebrew.

More information about the GIIS is available at:

THE GILDENHORN INSTITUTE FOR ISRAEL STUDIES

4137 Susquehanna Hall, University of Maryland, College Park, MD 20742 U.S.A.

www.israelstudies.umd.edu

ABOUT THE AUTHOR

Laurence Wolff was a senior education officer at the World Bank for 22 years. Since 1999 he has consulted for USAID, UNESCO, the Inter-American Development Bank and others. He is a frequent visitor to Israel, where his family is involved helping the Daniel Centers for Progressive Judaism in Tel Aviv. He may be contacted at larrywolff@live.com.

TABLE OF CONTENTS

LEARNING ASSESSMENTS IN ISRAEL: INCREASED USE AND GROWING CONTROVERSY.....	1
ISRAEL'S NATIONAL ASSESSMENT (THE MEITSAV)	3
Experience to Date.....	3
Towards Best Practice.....	5
ISRAEL'S PARTICIPATION IN INTERNATIONAL TESTS (TIMSS AND PISA)	6
Experience to Date.....	6
Towards Best Practice.....	9
ISRAEL'S END OF SECONDARY CYCLE EXAMINATION (THE BAGRUT)	10
The Current Situation.....	10
Towards Best Practice.....	11
CONCLUSION	12
BIBLIOGRAPHY AND REFERENCES.....	13

Learning Assessments in Israel: Increased Use and Growing Controversy

Assessing learning achievement has become a fundamental element of the education policy and decision making process of countries, states and provinces, cities, and schools. Most countries now have national learning assessments, given on a sample and/or universal basis. Over 60 countries now participate in international learning assessments where the progress of their students is charted. In addition most countries around the world have national tests at the end of the secondary cycle which, while not normally used for policy decisions, play an important role in determining a student's progression into higher education. While there is an increasing worldwide commitment to measuring achievement, especially in reading and mathematics, both nationally and internationally, there continue to be debates in many countries on the value of standardized large-scale tests, with professionals, decision makers, teachers, parents, and the public weighing in on these issues.

Israel's education system has moved in the direction of other countries in the world. A nationwide achievement test, the Meitsav, (the Growth and Effectiveness Measurement of Schools) has been given to Israel's students since 2002. Israel regularly participates in the international tests given by PISA and the IEA (TIMSS and PIRLS). For a long time Israel has had a high stakes end of secondary cycle test, the Bagrut. At the same time Israel has seen a vigorous and often noisy debate on the value of standardized tests.

Reports and analyses over the last ten years have argued that the quality of Israel's elementary and secondary schools is mediocre,¹ compared to the expectations of its own population and its performance compared with other OECD countries. Israel has regularly scored lower than most OECD countries in international tests. In a striking change, and as widely reported in the Israeli press in late 2012, Israel's standing in the 2011 Trends in the International Mathematics and Science Study (TIMSS), improved from 26th to 6th between 2007 and 2011. In late 2012, when the results were in, the minister at the time (Gideon Saar) said that there had been a "revolution in basic education in Israel" (IMRA). But this reported improvement was met by skepticism from some sectors of the general public as well as in the press. Some argued that there was cheating or manipulation and others expressed dismay at how overwhelmingly important test scores had become.

After the elections of February 2013, the Yesh Atid party sought and gained the education portfolio. The new Minister of Education, Rabbi Shai Piron, a trained educator and former head of Hakol Chinuch, an organization which has lobbied for a decade for more education funding and attention to equity, has recently argued that "testing is not learning." He has proposed that "values" should be a more important educational objective than has been the case up to now.

In August 2013 Minister Piron announced that he was suspending implementation of the year's external Meitsav test. A number of public commentators agreed that "testing" had become a "golden calf" at the expense of quality teaching and values (Skop). In a response reported in the Jerusalem Post, previous Minister Saar (now Minister of the Interior), said that, "When we decide to cancel nationwide standardized tests that serve as the basis for the education system's data and are used to improve schools, we are replacing a culture of measurement with one of pretending." Minister Piron responded that he was

¹ See for example the author's report "Education in Israel: the Challenges Ahead."

not canceling the Meitsav but “freezing” it. “I am not against measurements,” he said. “What I’m against is the superficial culture of ‘A Star is born’” (Harkov 1).

The new Minister has also been particularly concerned with what he believes is the perverse impact of the “Bagrut” examination, Israel’s end of secondary cycle examination, given to most students who complete secondary education, and an important determinant of which higher education institutions they will attend. The Bagrut is a high stakes end of cycle examination, similar to the BAC in France, the Abitur in Germany, and the General Certificate of Education in the UK. Every student must take tests in seven subjects. The Minister has proposed to reduce the number of required subjects externally tested from seven to four—English, mathematics, Hebrew or Arabic, and one optional subject (Erez). Some stakeholders, including the head of the teachers union (Erez) as well as Prime Minister Netanyahu (Harkov2), have expressed concern that, if only four subjects are tested, then students will pay little or no attention to other subjects.

This report seeks to put these controversies in context both within Israel and internationally. It summarizes Israel’s experience with its national assessments, the Meitsav, reviews what actually happened to account for Israel’s recent improvement in international tests, and discusses issues related to the Bagrut. It comments on recent policy changes and discusses how to incorporate best practice worldwide in educational assessments into Israel’s education system.

Out the outset it should be understood that, while assessing learning achievement is increasingly seen as an essential component of education systems, their widespread adoption has led to some negative consequences. In 1976 a social scientist (Donald Campbell) recognized the drawbacks, and coined “Campbell’s Law,” which states that “The more any quantitative social indicator (or even some qualitative indicator) is used for social decision-making, the more subject it will be to corruption pressures and the more apt it will be to distort and corrupt the social processes it is intended to monitor Achievement tests may well be valuable indicators of general school achievement under conditions of normal teaching aimed at general competence. But when test scores become the goal of the teaching process, they both lose their value as indicators of educational status and distort the educational process in undesirable ways” (Campbell).

In education these distortions can include narrowing the curriculum to the tested subjects, using most of classroom time on items which are surmised to be on the test, increased time allocated to tested subjects, reduction in teachers’ flexibility and creativity, and subtle and not so subtle attempts at cheating or “gaming” the system. Poorly constructed assessments often measure recall and “lower order” learning skills rather than the ability to analyze or synthesize knowledge (see for example, Koretz, on these problems in the US). Values, citizenship, creativity, artistic and musical expression, etc. can become devalued.

Around the world these issues are not uncommon. The public is often obsessed with “league tables,” which are inevitably highlighted in newspaper reports and in public perceptions, rather than with how testing results can be used to reform pedagogy, management, learning goals, opportunity to learn, and teacher training. In addition, the objectives of education are far greater than those of numeracy and literacy. A recent international committee defined six clusters of education objectives—literacy and communication, numeracy and mathematics, science and technology, physical well-being, social and emotional growth, approaches to learning and cognition, and abilities in and knowledge of culture and the arts (UNESCO). It is essential for schools to pay attention to all of these clusters, as well as values, citizenship, creativity, capacity to cooperate and to work together, etc.

“Campbell’s law,” is a “tendency” rather a law. But it underscores the need for alertness, care, and constant tinkering to make sure that unintended consequences are kept to a minimum.

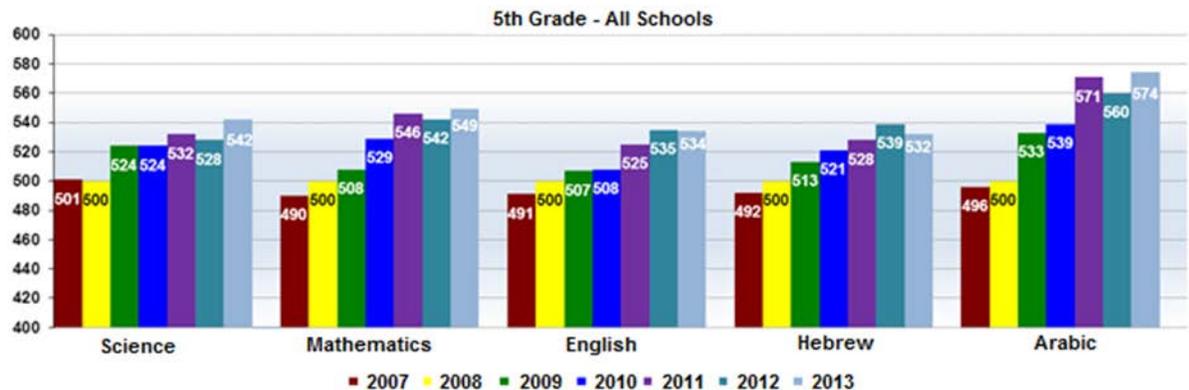
Israel’s National Assessment (the Meitsav)

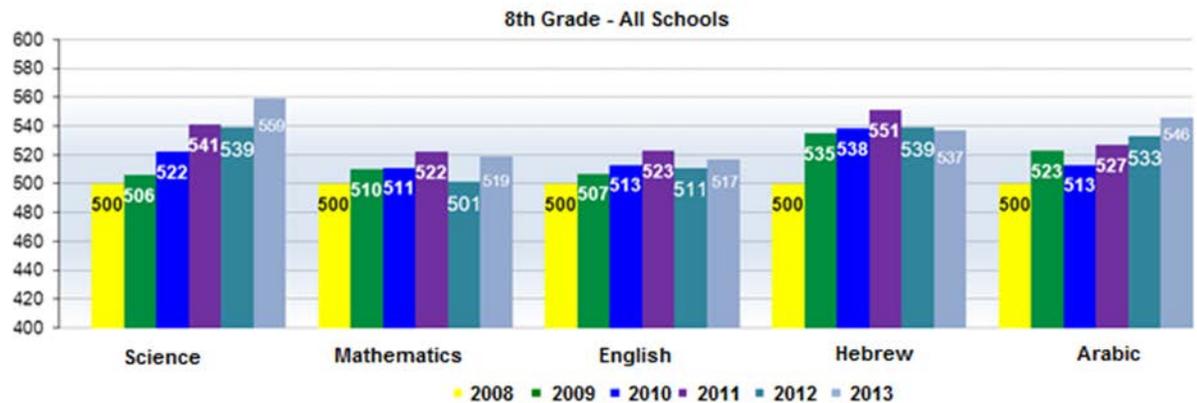
Experience to Date

In 2002 Israel began to assess learning through the “Meitsav.” In 2006 it established a semi-autonomous national center for educational assessments, the “National Authority for Measurement and Evaluation in Education (RAMA in Hebrew), with the declared purpose to “implement a culture of ‘assessment in the service of learning’” (Beller1).

Beginning in 2007, technical refinements introduced by RAMA have made it possible to compare Meitsav scores from one year to the next. As can be seen in chart 1, Israel has seen relatively consistent improvements in scores on these tests since 2007, especially at the fifth grade level but also at the eighth grade level. All of the scores were set at a standardized mean of 500 in 2007. In 2012 the average of all scores in science, math, English, Hebrew and Arabic was 541 for fifth grade and 525 for 8th grade. There were modest declines in results in most of the assessments from 2011 to 2012 and modest increases in 2013. Over time the gaps between Jewish and Arab students have declined, although they remain significant, with the lowest scorers Bedouins in the South. Much of the gap between Jews and Arabs disappears when socio-economic status of students’ parents is taken into account.

Chart 1 MEITSAV SCORES 2007-2013





Source: RAMA

Under the system developed by RAMA, external tests are given in science, math, English, Hebrew (for Hebrew speakers) and Arabic (for Arabic speakers) in grades 5 and 8. These tests are given on a staggered basis. Israel's schools are divided into four "groups" each of which is representative of the country as a whole. Individual schools are tested once every two years in external tests in two subjects. Two years later the schools are tested in the two other subjects. In the years when a school is not externally tested, sample tests are provided to principals and teachers for their own use within the classroom. The idea behind internal exams is to provide teachers and principals with a snapshot of the conditions of learning in their classroom without the negative consequences of a higher stakes external test.

In its documentation, staff of RAMA have regularly expressed their awareness of the potentially unintended negative consequences of testing. They designed the Meitsav as minimally intrusive test which would serve as a pedagogical tool for teachers, emphasizing that assessment should be "in the service of learning."

RAMA sends a detailed report to each participating school. It provides guidelines to teachers and principals for utilization of results, including improving teacher-developed classroom evaluations. Each school selects one of its staff as an assessment coordinator, who helps teachers to utilize test results. Concerned with possible neglect of untested subjects, RAMA has begun efforts to test other subjects. RAMA also began an applied research program which uses assessment results to examine policy issues such as equity, financing, and the impact of pre-schooling. RAMA also surveys student and teacher perceptions of "school climate," violence in schools, and relations between students and teachers. These have shown improvements since the surveys were initiated in 2007.

Until 2012 scores by region, ethnicity, and types of schools were provided to the public, with scores provided to individual schools on a confidential basis. In 2012 a lawsuit sought to require the Ministry to publish school by school results. A Ministry of Education committee concluded that publication by school might have more negative consequences than positive ones. In September 2012 the Israeli

Supreme Court ruled that the Ministry had to make all results public, which was done shortly thereafter. As a result "league tables" popped up immediately in the papers.

As noted above, In August 2013 Minister Piron announced that he was suspending implementation of the 2013 external Meitsav test (the internal sample tests continued to be made available to schools). One reason for the suspension, he said, was that continued publication of school by school results might lead to excessive preoccupation with the "league tables" by stakeholders as well as the general public.

While it appears that many teachers and administrators have welcomed the positive elements of the current testing regime and use the tests to improve their pedagogy, newspaper reports regularly highlight distaste of learning assessments and skepticism about the reliability of test results. This could be in part a result of lack of understanding by reporters or their preference to highlight negative news. But it also suggests that, in the past, the positive pedagogical purposes of the Meitsav have not been adequately marketed to the public.

Towards Best Practice

In best practice worldwide, assessment is a tool to improve learning rather to punish "poor" performers, with feedback designed to strengthen the teacher's ability to improve his/her pedagogy and to undertake classroom assessments. It is but one necessary element of a wider set of policies and actions to increase learning through changes in the implemented curriculum, teacher training, and standards setting and implementation. Best practice assessment includes consensus, public understanding, transparency, commitment, continuity, a sense of shared responsibility, adequate financing of assessment efforts, high quality staff, and a strong emphasis on dissemination and use. Best practice militates against cheating and gaming the system through ensuring that test items reflect real curriculum goals and through rigorous controls and oversight. It keeps the number of external tests to the minimum needed to measure progress and inform decision making at the school, local, regional, and national levels.

Israel's Meitsav fulfills most of the elements of best practice. Compared with a number of other countries, Israel is not "over tested," since only two grades are tested externally and not every year. This compares, for example, with the USA where the "No Child Left Behind" law mandates external testing of every grade every year. RAMA's approach, especially in feeding back information to teachers, is worthy of replication in other countries. RAMA carefully monitors the test taking process and reports incidences of cheating or manipulation. A major exception to best practice is the fact that most ultra-orthodox students, who account for around 20% of primary school enrollment, do not study the common core curriculum and do not take the Meitsav exam. The Government is making a serious effort to change this practice (see, for example Dattell and Lior).

Many countries publish school by school testing results, while others keep this information confidential. This issue is often hotly debated around the world and countries follow different paths (see McKinsey, p. 55). The proponents of publication argue that information on school results will lead to pressure on poorly performing schools to improve. The detractors say that test score information is often unfairly used to "blame" teachers for poor student performance, and can lead to migration of middle class students away from some schools, leading to greater social stratification. In fact, the nature of the impact on parents' decision making depends on how these results are presented. When a school's clientele consists of children of middle and upper class families, then it is expected that students will score well on measures of achievement. To be more useful, individual school data should be presented in a way that connects the socio-economic background of students with scores. The schools which have the highest scores compared to "expected" scores, given the socio-economic status of their students, should be defined as "effective" schools. The information that Israel feeds back to individual schools does help

them to recognize the impact on learning of the socio-economic status of their students (see RAMA3 for an example of feedback to an individual school).

Given the technical high quality of the Meitsav, its focus on the positive elements of learning assessments, and the likely limited impact of published school by school results, there was no good reason for canceling the 2013 Meitsav. If Israel were eventually to abandon the Meitsav completely, it would not be able to determine whether its education policies have an impact on learning. Sample based assessments, similar to the NAEP (National Assessment of Education Progress) in the US, would measure the extent to which Israel is reaching national goals, but schools would not know how they performed compared to the nation as a whole. In any event, to ensure continued appreciation of its value, the Meitsav ought to be much better marketed to Israel's stakeholders, education decision makers, and the public at large. It would also be worthwhile to undertake a study of the impact on parents of publication of individual school results. In addition RAMA's efforts to measure school climate are an innovative effort to measure "values" in the school system, which ought to be continued. These efforts could be expanded to cover civics and attitudes towards the state and its history, but the task is made difficult by a lack of consensus in Israel on many of the specific goals in these areas.

Israel's Participation in International Tests (TIMSS and PISA)

Experience to Date

TIMSS. Mathematics and literacy assessments have become internationally recognized as important not simply because they are relatively easy to measure but also because they are believed to be fundamental for a productive life in the modern world. In keeping with this awareness, Israel has regularly participated in the TIMSS, which measures mathematics and science achievement, the PIRLS (literacy) tests of the IEA, and the PISA tests of language, math, and science managed by the OECD. The results of the TIMSS 2011 assessment were reported in November 2012. The 40 participating countries worked together to design a comprehensive framework based on mathematics concepts and skills that are taught in the majority of participating countries. The test is organized around two dimensions. A "content dimension," specifies the domains or subject matters to be assessed within mathematics, which at the eighth grade level include Number (30%), Algebra (30%), Geometry (20%), and Data and Chance (20%). A "cognitive dimension" – "knowing" (35%), "applying" (45%), and "reasoning (25%)," is defined within each of these domains, so as to measure thinking processes expected of students as they engage with the mathematics content. Given the framework's broad coverage goals, the mathematics assessment item pools are large—217 assessment items at the fourth and eighth grades, respectively—with about half being multiple choice and half constructed response items where students write their answers. Each student is tested on a small percentage of these items, and statistical techniques are used to ensure all results are comparable. Another test, the PIRLS (Progress in Reading Literacy Study) measures reading skills of fourth graders. While usually only 4-8000 children in each country are tested, TIMSS staff have an elaborate set of controls to ensure that samples are randomly selected and representative of the country as a whole (see IEA1).

In the 2011 TIMSS mathematics test given to eighth graders, Israel raised its score from 463 in 2007 to 516 in 2011, an increase of 53 points,² landing 6th in the world instead of 24th four years earlier (IEA1).

² Over half a standard deviation.

Israeli Arabs also improved their scores, which are now higher than any other Arab country, although still significantly lower than Jews. Israel’s increase in scores from one test to another was one of the highest in the 45 year history of the TIMSS assessment. Between 2007 and 2011 the only countries with similar improvements were Qatar and Saudi Arabia but they began at a much lower level. Among middle level countries, Russia had the next greatest increase (27 points) after Israel. Table 1 summarizes Israel’s scores in mathematics in the years since 1999 for all those who took the tests as well as for Jews and Arabs.

Table 1. Israel’s TIMSS Math Scores 1999 to 2011

	Arabic	Hebrew	National
1999	397	482	466
2003	465	505	496
2007	408	484	463
2011	465!	536!	516!

Sources: IEA1 and 2, TIMSS 2007 and 2011. In Israel schools are separated between those where the language of instruction is Hebrew or Arabic. A very small percentage of Arabs attend Hebrew speaking schools.

Israel’s scores in the TIMSS science assessment also improved significantly (from 455 to 516) (IEA2), and in the Progress in Reading Literacy test (PIRLS) given to fourth graders (from 512 to 541) (IEA3).

This improvement triggered a very lively discussion in Israel. In newspaper articles and among the general public there were suspicions of cheating and/or manipulation of the system, including biasing the sample, asking weaker performing students to stay home, and adding classes, time, and test prep for the sample schools.

Beyond individual opinions, there is no evidence of cheating or manipulation. According to RAMA staff, the sample design was prepared by the international agency overseeing the assessment. The small number of schools that refused to participate in the assessment was replaced by schools in the same region and at the same socioeconomic level. Names of students in classes were identified in advance so as to ensure that these same students took the tests. Outside proctors were hired to oversee test implementation. The schools themselves were informed only one and half months before the test administration that they were in the sample, so the time available for test prepping was minimal. The process was administered under the leadership of highly professional staff at RAMA, Israel’s learning assessment agency. Some newspaper articles noted that the results were “distorted” since “ultra-orthodox” students, who comprise 20% of the school enrollment, not only do not study English, mathematics or science but also for the most part do not participate in any national or international assessment (Arlosoroff). But this has been an issue for Israel for many years and so had no incremental impact on the test results.

As described in public documents and explained by the Secretary General at the time, in 2009 the Ministry set specific goals to be achieved, including that of higher academic achievement and “imparting knowledge necessary in modern society” (Beller1). One measure of the extent to which these goals could

be achieved would be an improvement in Israel's scores in international tests. The goal in 2009 was that, within 6 years, Israel would be among the ten leading countries in international achievement.

The TIMSS results were to a great extent the result of a major investment in curriculum reform, teacher support and training begun under former Minister Yuli Tamir (2006-2009) but emphasized more strongly by Minister of Education Gideon Saar (2009-2013). Under the leadership of the Secretary General (Shimshon Shoshani), the Ministry set out to achieve this goal, focusing initially on seventh and eighth grade math and science and fourth grade language, since these were the assessments which would be given in 2011. The mathematics curriculum was revised to include higher order concepts, based to a great extent on the curriculum underlying the TIMSS assessments. An additional hour a week was added for 7th and 8th grade mathematics, teachers were trained to understand the higher learning expectations, and a "learning coordinator" was hired to help individual schools. Seminars impressed upon school principals the importance of the test. The estimated additional cost for this program, according to those responsible, was 400 million shekels (US\$100 million).

The international TIMSS report provides external evidence that many of the actions described by the Minister and the Secretary General of the time did take place.³ In 2011 Israel taught 89% of the subjects in the TIMSS assessment compared to only 69% in 2007. This would be a reflection of attempts to align its mathematics curriculum and standards to the knowledge and skills underlying the TIMSS assessment. Israel is not alone in this process of universalizing mathematics curriculum. The percentage of subjects in the TIMSS test covered by all countries increased from 72% in 2007 to 80% in 2011. The instructional time in Israel given to mathematics increased from 12% of weekly hours in 2007 to 15% (165 hours in the school year) in 2011. This compares with 13% (138 hours) for all countries in 2011.

According to the TIMSS report, 79% of Israel's mathematics teachers participated in professional development courses related to mathematics content in the two years before the TIMSS exam, compared with 59% of Israel's teachers in 2007 and an average of 55% for all participating countries. Israel's collaborative efforts increased dramatically from 2007 when only 9% of teachers met on a weekly basis. In 2011 62% of mathematics teachers in Israel had interactions with other teachers at least 1-3 times per week. This was the highest level of collaboration among all participating countries, where the average was only 28%.

The TIMSS 2011 report does not provide charts comparing Israel's scores between 2007 and 2011, as it does for nearly all participating countries. This was the result of the fact that Israeli authorities had decided to re-translate the Arabic version of the test, since it was felt that the 2007 version was inaccurate. The IEA decided to err on the side of caution, since the text of the Arabic language "anchor items" in 2007 was different from that of 2011. Subsequent research by RAMA has confirmed that the revision in the translation had only a very minor impact on test scores. Perhaps more importantly the 2007 assessment took place shortly after a long teachers' strike which adversely affected the functioning of junior high schools. This may have resulted in fewer hours of teaching and learning and both teachers and students may have paid little attention to doing their best on the assessment (RAMA1).

Israel's 2003 TIMSS score was 496, but there was a 33 point decline to 463 from 2003 to 2007. If all of this decline were the result of teachers' strikes and school disruptions, then Israel's improvement from one test to another would be a more modest but certainly respectable 20 points.

³ The information which follows is from IEA1, TIMSS report, pages 190-310.

PISA 2012. PISA 2012 scores for math and reading were reported in December 2013 (OECD3). Below are the results for Israel since 2003.

Table 2. Israel’s Scores on PISA, 2003 to 2012

	PISA Math			PISA Reading		
	Arabic	Hebrew	National	Arabic	Hebrew	National
2003	344	449	418	378	465	452
2006	372	460	442	372	456	439
2009	367	470	447	392	498	474
2012			466			486

Source: PISA 2012 Results and RAMA

Israel shows real positive improvements since 2003, and is cited in the PISA 2012 report as one of the four or five countries which have improved most over time. Nonetheless it is still below the median for OECD countries, and socio-economic differences in achievement continue to be higher than most OECD countries. A number of OECD countries decided not to participate in the TIMSS, preferring to confine their international testing efforts to PISA. This means that the universe of countries in TIMSS is not as highly achieving as in PISA. PISA assessments focus on applied use of math and reading in real life situations and therefore are not as closely linked with curriculum as is TIMSS. Nonetheless scores on the two assessments are highly correlated, and Israel has improved its position in TIMSS and PISA over time by 50 points in math, which is the rough equivalent of one school year.

Towards Best Practice

International assessments are authentic, reliable, and valid measures of knowledge and understanding. They have had an enormous and mostly positive impact on education policy worldwide. After Germany scored very poorly on TIMSS and PISA ten years ago, it implemented an extensive reform designed to ensure that students in technical/vocational schools received much more intensive grounding in language and mathematics. The development of the Common Core Curriculum in the USA is a direct result of the performance of US students in international tests. In both countries a wide consensus was sought and found before proceeding with these major reforms. At the same time performance on these tests has become increasingly important for countries. One result is that, especially in mathematics, country curriculum goals are becoming increasingly common to all countries. While TIMSS and PISA staff diligently seek to ensure that outright cheating or manipulation does not take place, anecdotal reports of countries “gaming” the system are often heard.

Israel’s efforts made in 2009-2011 to raise learning and scores in the 2011 TIMSS to a great extent used a best practice approach, including focusing on higher order learning skills and the implemented curriculum, encouraging teachers to work together, and increased time on task. There is no evidence of “cheating,” manipulating samples, or asking poorly performing students to stay home.

But Israel's approach had several failings. In the first place it appears that the marketing of the program focused on specific numerical improvements in scores in the international tests given in 2011, rather than on meeting a longer term objective of increasing students' knowledge of mathematics. Israel is not the only country that has taken this focus. But this has meant that teachers may not be motivated to work hard in years when the international assessments are not given and that less emphasis may be placed on learning in the ninth to twelfth grades which are not internationally tested. A second miscalculation was the linkage of the effort to improve scores with the Minister of Education at the time, rather than on an agreed upon consensus among all stakeholders. One result was partisan mistrust leading to a risk that the positive elements of the effort might not be continued. A need for cuts in the education budget will likely result in reduction of funding or even elimination of mathematics and science "coordinators" as well as assessment advisers. A third problem has been a lack of follow-up on the results. RAMA recently published a detailed report analyzing the implications of the TIMSS results for Israel's schools but there was no newspaper coverage of it and only limited efforts to implement its recommendations. If Israel abandons the approaches it took to improve its performance on the 2011 TIMSS, then it will likely fall back into the "middle of the pack" in international testing. Discontinuity in policies could mean that only the one or two cohorts who prepared for the international tests in 2011 would have better knowledge of math, science, and language, compared to subsequent cohorts.

Israel should continue to participate fully in the PISA and IEA (TIMSS and PIRLS) assessments. To end its participation would be contrary to the commitments of every one of its OECD partners and would leave Israel without a capacity to understand what is happening in its schools compared to its competitors. Failure to take performance in these tests seriously would compromise the knowledge capital of future generations. Given the current concern with "values," Israel ought to consider participating in the 2016 Civic Education study of the IEA. This study measures attitudes and knowledge related to tolerance, democracy, respect for others, and belief in the efficacy of the state. Israel could add a module related to the particular issues it faces regarding the role of the state and the relationships among its different ethnic and religious groups. In addition, as part of its desire to measure skills beyond mathematics and language, Israel should consider participating in a new PISA 2015 study of students' collaborative problem solving skills (OECD2).

Israel's End of Secondary Cycle Examination (the Bagrut)

The Current Situation

Israel's end of secondary cycle examination, the Bagrut, is given to most students who complete secondary education and is an important determinant of which higher education institutions they will attend. It is not designed to be a "learning assessment" to impact on policy or pedagogy. Students are expected to take tests in seven subjects with English, math, Hebrew or Arabic, history, and civics required and two optional subjects. Israel has 157 individual Bagrut tests, many of them in narrow focus technical/vocational areas, given over three years on five separate occasions, a number which is among the highest in the world. Individual subjects are given at various "levels of difficulty," which are defined as "units"—e.g. a five unit mathematics test is much more difficult than a "two unit" mathematics test.

In 2010 12% of 12th graders did not sit for the examination, and 9% dropped out before entering 12th grade. 77% of those who took the test scored high enough to be eligible for a "matriculation" certificate (Beller 1). The remaining 23% received a certificate for completion of secondary school which allows them to enter certain non-academic or training courses, but in general they are unable to enter most colleges and universities offering Bachelor's degrees. The "matriculation" rate for each school as well as the country as a whole is carefully followed by parents and stakeholders, and political leaders often seek ways to increase this ratio. One result has been that over time students have been allowed to take the

Bagrut as many times as they wish. Bagrut cram schools have mushroomed to help dropouts and adults to sit for these exams. Recently increasing numbers of students have been requesting special “accommodations” to account for learning disabilities. These accommodations can include extra time, ignoring spelling mistakes, and sometimes giving the test orally. Since the tests are not prepared with “anchor items” from one year to another, it is not possible to confirm whether one test is more or less difficult than a previous one.

Most observers and stakeholders agree that the Bagrut distorts much of the teaching at the secondary level. Too many tests are given in all subjects, beginning in tenth grade, with options for students to take the tests as many times as they wish. In addition, the Israeli testing regime is unusual in that it has both a European style high stakes end of secondary cycle exam as well as an American style aptitude test (the Psychometric Exam”) similar to the SAT which is required for most universities.

Towards Best Practice

Israel’s Bagrut is similar to the Baccalaureate exam in France, the Abitur in Germany, and the O and A levels in the UK. The United States does not have end of cycle exams of this sort except for New York State (the Regents exam) and now a few other states, although the Advanced Placement examinations increasingly play a similar role for high school students hoping to attend selective colleges and universities.

International research has tended to confirm that countries with strong end of cycle exams have students with a deeper knowledge of language, math, and science than those without such exams. At the same time most countries around the world struggle to deal with the negative impacts of high stakes end of secondary cycle examinations. In England students take 7-10 exams after studying for two-three years and receive certificates in individual subjects. These tests determine which higher education institutions they will attend. The French baccalaureate exam is offered in four “streams”, scientific, humanistic, economic and social, and technical. Over 80% receive pass the exam, which guarantees that a place will be found for them in higher education, although over one third do not sit for the exam. France has encouraged the development of a vocational stream for less academically inclined students. The German Abitur tests four or five subjects and is a prerequisite to university attendance. Only 49% of German secondary school students graduate with the Abitur. In South Korea, high school students study day and night, often in costly cram courses (“hogwons”), to get the highest scores on the university entrance exam. Recently Korea has taken steps to reduce the percentage of secondary school graduates who attend universities and encourage them to enter vocational technical institutes. Increasingly in the US “Advanced Placement” courses are needed for acceptance to the best colleges and universities. The New York State Regents exam, which covers four or five subjects, is required for graduation. Recently the number of subjects tested was reduced and the exam will soon be aligned with the Common Core State Standards. Efforts to reduce the negative effects of end of cycle exams include: using teacher evaluations of student learning to count for part of the end of cycle grade; offering more exams which measure vocational/technical skills and knowledge; raising the prestige of post-secondary non-university courses and programs connected with the labor market; providing certificates and diplomas even to those students who do poorly on the exams; reducing the number of tested subjects; and providing extra support for underprivileged students. These efforts are very highly country specific.

In Israel as elsewhere these tests will continue to be necessary for the purpose of selection. Israel could reduce some of the pressure and distortions brought about by the test through: (a) giving students only two opportunities to retake the test; (b) offering Bagrut level tests only in grades 11 and 12 rather than in grades 10-12; (c) raising up the public perception of the value of post-secondary courses that do not necessarily lead to a bachelor’s degree; (d) more carefully controlling poorly documented requests for special “accommodations;” (e) reducing the number of narrow based vocational subjects and broadening

the content of vocational curricula; (f) offering only the option of basic or advanced level tests; and (g) increasing the importance of teacher based evaluation for the final grade. Within the Israeli context, reducing the number of subjects tested may not be advantageous. To reduce inequity, financial support for underprivileged students who do not have the means to pay for private Bagrut prep courses should be increased. To strengthen classroom based assessments, Bagrut results could be fed back to schools, with explanations of problems that students faced in answering questions, presented in such a way that teachers are helped to improve their pedagogy. Efforts could be made to keep some test items confidential so that they could be given in subsequent years as “anchor items” permitting more reliable comparisons of test difficulty from year to year.

In January 2014 the Ministry of Education proposed a reform of the testing regime which incorporates a number of the above suggestions as well others. In the short term the number of mandatory subjects will not be reduced, but the number of required exams will be reduced from 15 to 10. English and math will have two optional exams rather than three. Efforts will be made to reduce the near universal requirement to sit for the “Psychometric” exams for university entrance (see Skop, *Haaretz*).

Conclusion

Israel has done well in using its learning assessments effectively and in facing issues of unintended consequences in their use. It should not abandon its progress to date but instead ought to consider how to effectively market and better use its assessment programs “in the service of learning.” It should treat the results of international assessments as an opportunity to improve its implemented curriculum, school management, and teacher training so as to promote higher order learning. In addition efforts currently under discussion to simplify the Bagrut and reduce the “backwash” impact of these exams on the entire secondary school system are beginning to show positive results. Israel’s education leaders should recognize that there is no contradiction between subject matter knowledge and creativity. In fact, while it is true that those with subject matter knowledge are not necessarily creative, the opposite is not necessarily true. It is likely that a substratum of deep knowledge in a field is one of the prerequisites to creativity in that same field.

Overall Israel needs to improve the quality of its primary and secondary school system. Teachers are relatively underpaid and have difficult working conditions. Demographic challenges include the increase in numbers of ultra-orthodox students, many of whom do not study English, mathematics, or science, as well as of Arab students, who perform significantly worse than their Jewish counterparts. Israel’s government has taken important steps to improve teacher working conditions, provide pre-schooling to underprivileged children, increase investment in schools with Arab and other underprivileged populations, and encourage the ultra-orthodox to broaden their curriculum. But if Israel is to have an increasingly educated citizenry, it must also measure the extent to which it is achieving its goals. While it is certainly true that an education system should be people centered and values oriented, in today’s world it must also be data driven.

Bibliography and References

Arlosoroff, Meir, “Haredi Education is Dragging Israel Closer to the Third World,” *Haaretz*, Jan 7, 2013.

Beller, Michal (1), “Assessment and Evaluation of the Israeli Education System,” Power point Presentation. Ramat Gan, Israel, April 2013.

Beller, Michal (2), “Assessment in the Service of Learning: Theory and Practice,” RAMA, Ramat Gan, Israel, March 2013.

Campbell, Donald, “Assessing the Impact of Planned Social Change,” The Public Affairs Center, Dartmouth College, 1976.

Dattell, Lior, and Yarden Skop, “Shas schools to be part of public school system,” *Haaretz*, January 9, 2014

Erez, Ran, “Reducing the number of exams,” www.irgun-hamorim.org.il, June 13, 2013.

Harkov, Lahav (1), “Ex-minister assails successor for banning test,” *Jerusalem Post*, August 14, 2013.

Harkov, Lahav (2), “Piron: encourage students to excel with love, not exams.” *Jerusalem Post*, August 27, 2013.

Hartman, Ben, “Israel children post impressive improvements,” *Jerusalem Post*, Dec. 12, 2012.

Hutchison, Dougal, and Ian Schagen, “Comparisons between PISA and TIMSS –Are We the Man with Two Watches?” National Foundation for Educational Research, n.d. n.p

IEA1, Mullis, Ina, Michael O. Martin, Pierre Foy, and Alka Arora, TIMSS 2011 International Results in Mathematics, Boston College and IEA, Chestnut Hill, MA, 2012

IEA2, Martin, Michael, Ina Mullis, Pierre Foy, and Gabrielle M. Stanco. *TIMSS 2011 International Results in Science*, Boston College and IEA Chestnut Hill, MA, 2012

IEA3, Mullis, Ina, Michael Martin, Pierre Foy, and Kathleen Drucker, PIRLS 2011 International Results in Reading, Boston College and IEA, Chestnut Hill, MA, 2012.

IEA4, Mullis, Ina, Michael O. Martin, and Pierre Foy, TIMSS 2007 International Mathematics Report. Boston College and IEA, Chestnut Hill, MA 2008

IMRA (Independent Media Review and Analysis), “Israeli pupils jump ahead in world ranking data from international TIMSS and PIRLS examinations show: dramatic increase in achievements of Israeli pupils – Hebrew and Arabic speakers in all fields,” December 24, 2012 <http://www.imra.org.il/story.php?id=59524>

Kashti, Or, "Israeli eighth graders math scores plummet," *Haaretz*, October 16, 2012.

Koretz, Daniel, *Measuring Up, What Educational Testing Really Tells Us*, Harvard University Press, 2009.

Mckinsey and Co., "How the world's best-performing school systems come out on top," September 2007.

National Center for Education Statistics, "Comparing NAEP, TIMSS, and PISA in Mathematics and Science," 2004, Washington DC.

OECD1, *PISA 2009 Results: Volume I. What Students Know and Can Do: Student Performance in Reading, Mathematics and Science*, OECD, Paris, 2010

OECD2, "PISA 2015, Draft Collaborative Problem Solving Framework," March 2013, Paris

OECD3, *PISA 2012 Results: What Students Know and Can Do*, December 2013, Paris

RAMA1, Ministry of Education, "Key Findings from Research TIMSS 2011," Ramat Gan, Israel, 2012

RAMA2, Ministry of Education, "Trends in Mathematics and Science Study, 2011, An Israel Perspective" (in Hebrew), Ramat Gan, Israel, 2013

RAMA3, "Meitsav 2012, Efficiency Measures, School Growth, Achievement" (in Hebrew), Ramat Gan, Israel 2013.

RAMA4, "Meitsav 2012, Report to Weitzman School, Jaffa (in Hebrew), Ramat Gan, Israel, 2013.

Skop, Yarden, "Israel suspends standardized tests," *Haaretz*, August 13, 2013.

Kop, Yarden, "Education minister unveils plan to streamline school testing, college admission requirements," *Haaretz*, Jan 9, 2014.

The Economist, "Education in South Korea: Class Struggle," October 29, 2013.

Nesher, Talila, "Proficiency test scores for Israeli students on the rise for first time ever," *Haaretz*, October 5, 2011.

UNESCO Institute for Education and Center for Universal Education, Brookings Institute, *Toward Universal Learning: Recommendations from the Learning Metrics Task Force*, 2013

Velmer, Tomer, "Meitsav exam results show improvement in all subjects." *YNET News*, October 4 2011.

Wolff, Laurence, with Elizabeth Breit, "Education in Israel: the Challenges Ahead," Israel Studies Center, University of Maryland, May 2012, www.israelstudies.umd.edu.