

SCIENCE AND MATHEMATICS EDUCATION CENTRE

**THE EFFECT ON TEACHERS OF USING MATHEMATICAL INVESTIGATION
TASKS AS TOOLS FOR ASSESSMENT**

Jeanne Albert

This thesis is presented in fulfillment
of the requirements for the award of the degree of
Doctor of Philosophy
of the
Curtin University of Technology

June 2002

Declaration

I certify that this thesis does not incorporate without acknowledgement any material previously submitted for a degree or diploma in any university; and that to the best of my knowledge and belief it does not contain any material previously published or written by another person where due reference is not made in the text.

Signed:

Acknowledgements

I would like to acknowledge the guidance, support and friendship provided by my supervisor, Professor John Malone. The feedback which I received was always very helpful and promptly delivered, which encouraged me to press on to completion.

I am grateful to all the teachers and teacher-leaders who participated in the School-Based Assessment Project and whose cooperation and feedback made the study possible. Their professionalism and their willingness to be involved made my work with them enjoyable and highly rewarding.

Finally, I would like to thank my family and my friends who, each in their own way, gave me encouragement and support throughout this project.

Abstract

This study set out to determine the relationship between assessment practices and teaching methods. I wanted to investigate whether making mathematical investigation assessment tasks available to elementary-school mathematics teachers would have a positive effect on their teaching. Research tells us that standardized tests influence instruction. My research explored whether a national Assessment Task Bank of mathematical investigative tasks could influence teachers.

With these aims in mind, the following research questions were formulated:

1. Will the teachers' use of mathematical investigation tasks for assessment purposes influence their view of mathematics?
2. Will the teachers' use of mathematical investigation tasks for assessment purposes influence the way they teach, and if so, in what ways?
3. Will the teachers' use of mathematical investigation tasks for assessment purposes influence the way they assess their students, and if so, in what ways?

My research was divided into two parts: 1) a national study involving teachers-leaders throughout the country; and 2) an intensive study in a small Israeli community, called Sharon. The first part examined how the national courses on assessment that I conducted affected the participating teacher-leaders in terms of their concept of mathematics, their teaching methods and their assessment practices. The second part examined the same issues with regard to the mathematics coordinators in the Sharon community. In each case, I have detailed my experiences so that the reader can gain a view of all facets of the study.

The research methodology adopted was based on a constructivist paradigm, sometimes referred to as a "naturalistic inquiry", utilizing ethnographic principles wherein the data collection and analysis procedures were eclectic. In the course of the five years of my research, I used many strategies of data collection – for example, unstructured participant-observations, interviews, questionnaires and content analysis of artifacts (tests and tasks written by teachers).

The ideas of reform mathematics (as defined in Ch 2 of this thesis) are based on a broadened vision of mathematics with emphasis on higher-order thinking. My research indicated that the use of mathematical investigation tasks helped the teachers in my study reach the awareness that mathematics, even on the elementary school level, involves generalizations, justifications and even creativity.

Prior to my research, and because of my position, I was aware that Israeli teachers were concerned primarily with teaching routine procedures and that their work sheets for the most part involved single-answer exercises. My research indicated that the use of mathematical investigation tasks indeed influenced the way teachers teach. Verbalization—having the students explain “why”—has become integral to the teaching practices of the participants in my study. Nowadays, the Israeli teachers I worked with use “authentic tasks” in their classrooms: real-life situations that involve some mathematics. Unfortunately, these tasks are not always planned properly.

My research demonstrated that teachers attending my professional courses found the mathematical investigation tasks to be useful for assessment purposes, providing them with additional information about their pupils, not obtainable through conventional assessment methods. The additional criteria (I introduced) for evaluating the pupils’ work aided in defining these additional areas. I found that while teachers were quite willing to use the mathematical investigation tasks to supplement the conventional tests, they were reluctant to use them as replacements.

Exposure to the Assessment Task Bank influenced, to a certain degree, the way the teachers in my study assessed their students. The tests of the teachers who were participants in my study now regularly include elements that were previously absent: questions requiring explanations and questions with more than one possible answer.

Although the teachers of my study were increasingly using questions that required higher-order thinking, the tendency was to use the tests in a summative manner, rather than formatively. In other words, many teachers found it difficult to use test results for planning their subsequent lessons. While they were able to analyze their students’ work and could report in some detail on each student’s performance, they failed to understand how this should affect their teaching. Before they were exposed to the tasks they had administered tests merely in order to provide grades, whereas now the teachers were often trying to understand the students’ thinking.

While long-term change is still elusive, my research has demonstrated that exposure to reform mathematics through the mathematical investigative tasks of the Assessment Task Bank did have some influence on the teachers’ view of mathematics, as well as their teaching and assessment practices.

Table of Contents

Chapter 1: Introduction	1
1.1 My Professional Background	1
1.2 The Background for my Research	5
1.3 Mathematics Education in Israel	8
1.4 Assessment in Israeli Education	10
1.5 The Tasks and their Grading	14
1.6 Aims of the Study and Research Questions	17
1.7 Research Methodology	17
1.8 The Significance of the Study	19
1.9 Overview of the Study	20
Chapter 2: Review of the Literature	22
2.1 The Kind of Mathematics We Want Today	23
2.1.1 Research on How Children Learn Mathematics	24
2.1.2 The Growth of Technology and Its Influence on Society	25
2.1.3 The Failure of the New Math Revolution of the Sixties	26
2.2 Reform Mathematics	28
2.3 Assessment	31
2.4 The Professional Development of Mathematics Teachers	36
2.4.1 Teacher Change	36
2.4.2 Professional Development	38
Chapter 3: Research Methods	44
3.1 The Research Questions and Working Assumptions	44
3.2 Investigative Assessment Tasks	46
3.3 Criteria for Scoring	48
3.4 The Development of the Investigative Assessment Tasks	49
3.5 The Research Paradigm	50
3.6 The Background for the Collection of the Data	53
3.7 Teacher Quotes: Coding Used	55

Chapter 4: Data Collection in The National Study	56
4.1 The Method of Data Collection	56
4.1.1 Summary of Data Collected	56
4.1.2 Construction of the Questionnaires	57
4.1.3 Interviews with Teachers	60
4.1.4 Sample Tasks and Tests	60
4.2 The First National Course on Assessment	61
4.2.1 The First Workshop	61
4.2.2 The Second Workshop	62
4.2.3 The Third Workshop	63
4.2.4 The Fourth Workshop	64
4.3 The Second National Course on Assessment	65
4.3.1 The First Workshop	65
4.3.2 The Second Workshop	67
4.3.3 The Third Workshop	68
4.3.4 The Fourth, Fifth, and Sixth Workshops	69
4.3.5 The Final Workshops (the Seventh, Eighth and Ninth)	70
4.4 The Third National Course on Assessment	71
4.4.1 The First Workshop	72
4.4.2 The Second Workshop	72
4.4.3 The Third Workshop	74
4.4.4 The Fourth Workshop	74
4.4.5 The Fifth Workshop	76
4.4.6 The Sixth Workshop	77
4.4.7 The Seventh Workshop	78
4.4.8 The Eighth Workshop	80
4.5 Follow-up Work with Course Graduates	81
4.5.1 <i>Ofek</i>	81
4.5.2 Additional In-Service Workshops	83
4.6 Interview with Teachers	83
Chapter 5: Data Collection in the Sharon Community	85
5.1 The Method of Data Collection	85
5.1.1 Summary of Data Collected	85

5.1.2	Construction of the Background Questionnaire	86
5.1.3	The Follow-up Questionnaire	88
5.1.4	Interviews with Teachers	88
5.1.5	Interviews with Teacher Coordinators	89
5.1.6	Sample Tasks and Tests	90
5.2	The First Year of the Sharon Project	90
5.3	The Second Year of the Sharon Project	90
5.3.1	Assessment	91
5.3.2	The First Workshop	91
5.3.3	The Second Workshop	92
5.3.4	The Third Workshop	94
5.3.5	Workshops at the Individual Schools	94
5.3.6	Sample Tests	96
5.3.7	Interviews	96
5.4	The Third Year of the Sharon Project	96
5.5	The Fourth Year of the Sharon Project	97
5.6	The Follow-up Questionnaire and Interviews	98
Chapter 6: The Results of the National Study		99
6.1	The Teachers' Views of Mathematics	99
6.1.1	The First National Course on Assessment	100
6.1.2	The Second National Course on Assessment	102
6.1.3	The Third National Course on Assessment	103
6.2	The Way Teachers Teach	108
6.2.1	The First National Course on Assessment	108
6.2.2	The Second National Course on Assessment	110
6.2.3	The Third National Course on Assessment	111
6.3	The Teachers' View of Student Assessment	117
6.3.1	The First National Course on Assessment	118
6.3.2	The Second National Course on Assessment	122
6.3.3	The Third National Course on Assessment	127
6.4	Follow-up Work with Course Graduates	139
6.4.1	<i>Ofek</i>	139
6.4.2	Additional In-Service Workshops	148

6.5	Interviews with Teachers	150
6.5.1	Interview with Teacher BC	150
6.5.2	Interview with Teacher AG	151
6.5.3	Interview with Teacher CJ	153
6.5.4	Interview with Teacher BD	156
6.6	Test/Task Analysis	159
6.7	Conclusions of the National Study	168
 Chapter 7: The Results of the Sharon Study		170
7.1	The Second Year of the Sharon Project	170
7.1.1	Background Questionnaire	170
7.1.2	Sample Tests	183
7.1.3	Interviews	184
7.2	The Third Year of the Sharon Project	192
7.3	The Fourth Year of the Sharon Project	194
7.4	Follow-up Questionnaires and Interviews	194
7.4.1	Teacher ED, Math Coordinator at School O	195
7.4.2	Teacher DL, Math Coordinator at School H	195
7.4.3	Teacher DA, Math Coordinator at School K	196
7.4.4	Teacher DF, Math Coordinator at School A	196
7.5	Conclusions of the Sharon Study	197
 Chapter 8: Conclusions		199
8.1	Implications for Teaching and Learning	204
8.2	The Limitations of the Study	208
8.3	Areas for Further Research	210
8.4	Discussion	212
 References		220
 Appendices		
A	Sample Investigative Tasks	
B	Questionnaires	
C	Teachers' Responses to Questionnaires	
D	Paper from SEMT '01	

CHAPTER 1: INTRODUCTION

This study was an attempt to determine the relationship between good assessment practices and good teaching methods. I wanted to investigate whether making good assessment tasks available to elementary-school mathematics teachers would have a positive effect on their teaching.

It is commonly accepted that standardized tests influence instruction. This is usually considered problematic, because it is generally thought that standardized tests cannot contain “good mathematics” (Silver & Kenney, 1995, p.43). But could the opposite occur? Could the creation of a national bank of mathematical investigative tasks for school-based assessment—tasks which contain “good mathematics”—also influence instruction but in a positive manner? Is it possible by these means to influence the mathematics that teachers teach, the assignments they give, and the types of tests they write? The research presented here addresses these questions.

1.1 My Professional Background

I decided to become a teacher over 30 years ago after completing my B.A. in economics. Despite my supervisor’s efforts to encourage me to teach at the high school level, I preferred to work with younger children. My desire to work with children and help them develop a love for mathematics overrode my interest in higher-level mathematics.

I began teaching junior high school mathematics in the United States in the mid-sixties, during the period of New Math. As a new teacher, looking for ways to deal with my pupils’ behavioral problems, I began to use mathematical games as a way of getting students involved in learning. I was influenced along the way by the conferences held by the National Council of Teachers of Mathematics and other professional workshops I attended, for example on Cuisenaire rods. In addition, my college supervisor had introduced me to professional journals and I continued to read extensively. After eight

years of teaching I decided to “make *aliyah*” (to immigrate to Israel). I continued to teach, both at the junior high and the elementary school levels, and began working with in-service and pre-service teachers as well. I have gradually become increasingly involved in teacher education, to the point where I no longer teach children on a regular basis.

I believe that the most important asset a teacher can possess is respect for the learner. This holds true for pupils of all ages, from kindergarten children to teachers attending professional courses. I also learned early on that the important thing was for the students to be learning, rather than for me to be teaching (a philosophical stance that led to my losing my first teaching job!). I try to listen to each student, to understand what the learner is saying, what he or she understands, with which topics he or she is experiencing difficulty and where his or her interests lie. This is what creates a true dialogue with the learner.

I believe that everyone is capable of learning mathematics. My job is to convince my students of this fact—and nobody needs more encouragement than pre-service elementary teachers who have assimilated the stereotypical view that mathematics is a “difficult subject”. This view is a misconception: after all, most children are interested in numbers and shapes before they start attending school, but somewhere along the line they become convinced that the subject is difficult.

I try to meet my students on their own level and help them progress from that point. To achieve this aim I use principles of a constructivist learning approach: using interesting problem situations, employing varied teaching methods, working with small heterogeneous groups, using manipulatives, games and tools of alternative assessment (Davis, Maher & Noddings, 1990; Delaney, 2001; Duit, Treagust & Mansfield, 1996; Hand, 1996; Malone & Taylor, 1993; Simon, 1995; Tobin, 1993). It seems to me that the most important emphasis in the learning of mathematics should be on thought processes. As Romberg and Kaput (1999, p. 16) wrote, the aims of mathematics teaching are “teaching students to use mathematics to build and communicate ideas, to

use it as a powerful analytic and problem-solving tool, and to be fascinated by the patterns it embodies and exposes". Hiebert et al. (1996, p.12) presented a similar view: "students should be allowed to make the subject problematic. ... Allowing the subject to be problematic means allowing students to wonder why things are, to inquire, to search for solutions, and to resolve incongruities. It means that both curriculum and instruction should begin with problems, dilemmas, and questions for students."

For the last twenty years I have been a lecturer of mathematics and mathematics pedagogy at the David Yellin Teachers' College in Jerusalem, Israel. In this capacity I teach courses in mathematics and mathematical teaching methods to future teachers. I work with them while they gain "in-school practice teaching" before starting to teach on their own, and I work with experienced teachers who return to the college for professional development courses.

I believe that in order to learn to be a teacher one must gain actual teaching experience. One cannot learn to teach by watching someone else teach. Like learning to swim, one must take the plunge into the water!

When I began working at the teachers' college, those preparing to become junior high school math teachers spent most of their "practice teaching" watching experienced teachers in their classrooms and delivering a few lessons under the teachers' guidance. I created a different method—unique to our college—which has now been implemented in all subjects, and not just mathematics. According to this method, student-teachers are responsible for their own class, from their second year of college onwards. There are no regular teachers involved for these lessons. The students function as the class teacher in every aspect. Three or four parallel classes in a school are divided into smaller groups, each comprising some 20 pupils (Israeli classes generally contain approximately 40 pupils), with each student-teacher being assigned a group for one to two hours a week, for the duration of one school year. The trainee teachers teach their groups simultaneously, in the presence of the college supervisor, who helps to resolve any problems that arise. These practice teaching sessions are

followed by a lesson in pedagogy, given by the supervisor, in which the training teachers reflect on the lessons they taught, prepare for the following lessons, write tests and give grades. The topics that the trainees teach are determined at the beginning of the year in cooperation with the class teachers, so that each training teacher has his own responsibilities. This means that if a lesson doesn't go as planned the student needs to contemplate, together with his supervisor, what to do during the subsequent lesson in order to improve the situation. Unlike the conventional mode, the class teacher is not involved. Thus, although teaching only a few hours a week, the student is involved in an actual—rather than simulated—teaching situation. This method has been implemented for the past four years with those teachers who are training to be elementary school math teachers as well.

In addition to my work at the David Yellin Teachers' College, I hold a part-time position at the Department of Science Teaching, the Weizmann Institute of Science, where I work with other mathematics educators, writing curriculum, working on professional development and carrying out research. In this capacity, I belong to a community of mathematics educators, benefiting from all the advantages of peer group support. I am involved both in the writing of curricula (we are currently working on material for teaching 11–12 year olds the topics of ratios and proportions) and in teaching professional courses for teachers (in-service projects like the one in the Sharon community; see Chapter 5).

Arcavi (2000, p.146) recently discussed how one should approach the question of choosing a topic for research and what one gains from the process: “we do research to sharpen our wits and powers of observation of the learning process”. My motivations in undertaking this research were two-fold: on the public level, to make a contribution to the field of assessment in mathematics teaching and on the personal level, along the lines of Arcavi, “to sharpen my wits and my powers of observation”. The questions Arcavi presents—“Dare to ask why? What if? What if not? Can it be otherwise?”—are

questions that I have always asked myself, in a constant effort to improve my teaching and my work with teachers. My research is, in a sense, an extension of these questions.

Like all teachers, I have always been engaged in assessment of my students: by observing them, by asking them questions, by giving them quizzes and tests. But assessment only became of special interest when I was asked to head the mathematics staff for the School-Based Assessment Project (see below). At this point, I began to view assessment—its theoretical background, its application in conventional school settings and potential alternative applications—in a new light.

1.2 The Background for my Research

Many important changes have taken place in the content and teaching of mathematics in the schools over the last two decades. To mention but a few (see Chapter 2, Section 2.2, for an in-depth discussion of these changes and for references):

- Thought processes and problem-solving strategies have received an increased amount of attention, in contrast to computations and the technical aspects of mathematical concepts.
- Monitoring and discussing students' thought processes while they work on complex problems has become increasingly important, as opposed to the previous focus on the production of correct answers to sequences of short exercises.
- The teaching of mathematics has become a complex process, conducted in a wide variety of styles, including individual study, work in homogeneous or heterogeneous groups and full class discussions. This stands in contrast to the traditional method of teaching, which was carried out in a "frontal classroom", with the entire class listening to the teacher.
- There is a desire for authentic, contextualized tasks.

- The emphasis has shifted away from computations and algebraic manipulations as a result of the growing use of technological tools.
- The common denominator of all these changes is the requirement that “higher-order thinking skills” be the focus of instruction in mathematics.

All learning, of course, involves thinking, but traditional instruction in mathematics focused primarily on learning to name concepts and to follow specific procedures. Now, the emphasis for all students must shift to communication and reasoning skills. These skills are now popularly called “higher-order thinking skills” (Romberg, Zarinnia & Collins, 1990, p.22). Resnick (1987, pp.2-3) listed features of such higher-order thinking skills, which, as Romberg et al. (1990, p.22) claim, stand in stark contrast to the criteria currently used for assessment of mathematical ability:

1. Higher order thinking is non-algorithmic. That is, the path of action is not fully specified in advance.
2. Higher order thinking tends to be complex. The total path is not “visible”, from any single vantage point.
3. Higher order thinking often yields multiple solutions, each with costs and benefits, rather than unique solutions.
4. Higher order thinking involves nuanced judgment and interpretation.
5. Higher order thinking involves the application of multiple criteria, which sometimes conflict with one another.
6. Higher order thinking often involves uncertainty. Not everything that bears on the task at hand is known.
7. Higher order thinking involves self-regulation of the thinking process. We do not recognize higher order thinking in an individual when someone else calls the plays at every step.

8. Higher order thinking involves imposing meaning, finding structure in apparent disorder.
9. Higher order thinking is effortful. There is considerable mental work involved in the kinds of elaborations and judgments required.

These changes have been implemented, in varying degrees, in the domains of curriculum, learning and teaching, but considerably less has been done in the domain of assessment. As expressed by Romberg et al. (1990, p.23), "Assessment procedures need to be developed which portray not only the number of correct answers students can produce, but also the thinking that produced those answers."

A growing gap between the content and style of mathematics teaching and the content and style of assessment has been observed and criticized (MSEB, 1993b; NCTM, 1989, 1995; Shepard, 2000a). Over the past few years, considerable efforts have been made to close this gap by designing assessment tools and methods that correspond to the above-mentioned developments in the teaching of mathematics (Burton, 1996; Pandey, 1992). The adaptation of assessment tools to the new style of mathematics teaching involves the following (NCTM, 1995; Shafer & Romberg, 1999):

1. Changing the assessment items. A different battery of assessment items is now being developed in order to reflect the emphasis placed on the meaningful learning of concepts and on processes of thinking, modelling and problem solving.
2. Enriching the assessment tools. A wide variety of styles of tests and additional tools of assessment, such as journals, portfolios and extended investigations, is being developed in order to collect more information about student knowledge, achievement and learning abilities, as well as about teaching methods and curriculum.

3. Changing the methods of analyzing student work. Qualitative assessment of students' mathematical processes has been added to the previously used statistical methods of analyzing quantitative data.

The Balanced Assessment Project (1995, p.2) enumerates six goals of worthwhile assessment which are similar to those that were adapted by the School-Based Assessment Project:

1. Assessment focuses on important, grade level appropriate mathematics;
2. Assessments are themselves worthwhile learning activities—not digressions from learning;
3. The assessment maintains a focus on accessibility and equity for all students;
4. Assessment elicits scorable, informative student work;
5. As a whole, the assessment provides a way for students to show both the breadth and depth of their mathematical accomplishment;
6. The assessment is based on standards that are public and gives students and teachers a way to improve the level of student work.

1.3 Mathematics Education in Israel

Jewish Israeli elementary schools, run by the Ministry of Education, are divided into three streams: the secular, religious and kibbutz regional schools. The Arab Israeli population has separate elementary schools that are also run by the Ministry of Education. In the Jewish schools, the language of instruction is Hebrew and in the Arab schools, it is Arabic. In recent years a number of dual-language schools have been established, catering to Jewish and Arabic children together in classrooms with two teachers, but these are still rare.

The subjects of my research are all teachers in the Jewish school system, although over the past four years the School-Based Assessment Project has been expanded to the Arab sector, based on an Arabic translation of the Assessment Task Bank.

In June 1996, the Ministry of Education of Israel ran a battery of achievement tests in mathematics and language—Hebrew and Arabic in the respective school systems—for children in the fourth and eighth grades throughout the whole country. About 10% of the students in each of these grades was tested, about 200 classes. Together with the achievement tests in mathematics, the teachers filled out a teachers' questionnaire (Ben Simon, 1998). The findings arising from the responses on the fourth-grade teachers' questionnaire may serve to provide an accurate picture of the status of mathematics in the elementary schools. The results reported here relate only to the findings in the Jewish sector. These percentages are based on the 188 teachers who responded to the teachers' questionnaire.

1. Most of the teaching is frontal (66%) and only sometimes takes place in homogeneous or heterogeneous groups (40%). Approximately one-third of the time (35%) students work individually.
2. Most of the teachers (70%) claimed that they teach both arithmetic and geometry on a weekly basis.
3. Almost all the teachers (90%) teach mathematics five hours a week, and most are of the opinion that more hours are needed.
4. The number of teachers who received guidance from a school mathematics coordinator equaled the number who received guidance from an external teacher-leader. Approximately 30% received advice more than once a month, as opposed to some 40% who received advice on a less frequent basis.
5. A total of 30% of the teachers claimed that the students frequently used calculators in class, and 31% claimed that they used computer programs to help teach mathematics. (Ben Simon, 1998, p.86)

Nowadays in Israel, most elementary-school teachers teach mathematics, although their mathematical knowledge is often minimal. Those who did not major in mathematics at the teachers' college where they received certification may graduate with only one year of study in mathematics and mathematical methods (i.e., 120 academic hours). Nevertheless, their views have been affected by the changes in the content and style of teaching described in Section 1.2 (i.e., the emphasis on problem solving and higher-order thinking). In general, when one asks them what is important in teaching mathematics today, teachers tend no longer to favor "skills", but instead to place a relatively high value on "mathematical thinking".

Between 1996 and 1998, I opened each workshop on alternative assessment by asking the participants what they considered to be the most important goals in elementary-school mathematics. The main goal enumerated by all groups was "understanding and thinking". This is a step in the right direction: teachers seem to have assimilated the goals of alternative mathematical education even if these goals are not completely reflected in their teaching.

Another important characteristic of Israeli elementary school teachers is that, in general, they are not fluent in English. Consequently, documents like the NCTM Standards and English-language professional journals and resources are not widely read and have little influence. University and college mathematics educators, on the other hand, attend international conferences and are influenced by world-wide trends.

1.4 Assessment in Israeli Education

The Israeli Ministry of Education ran national achievement tests in mathematics and language (Hebrew or Arabic, depending on the child's native tongue) for all fourth- and fifth-grade students during the 1990–91 school year and for all third-grade students in the 1991–92 school year. These tests consisted entirely of multiple-choice questions and placed an emphasis on basic skills. Each child was considered to have

“passed” if he had fewer than five mistakes and to have “failed” if he had at least five mistakes. In addition, schools and local school districts were evaluated on the basis of their students’ achievements, with the national newspapers printing maps with red marks showing which regions had “failed the test”! The result was a subsequent increased emphasis on basic skills, with teachers drilling their students on similar tests.

This situation generated much dissatisfaction among Israeli educators. In 1993, a new Chief Scientist was appointed to the Ministry of Education and a committee was established to determine how to run national tests in the future.

Nowadays, the Ministry of Education carries out three types of assessment at the national level:

1. Matriculation exams in the final years of high school;
2. National assessment at the elementary and junior high school levels;
3. School-based assessment with the use of national assessment task banks.

These types of assessment are described below.

1. The Israeli matriculation exams are compulsory for those who wish to receive the high school graduation certificate. Success in these exams is a prerequisite for anybody who wishes to be eligible for higher education. These matriculation exams have a strong influence on the subject matter that is taught in high school.

2. National assessment tests were introduced in elementary and junior high schools in order to provide the Ministry of Education on data regarding the extent to which its educational goals are attained. The tests were administered to a representative sample of the student population. The identity of the schools and of the students remained confidential, in an effort to reduce anxiety.

The stated purposes of this national assessment test are:

- To provide information on a national scale regarding the extent to which the mandatory curriculum is being taught and the extent to which the students understand, retain and assimilate it.
- To provide such information to the executive board of the Ministry of Education, to the Knesset education committee, to school districts, to the schools themselves and to the general public.
- To focus public attention on the national curriculum.

This exam is administered in a four-yearly cycle, with different subjects from the curriculum elected each year. The subjects that under such tests are: Hebrew, Arabic, Mathematics, English, Sciences, Social Sciences and Biblical Studies.

A major change has been introduced in this exam in recent years is the attempt to assess performance, and not merely achievement. Open-ended questions and items requiring higher-order thinking are now incorporated in a greater proportion than ever before.

3. The school-based assessment program deals with exams administered within and by schools. The program is based on the use of assessment task banks, which have been developed by expert teams, each of which comprises experts in assessment and experts in the specific subject matter. The actual use of items from the banks is up to the individual teacher. The assessment task banks contain authentic performance assessment items, whose purpose is to stimulate the mind, to urge students to confront complex contextualized tasks, to cultivate open expression of views and to encourage the integration of different fragments of information. In short, the goal of these banks is to promote the use of higher mental functions (Nesher, 1996).

This is the official policy of the Israeli government today. In 1994, the Ministry of Education decided as a matter of policy that most assessment should take place on the school level, rather than be carried out on a national or regional scale. Thus, it

established the School-Based Assessment Project. This change of direction was based on the recognition that although the national government needed to know the achievement levels of students, there was no need to test each and every student to determine average levels. Henceforth, national evaluation would be based on samples. In order to aid the schools with their new role in assessment, the Ministry of Education funded a project whose purpose was to create task banks for use in school-based assessment. As mentioned earlier, I was chosen as director of the team that was to develop assessment task banks for elementary school mathematics.

As a result of these developments, the team faced an opportune situation, in which we had teachers beginning to be aware of the new changes in mathematics teaching, coupled with governmental provision of a bank of mathematical tasks intended for school-based assessment. The staff of the School-Based Assessment Project decided to take advantage of these optimal conditions. We decided to create mathematical investigative tasks which, we hoped, would influence these changes and help promote the teaching of reform mathematics (see Chapter 2, Section 2.2, for definition).

Although, as previously stated, many elementary-school teachers would attest that the goal in teaching mathematics is “understanding and thinking” (see Section 1.3), most elementary schools and teachers continued to use conventional tests for evaluation purposes. This means that the teachers based most of their formal decisions—such as grading, reporting to parents and principals—on tests that consist of short-answer questions, mainly stressing low-level thinking skills. The distinction between “higher-order thinking skills” (described in Section 1.2) and “low-level skills” is a crucial one. It is described by the NCTM *Assessment Standards* (1995) as follows:

A shift in the vision of learning mathematics toward investigating, formulating, representing, reasoning, and applying a variety of strategies to the solution of problems—then reflecting on these uses of mathematics—and away from being shown or told, memorizing, and repeating. This represents a shift from mechanical to cognitive work

and also assumes the acquisition of a healthy disposition toward mathematics. (p.2)

It turns out that, despite the teachers' awareness of the importance of mathematical thinking and understanding, this awareness was not reflected in their assessment tools. The aim of my research was to examine whether the provision of alternative tools of assessment—in the form of the assessment task bank—could ameliorate this situation.

1.5 The Tasks and their Grading

The mathematics investigative tasks created by the staff of the School-Based Assessment Project took into account the above-mentioned changes that were occurring in the mathematics community and the realities of the Israeli mathematics curriculum. We created tasks which on the one hand, Israeli students could succeed at, but which on the other hand, would encourage their teachers to place more emphasis on higher-level thinking (see Appendix A for examples of tasks).

Our guiding principles were the following (Albert & Friedlander, 1995, p.4; 1997, p.4):

1. To present mathematics as a subject that deals with authentic, worthwhile problems;
2. To emphasize the process of solving the problem, rather than the final solutions;
3. To emphasize processes of higher-level thinking, such as generalization, justification and estimation;
4. To downplay the use of calculation techniques and to encourage the use of calculators;
5. To introduce the students to unknown problems without algorithmic solutions;
6. To introduce the students to problems with many solutions;
7. To integrate various branches of mathematics;

8. To enable each student to attain his or her mathematical potential;
9. To enable even those students whose mathematical abilities are low to experience mathematical investigations.

We defined authentic tasks as tasks that require the use of mathematical knowledge and skills in a meaningful way. These could be real world, everyday practical tasks, or fanciful, abstract, or weird tasks. The crucial element is the use of “worthwhile” mathematics.

These ideas were based on principles that we found in similar projects (Heuvel-Panhuizen, 1996; MSEB, 1993a; NCTM, 1989; Pandey, 1991; Schwartz et al., 1995; Silver & Lane, 1993). The mathematics that we were aiming to promote can be divided into two types: “insights into the structure of mathematics” and “strategies or methods for solving problems” (Hiebert et al., 1997, p.23). Whether or not the task was derived from the daily experience of the student was not an important factor. As expressed by Hiebert et al. (1997):

Whether or not tasks leave behind something of mathematical value does not depend on how they are packaged—whether they are large-scale, real-life projects, clever puzzles, or simple, unadorned tasks. What matters is that, as students engage the tasks, they reflect on mathematical relationships and ideas. (p.162)

Another way we influenced the Israeli teachers was by teaching them to use rubrics (or pre-determined criteria) for the grading of their students’ performance on investigative tasks. This evaluation method involves basing student assessment on detailed profiles of the entire problem-solving process, rather than relating solely on the final results. These profiles were based on a set of well-defined criteria of assessment which were shared with the students. (See Chapter 3, Section 3.3.) The dimensions chosen for the criteria reflected the essential qualities of good performance in the specific subject (Herman, Aschbacher & Winters, 1992, p.57).

In the opinion of many researchers (Clarke, 1992b; Greenwood, 1993; Lambdin Kroll, Masingila & Tinsley Mau, 1992; MSEB, 1993b; NCTM, 1989), the traditional evaluation methods of attaching a numerical grade to the student performance on complex mathematical tasks fail to capture the wide variety of aspects of their work. Conversely, this innovative method of analyzing the students' work using criteria-based profiles enables teachers to obtain a rich diagnosis of mathematical abilities and dispositions, and consequently to determine what sort of intervention is needed.

We decided that by providing such criteria for assessment we could guide the teachers in evaluating important aspects of their students' work. The criteria we developed are the following (see Chapter 4, Sections 4.2 and 4.3, for the process of establishing these rubrics):

- Understanding the task;
- Understanding the concepts (in each task the relevant concepts are specified- e.g., the meaning of multiplication, the order of operations);
- Mathematical thinking (e.g., reversed thinking, generalization, justification);
- Calculations;
- Communication (written explanations, ways of displaying the solution);
- Variety and creativity.

Much literature is available today on creating rubrics (see, for example, Bush & Leinwand, 2000; Freedman, 1994; Herman et al., 1992). It has been demonstrated that such scoring is reliable (Baxter, Shavelson, Herman, Brown & Valadez, 1993).

Thus, we have teachers who were aware of the new changes in mathematics teaching, investigative tasks which incorporate these changes, and criteria for evaluating the students' performances. The question was whether the teaching practices of these teachers could be influenced by teaching them how to use the tasks for assessment purposes.

1.6 Aims of the Study and Research Questions

The main aim of this study was to determine the effect on teachers of using mathematical investigation tasks as tools for assessment. As mentioned above, it is accepted that traditional achievement tests influence instruction (Linn, 2000; Romberg et al., 1990)—in a negative fashion. My aim was to investigate whether the reverse holds true: whether good assessment tests can have a positive effect on instruction.

A secondary aim was to evaluate the efficacy of using criteria for the analysis of students' solutions on these investigative tasks. As described above, analysis by means of criteria requires basing student assessment on detailed profiles of the entire problem-solving process, rather than relating solely to its final results.

In order to provide a focus for these aims, the following research questions were formulated:

1. Will the teachers' use of mathematical investigation tasks for assessment purposes influence their view of mathematics?
2. Will the teachers' use of mathematical investigation tasks for assessment purposes influence the way they teach, and if so, in what ways?
3. Will the teachers' use of mathematical investigation tasks for assessment purposes influence the way they assess their students, and if so, in what ways?

1.7 Research Methodology

My research was divided into two parts: 1) a national study involving teachers throughout the country; and 2) an intensive study in a small Israeli community. I began by examining what was happening in the teaching and assessment of mathematics throughout the country, through national courses on assessment. The first nationwide workshops began in the spring of 1995. However, I discovered that in order to determine how teachers were reacting to the use of reform mathematics and how this

affected their teaching and assessment methods, it would be necessary to undertake a more intensive study. Thus, I embarked upon an intensive study in a community I will call Sharon where I was already working and where teachers had already been exposed to the concepts and ideas of reform mathematics. School-based assessment in mathematics was introduced to the community of Sharon in December, 1996—a year and a half after the first national course on assessment.

My research recounts the progression of the national courses on assessment and examines how these courses affected the concept of mathematics held by the participating teacher-leaders and the assessment practices they practice. Similarly, I recount the progression of the study in the Sharon community and its effect on the mathematics coordinators. In each case, I have detailed my experiences in order to provide the reader with a view of the entire picture.

This research project was based on a “naturalistic inquiry” paradigm (Lincoln & Guba, 1985; Moschkovich & Brenner, 2000).

In the course of the five years of my research, I used many strategies of data collection: unstructured participant-observations, interviews, questionnaires and content analysis of artifacts (tests and tasks written by teachers). All interviews were taped and subsequently transcribed. The teachers’ responses to each question on every questionnaire were translated from Hebrew into English, analyzed and classified into categories. Conclusions were then drawn from these classifications. The interview transcripts were similarly analyzed.

Most the data collected was qualitative. For each of my aims and accompanying main questions, I had corresponding questions on questionnaires and interviews. In addition to questionnaires and interviews, I collected evidence of instructional practices—that is, examples of student assignments and samples of teacher tests both before and after exposure to the investigative tasks in the assessment task bank.

1.8 The Significance of the Study

This study is significant for three reasons. First, it provides information about the effect the provision of school-based assessment tasks had on teachers. That is, it shows how the provision of teachers with investigative tasks which incorporate higher level thinking affected the types of assignments and tests they set and consequently, provides an insight into what happens in these teachers' classrooms. Second, it gives us information about the teachers' use of criteria for analysis of students' solutions. Third, it gives us information about effective professional development. Information which is lacking, as pointed out by Wilson and Berne, "we have little sense—save the collective and negative self-reports of generations of teachers about traditional in-service programs—of what exactly it is that teachers learn and by what mechanisms that learning takes place." (1999, p.174)

I believe that the results are relevant for various sectors of the mathematics education community. They will help the curriculum developers to develop an awareness of the importance of the materials they provide as examples for the teachers and how the tasks themselves affect the teachers' views of mathematics. They will enable the providers of professional development to learn how to offer more effective courses, thus influencing teachers to teach differently. Ultimately, of course, those elementary-school pupils studying mathematics will learn more when teachers teach reform mathematics.

Since all the background research and theories used in the study are from international findings, the results should have a broader application beyond the confines of Israel. The need for elementary school mathematics teachers to teach their students to achieve higher-level thinking is an international problem.

1.9 Overview of the Study

Chapter 2 contains the theoretical background for the study. This is divided into three issues. The first deals with the mathematical content desired in elementary schools today, often referred to as “reform mathematics” (Section 2.2). The second is concerned with the assessment of students’ achievements and the relationship of these assessment practices to the type of mathematics taught. The third issue relates to the professional development of mathematics teachers.

Chapter 3 describes in depth the purpose of my study and the research questions that were posed. Details of the investigative assessment tasks are provided: their characteristics, the criteria used for evaluating the results, and an explanation of how the tasks were developed. My research paradigm is presented, and followed by the background for the collection of the data.

As previously mentioned, my research was divided into two parts: 1) an extensive study of the country, in which data was collected from workshops given to select teachers throughout the country in the framework of national courses on assessment, and 2) an intensive study of a small Israeli community known as Sharon, in which all teachers attended the workshops.

Chapter 4 provides the methods of data collection for the national study, details about the instruments used for analysis, and about the different workshops given. Chapter 5 provides similar information for the Sharon study.

Chapter 6 provides the results of the national study. This part of my study was based on three national courses on assessment, in which I instructed teachers how to use the investigative tasks of the Assessment Task Bank and on the *Ofek* satellite courses, where the teacher-leaders who had graduated from the earlier national courses on assessment were involved in the planning and running of the workshops. In addition, the chapter provides details of four follow-up interviews with teachers and an analysis of tests collected throughout my research.

Chapter 7 provides the results of the intensive study of the Sharon community, in which all six elementary schools adopted the use of school-based assessment in mathematics. This intensive study was, in effect, a simulation of an ideal situation, in which all teachers and teacher-leaders were exposed to reform mathematics. This community was chosen because the leading teachers of each school were already involved in an in-service project, which placed emphasis on reform mathematics in an effort to transform their concept and vision of mathematics.

Chapter 8 contains the conclusions derived from my research, implications for teaching and learning, limitations of my study, and recommendations for future research.

CHAPTER 2: REVIEW OF THE LITERATURE

This chapter establishes the theoretical framework for my study. In addition to examining and discussing the relevant literary sources, the chapter describes where my research fits into the general body of literature on assessment and on the professional development of teachers, and indicates where it extends our knowledge in this area.

Three issues that dominate the literature provide a focus for this chapter. The first concerns the mathematical content that is desired in elementary schools these days. As aforementioned, the purpose of my mathematical investigation tasks was to influence teachers to teach “reform mathematics”. This chapter defines the term “reform mathematics” and explains how the investigation tasks that I devised helped achieve this goal.

This discussion leads directly into the second issue: the assessment of students’ achievements and the relationship of assessment to the kind of mathematics taught. Much has been written about the often negative effect of assessment on teachers’ practices, the result of the narrow view of mathematics reflected in standardized tests. My research demonstrates how this influence can be transformed into a positive force by using investigation tasks for assessment purposes.

The third issue relates to the professional development of mathematics teachers. The weak knowledge of mathematics of many Israeli elementary school teachers is a related issue here. Since my work is concerned with the effect that the use of mathematical investigation tasks as an assessment tool has on teachers, I needed to develop ways to introduce teachers to these tasks and to instruct them how to use them for assessment purposes. Professional development courses were thus planned with this purpose in mind. I will present in this chapter what is currently known about effective professional development and the ways in which I adapted and modified such courses towards achieving my goals.

2.1 The Kind of Mathematics We Want Today

When we consider what kind of mathematics we want children to learn today, we are influenced by documents such as the United States' National Council of Teachers of Mathematics *Standards* (1989) and the *Standards 2000*, the Australian *A National Statement on Mathematics for Australian Schools* (Australian Education Council, 1991), the English *Mathematics in the National Curriculum* (Department for Education, 1995) and *The National Numeracy Strategy* (Department for Education and Employment, 1999). All these documents call for a reform in the teaching of mathematics and present similar views of what kind of mathematics children will need in order to succeed in the twenty-first century (Burton, 1996; Davis, 1992; Hiebert et.al., 1996; Steen, 1990).

The reform recommended by these documents can be summed up as follows:

1. The emphasis should be placed on students "doing", rather than "knowing that";
2. Authentic contextualized tasks should be provided.
3. Students need to be able to use technological aids, such as computers and calculators;
4. Teachers should have a broader view of the subject of mathematics itself, not just the traditional arithmetic–algebra–geometry–precalculus–calculus sequence;
5. Students need to acquire the tools to enable them to choose between alternative methods;
6. The emphasis should be on higher-order thinking, as opposed to basic computation skills.

As summarized by the American Mathematical Sciences Education Board:

Students should be involved in finding, making, and describing patterns. They should construct mathematical models for both practical and theoretical situations—using technology when appropriate—learning to represent and reason about quantities and shapes, to devise and solve challenging problems, and to communicate what they have learned. Students also should encounter mathematics as a human endeavor,

learning something of its history in various cultures, coming to appreciate its aesthetic side, and understanding its role in contemporary society and its connections to other disciplines and areas of knowledge. (MSEB, 1993b, p.20)

These views of mathematics are often referred to as “reform mathematics”, and this is the use of the term that I subscribe to in this study. The choice of what mathematics to include in these documents was influenced by many factors, including 1) research on how children learn mathematics; 2) the growth of technology and its influence on society; and 3) the failure of the New Math revolution of the sixties.

2.1.1 Research on How Children Learn Mathematics

The last thirty years have produced a wealth of literature on teaching and learning mathematics (e.g., Brooks & Brooks, 1999; Ferrini-Mundy, 1998; Hiebert et al., 1997; Kieran, 1994; Kilpatrick, 1992; MSEB, 1989). One early document, *Objectives for Mathematics Learning* (Avital & Shettleworth, 1968), which outlines a taxonomy of objectives for mathematics is of crucial relevance to this study. In addition to professional journals, there are many international organizations, that provide professional support, some holding annual conferences (e.g. PME, MERGA and ICMI).

Goldin (1993) reports on the need for cognitively-oriented learning as follows:

There is already a wide consensus among researchers and leading mathematics teachers that powerful mathematics learning cannot be achieved through traditional rote methods. Most children do not develop conceptual understanding when they are presented with low-level, discrete skills and algorithmic procedures to imitate, procedures in which they are then given “drill and practice” and on which they are subsequently tested. ... Current research directs us toward far more sophisticated, cognitively oriented approaches: approaches based on the construction of meaning for mathematical ideas, discovery learning and problem solving, the building of powerful systems of cognitive

representation, and the provision of structured learning environments that foster such outcomes. (p. 3)

Noddings (1993) relates to the important role played by teachers:

Turning students loose “to construct” will not in itself ensure progress toward genuinely mathematical results. Teachers must ask questions that challenge ill-formed hypotheses and weak conjectures; they must pose new problems that require the revision of old constructions; and, sometimes, they simply must show how things are done. In the last case, wise teachers take note of their own decision to tell or show and watch for later opportunities to encourage construction. (p. 38)

My study used these summaries of research findings on how children learn mathematics in two ways. First, the teachers learned mathematics and didactics in cognitively-oriented ways. Second, they reflected on this way of learning and applied it to their teaching with children. One specific example of applying Noddings’ ideas (1993) is the importance I placed on the discourse teachers had with their pupils.

2.1.2 The Growth of Technology and Its Influence on Society

The development of technology—and in particular, the widespread use of computers and calculators—has influenced the teaching of mathematics in two main ways. For one, the fact that we now have cheap and widely available machines that can perform mathematical calculations quickly and accurately begs the question whether it is desirable for children to spend hours and sometimes weeks or months learning calculation algorithms, such as how to add simple fractions. Given the time constraint of schooling and the expansion of knowledge in the modern world, educators are asking themselves how to utilize the time allocated to mathematics lessons most efficiently. If we forego the teaching of calculation algorithms, then time is made available for important new domains, such as data analysis and authentic problem solving. The other way in which the development of technology has affected

mathematics teaching is in the consideration of what preparation students need to be given to become productive workers in the modern world. A person who can calculate rapidly and accurately is no longer guaranteed a job; nowadays, the same person needs to know what to calculate, rather than merely how to calculate (Leinwand, 2000). This is widely referred to as mathematical literacy (Paulos, 1988; Steen, 1997), or “numeracy”:

To function adequately in today’s society, mathematical literacy— what the British call “numeracy” —is as essential as verbal literacy. ... without the ability to understand basic mathematical ideas, one cannot fully comprehend modern writing such as that which appears in the daily newspapers.

Numeracy requires more than mere familiarity with numbers. To cope confidently with the demands of today’s society, one must be able to grasp the implications of many mathematical concepts – for example, change, logic, and graphs - that permeate daily news and routine decisions (MSEB, 1989, pp. 7-8).

As Leinwand (2000) explains:

There is ample evidence that, despite pockets of excellence and achievement, the traditional program fails to respond to these changing conditions and that mere tinkering at the edges will not produce the mathematical outcomes we need. (p. 6)

These influences affected my study. All the investigative assessment tasks used put emphasis on authentic problem solving rather than mathematical calculations (see Section 1.5 for definition of “authentic”). This meant that teachers needed to understand the changed priorities.

2.1.3 The Failure of the New Math Revolution of the Sixties

In the 1960s, there was a demand for better mathematics and science instruction in the United States after the Russians succeeded in sending a man into space. Researchers and mathematics departments in the universities, often funded by the National Science

Foundation (NSF), were recruited to write new instructional materials. Many different approaches were used. In general, there were two reasons for the lack of success of these materials: not enough was known about how children learn mathematics and the basic philosophy of “teacher-proof” materials was inappropriate (Blane, 1992; MSEB, 1989). “Questions were raised about the efficacy of the concepts on which the institutes were modeled, with their emphasis on ‘top-down’ instruction by eminent scientists and their focus on subject matter expertise to the neglect of pedagogic technique and learning theory” (Frechtling, 2001, p. 21).

At the same time, there were many attempts to re-educate teachers through in-service programs sponsored by the NSF and its institutes, which aimed at increasing teacher knowledge. These attempts, in general, were not successful. There was little concern from staff about implementation of institute precepts in the school settings in which the teachers functioned and little evidence that participation had affected teacher behavior and student learning and achievement (Frechtling, 2001).

One prominent mathematician, Professor Morris Kline, expounded in detail on the reasons for the failure of these reforms in his 1973 book, *Why Johnny Can't Add—The Failure of the New Math*. He noted, among other factors, the emphasis on rigorous formal logic and on precise language, and the denial of any connection between mathematics and the physical world.

Despite the failure of the New Math revolution, it is worthwhile examining more closely the recommendations of the documents mentioned above, pertaining to the problems of the educational system in the sixties, since these may have some bearing on the situation existing today.

2.2 Reform Mathematics

There is now a fairly widespread consensus on the type of mathematical activity that we would ideally like to see in classrooms. This has been exemplified by Swan (1993, quoting The National Curriculum Council), in the following list:

1. Activities should bring together different areas of mathematics;
2. The order of activities should be flexible;
3. Activities should be balanced between tasks which develop knowledge, skills and understanding, and those which develop the ability to tackle practical problems;
4. Activities should be balanced between the applications of mathematics and ideas that are purely mathematical;
5. Activities should be balanced between those of limited duration and those with scope for development over an extended period of time;
6. Activities should, where appropriate, use the pupil's own interests or questions, either as starting points or as further lines of development;
7. Activities should, where appropriate, involve both independent and cooperative work;
8. Tasks should be both of the kind which have an exact result or answer and the kind which have many possible outcomes;
9. Activities should be balanced between different modes of learning: doing, observing, talking and listening, discussing with other students, reflecting, drafting, reading and writing, among others;
10. Activities should encourage students to use mental arithmetic and to become confident in the use of a range of mathematical tools;
11. Activities should enable students to communicate their mathematics;
12. Activities should enable students to develop their personal qualities;
13. Activities should enable students to develop a positive attitude to mathematics. (p.199)

Wheeler (1993) dealt with the same issues:

... let me suggest that our specific aims in teaching mathematics should be to give every student the opportunity to:

1. Function “like a mathematician”;
2. Develop a repertoire of general mathematical “know-hows”;
3. Appropriate some mathematical knowledge. (p.91)

Wheeler felt that the primary goal should be the first of these three. This goal should not be confined to those students who may go on to become professional mathematicians or the like. The underlying assumption is that everybody is capable, to a certain extent, of thinking and acting “like a mathematician”; one of the aims of schooling is to help students realize that potential. Proposals that include mathematical problem solving or mathematical investigations in the school curriculum are efforts to attain this goal.

The main concepts underlying reform mathematics can be summarized as follows:

- a) More mathematics for **all** students;
- b) New topics;
- c) Learning through investigations—based on theories of constructivism;
- d) Different ways of learning / teaching, heterogeneous classrooms, cooperative group work—based on theories of Vygotskii;
- e) The use of technology in the classroom.

Another aspect of reform mathematics is the requirement for students to learn to “think mathematically”. Different authors define this in different ways. Schoenfeld (1994) writes “Learning to think mathematically means (a) developing a mathematical point of view—valuing the processes of mathematization and abstraction and having the predilection to apply them, and (b) developing competence with the tools of the trade and using those tools in the service of the goal of understanding structure—mathematical sense-making.” (p. 60)

In *Adding It Up*, Kilpatrick, Swafford and Findell (2001) place the emphasis on understanding. “Today it is vital that young people understand the mathematics they are learning. ... Citizens who cannot reason mathematically are cut off from whole

realms of human endeavor. Innumeracy deprives them not only of opportunity but also of competence in everyday tasks. *All young Americans must learn to think mathematically.*" (p. 16)

Ma's analogy with a taxi driver (1999) provides a succinct description of what it means to think mathematically and to have a "profound understanding of elementary mathematics".

Conducting interviews for my study made me think of how people know the town or city they live in. ... Some people—for example, newcomers—only know the place where their home is located. Some people know their neighborhoods quite well, but rarely go farther away. Some people may know how to get to a few places in the town ... Yet they may only know one way to get to these places, and never bother to explore alternative routes. However, some people, for example, taxi drivers, know all the roads in their town very well. They are very flexible and confident when going from one place to another and know several alternative routes. If you are a new visitor, they can take the route that best shows the town. If you are in a rush, at any given time of day they know the route that will get you to your destination fastest. They can even find a place without a complete address. In talking with teachers, I noticed parallels between a certain way of knowing school mathematics and a certain way of knowing roads in a town. The way those teachers knew school mathematics in some sense seemed to me very like the way a proficient taxi driver knows a town. (p.123)

These concepts lead to the central role of the mathematical tasks, which require higher-order thinking from the students (NCTM, 2000).

How they spend their time is determined by the tasks that they are required to complete. The tasks make all the difference.

Students also form their perceptions of what a subject is all about from the kinds of tasks they do. ...If we want students to think that doing mathematics means solving problems, they will need to spend most of their time solving problems. (Hiebert et al., 1997, pp. 17-18)

Good tasks, however, are not sufficient by themselves, as suggested by Kazemi (1998), who demonstrated how the quality of the discourse after the students work on

the task can promote or hinder real learning. Likewise, Stein and her colleagues from the Quasar project described how the cognitive demands of high-level tasks had a tendency to decline during their implementation in the classroom (Stein, Grover & Henningsen, 1996).

The mathematical investigative tasks I used in my research were devised in an effort to require higher-order thinking of the students. The emphasis was placed on “doing mathematics”, rather than recalling information. Teachers were directed to allow calculator use for computations wherever appropriate. Thus, the tasks embody all aspects of reform mathematics. The innovation in this study was their use for assessment as a means to influence teachers to use reform mathematics in the classroom.

2.3 Assessment

Assessment consists of the complete accounting of a student’s knowledge. Any form of assessment has five common features: 1) the situation, task, or question; 2) the response; 3) an interpretation of the student’s response by a teacher; 4) the assignment of some meaning to this interpretation; and 5) the reporting and recording of the findings (Webb, 1993, pp. 3-4). Good assessment is essential for planning future learning experiences:

Because mathematics is a dynamic, interconnected system, students’ knowledge of mathematical concepts and procedures, problem solving, and reasoning develop and mature over a period of years. ... It is essential for effective teaching to know the meanings students are assigning to the mathematical ideas as they are learning to ensure that a solid foundation is being formed. Assessment, then, must be an interaction between teacher and students, with the teacher continually seeking to understand what a student can do and how a student is able to do it and then using this information to guide instruction. (Webb & Briars, 1990, p. 108)

Moreover, “it is through our assessment that we communicate most clearly to students which activities and learning outcomes we value” (Clarke, Clarke & Lovitt, 1990, p. 118).

In recent years, the importance of the role that assessment plays in education has been acknowledged. Many recent books on assessment relate it to learning and cognition (e.g., Linn & Gronlund, 1995; Pressey & McCormick, 1995). Current mathematical handbooks all contain chapters on assessment (Clarke, 1996; Webb, 1992). In addition, there are many books on alternative assessment or performance assessment, some specifically related to mathematics teaching, with chapters on various methods of assessment, such as portfolios, exhibitions and journals. (Ann Arbor Public Schools, 1993; Barber et al., 1995; Blume & Heid, 1993; Brandt, 1992; Johnson, 1996; Kuhs, 1992; Kulm, 1994; Linn & Herman, 1997; Stenmark, 1989, 1991). Textbook series now provide examples of performance assessments (e.g., Glencoe’s *Extended Assessment*, 1995; Houghton Mifflin’s *Assessment Options*, 1995; Manfre, Moser, Lobato & Morrow, 1996), and since the early 1990s, professional journals have continued to feature articles on this important subject.

Assessment is an integral part of reform mathematics (Darling-Hammond, Aness & Falk, 1995; Lange, 1992; Leadership Development and Enabling Change Project, 1994; Leder & Forgasz, 1992; MSEB, 1993b; Rothman, 1995; Webb & Romberg, 1992). Mokros, Russell and Economopoulos (1995) describe the difference between traditional methods of assessment and the requirements imposed by the reform movement:

[The latter] asks that teachers truly examine students’ mathematical work, question students about their thinking, and observe their strategies for solving challenging, multifaceted problems. (p. 84)

Clarke (1997) shows how teachers can use a wide variety of techniques to make assessment an integral and constructive part of their mathematics instruction. Heuvel-Panhuizen (1994) uses examples involving percentages to show that the “problems are

the most crucial part of the assessment". (p. 342) She demonstrates how rich, open problems can reveal student's understanding and abilities.

While standardized, norm-referenced, multiple-choice tests influence teaching in that they emphasize content coverage, fact recall and rote performance of skills, assessment—like curriculum and instruction—could be harnessed to attain different educational purposes: to develop the ability to use critical skills and information in order to "have wonderful ideas" (Duckworth, 1987, p. 1)—to inquire and discover, to pose and solve problems that not only make sense of the world, but can have an impact on it. Assessment can be used as a valuable tool to inform instruction and support powerful learning (Barnes, Clarke & Stephens, 2000; Falk, 2000).

The changes necessary in order to institute reform mathematics are unsuccessful unless they are supported by the national assessment systems (Kulm, 1990). Unless the high-stakes testing strongly supports reform mathematics, teachers will have no incentive to implement these new ideas. School teachers and school systems tend to be conservative, "preserving", rather than transforming culture; thus, if the community is interested in instituting reform mathematics—a new concept for teachers and not easy to implement (Romberg, 1997), then a clear message must be conveyed.

... in spite of impassioned arguments through the years about broader goals for mathematics teaching, teachers—in response to society's demands—have primarily concentrated on computational facility, on the mastery of discrete skills, and on the ability to solve problems similar to those presented in the textbook or in class. Reinforcement for these emphases came from the fact that standardized tests rarely included items designed to assess higher-order thinking. (Lambdin, 1993, p. 11)

Davis and Maher (1993, p. x), comparing "modern" pedagogical methods to "traditional" ones, note that schools used to be thought of as conveyors of information. Teachers (and adults in general) knew things that children did not know, and this knowledge had to be imparted to children. "Telling" was a major pedagogical method—perhaps the primary method. Modern pedagogy, in contrast, views the role

of schools as helping children to develop their own thinking. We need to find yet additional ways to help teachers pay closer attention to children's way of thinking. Lange, Burrill, Romberg and Reeuwijk (1993) challenge us to provide an alternative to traditional evaluation, by using more "open-ended" tasks that make meaningful assessment of mathematics possible. Such an approach to assessment complements pedagogical methods that support the development of thinking in children and a shift in the role of the teacher to coach—one who helps each student to see what he or she is doing, and who helps each student find ways to improve their achievements even further.

There are a number of cases in which national assessment was employed to promote reform, but these are few and far between (Bazzini, 1993; Linn, 2000; Niss, 1993b). One such case, that took place in the U.K. and was successful, is described by Brown (1993):

Teachers have found operating such a system over the whole 11–16 age group to be a complex but rewarding task. In GAIM [*Graded Assessment in Mathematics*] and in similar schemes there have been considerable gains in a number of areas. Teacher professionalism has increased, with teachers becoming much more aware both of the nature of the mathematics they are teaching and of their students' individual achievements and weaknesses. This has often resulted in provision of a curriculum which is more appropriate to students' needs. (p. 76)

Most nations are still sending out mixed signals: on the one hand, documents produced by the professional community push for reform mathematics, but on the other hand, local and national standardized tests are still based on conventional mathematics (Joffe, 1992; Kohn, 2000; Medina & Neill, 1990). The situation in the U.K. is a case in mind:

The UK now seems to be heading back to our previous position where the curriculum becomes subservient to the requirements of regular routine written examinations, which the Cockcroft Committee in 1982 identified as a major cause of low standards of motivation and achievement. (Brown, 1993, p. 82)

Hargreaves, Earl, and Schmidt (2002) examined alternative assessment reform and concluded that it includes:

opportunities to make assessment, learning, and teaching more technologically sophisticated, more critical and empowering, more collaborative and reflective, than they have ever been. Each also highlights the risks that alternative classroom assessment might extend into apparently endless surveillance, degenerate into narcissistic self-indulgence, or crowd out deeper learning and classroom caring. (p. 92)

That is, education systems and educators must choose carefully and think through the results of their choices, and not just assume that “alternative” is better!

This, in addition to increasing demand for teacher and school accountability in many countries, places the teacher in a “no-win situation”.

The methods that teachers use for assessment in their classrooms are influenced by the outside high-stakes testing being carried out (Dossey & Swafford, 1993; Resnick & Resnick, 1992; Shepard, 2000b). There are many articles on the negative effects of high-stakes testing (e.g., Jones, 2001; Nathan, 2002; Paris, 1998). In areas in which a clear, unequivocal message is being conveyed in favor of reform mathematics, with its emphasis on mathematical thinking and problem solving, in-class assessment has broadened to include mathematical investigations and even the use of portfolios (Clarke & Stephens, 1996). Where the high-stakes testing requires only basic manipulative skills and low-level thinking, teachers continue to use conventional tests (Dossey & Swafford, 1993; Niss, 1993b; Swan, 1993).

The connection between learning, teaching and assessment is well founded and has been discussed extensively (Baron, 1996; Black & William, 1998; Ginsburg, Jacobs & Lopez, 1993a; Grouws & Meier, 1992; Howson, 1993; Lambdin, 1998; Morgan, 2000; Niss, 1993a; Shannon, 1999; Shepard, 2000b; Stephens & Money, 1993; Wheeler, 1993). It seems to be reciprocal: good assessment may influence teachers to change their methods, and, conversely, as teaching methods change, the demand for different types of assessment grows (Linn, 1991). This, however, is often a lengthy

process and is dependent on additional factors (Clarke, 1999; Graue & Smith, 1996; Smith, 2000). Moreover, the change in teaching methods generates different kinds of classrooms with better learning going on (Lambdin, 1998). There are many cases of projects in which mathematics assessment was the focus (see, for example, Bright & Joyner, 1998; Herrington, Herrington, Sparrow & Oliver, 1998; Madfes & Muench, 1999; Moon, 1997; Moon & Schulman, 1995; NCSM, 1996). These changes are, of course, gradual, sometimes taking years; nevertheless, changed assessment demands do influence what teachers do in their classrooms, especially if the changes are on high-stake tests, which result in community pressures (MSEB, 1989; Niss, 1993a; Stephens & Money, 1993).

These two main issues form the background for my study: the “new” mathematics that we now wish to implement and the knowledge that “what you test is what teachers teach”. One purpose of the School-Based Assessment Project (see Chapter 1) was to influence teachers to teach reform mathematics. The degree of this influence is one of my major research questions, of importance for extending the knowledge of the mathematics education community. The School-Based Assessment Project does not involve high-stakes testing; thus its influence on teachers depended on its implementation.

2.4 The Professional Development of Mathematics Teachers

2.4.1 Teacher Change

The process of change is a gradual one. In order for change to take place, the people involved need to be aware of the possibility that things could be different. Their underlying beliefs need to change. (See, for example, *Learning and Testing Mathematics in Context*, a description of a statistics teaching experiment in Wisconsin, USA, by the staff of the Freudenthal Institute [Lange et al., 1993].)

Often we have had a view that educational problems can be fixed by changing one aspect, such as the curriculum, preparation of teachers, or assessment. ... Moving to classrooms that encourage understanding requires more than fiddling with one aspect, or adding more on to what is being done. It requires more substantive, long-term changes. It also requires a change in attitude and beliefs as well as in practice and expectations. (Hiebert et al., 1997, p. xiv)

The process of change is quite complex and often very threatening. Successful professional development which generates long-term change must not only address the cognitive needs of teachers, but must also deal with their feelings about themselves and about their teaching. Working with elementary school teachers in the subject of mathematics is particularly complex because of the teachers' own lack of knowledge and even personal failure in this subject.

As Fullan (1999) wrote, innovation is a complex process and success is dependent upon local conditions. Fullan, (with Stiegelbauer, 1991), writing about teacher change in the context of school reform, discussed change as a linear process, but later he wrote about the complexity of the change process (Fullan, 1999). In addition, he discussed the reasons for the failure of previous reforms:

First, ... I maintain that these failures have occurred because the theories of change underpinning them are simplistic or absent altogether. ...

Second, we don't fully know what strategies based on complexity and chaos will look like in education reform, partly because we are at the early stage of these new discoveries and largely because educational systems operate too much like political bureaucracies. They seemingly dwell in chaos, but they do not purposefully live and learn on the creative edge of chaos. They have not yet proven capable of balancing 'too much and too little structure' on the way to continuous learning.

Third, while we don't know what these strategies would 'fully' entail, we are now seeing more sophisticated examples in operation in public school systems. And, there is early evidence of their success. (Fullan, 1999, pp. 29-30)

Fullan's comments are especially pertinent to my research, since his study—like mine—is concerned with reform on a large scale. My work is primarily concerned with the change in assessment practices that individual teachers undergo, rather than with systematic change at the level of the school or at a more global level. The latter is an issue for future research that will be addressed briefly in the final section.

It is widely accepted that knowledge and understanding of students' mathematical learning is important for teaching. (Ball, 1997; Carpenter, Fennema, Franke, Levi & Empson, 1999; Carpenter, Fennema, Peterson, Chiang & Loef, 1989; Even & Tirosh, 2002; Hiebert et al., 1997). One major main idea behind my professional development courses was teaching teachers to analyze students' work on the investigative assessment tasks. This was my way of deepening the "teachers' appreciation for what is worthwhile in mathematics and to do so in ways that can take into account the full context of student understanding and experience." (Driscoll, 1999, p. 79)

As Solomon and Morocco (1999) wrote:

Teachers who are committed to diagnosing individual students' skills and understandings are *using* that diagnosis to inform their teaching practice are likely to have a fundamentally different view of teaching and learning. (p. 235)

I planned professional development courses that would be appropriate for elementary school teachers in Israel and that would be relevant for the Israeli school system. Nevertheless, these courses were based on the most recent theories of teacher development. Thus, the findings are pertinent not only to Israel, but may be adapted to educational systems elsewhere.

2.4.2 Professional Development

Nowadays there are many resources discussing effective professional development (see, for example, Friel & Bright, 2001; Loucks-Horskey, Hewson, Love & Stiles, 1998; Nesbit, DiBiase, Miller & Wallace, 2001; Solomon, 1999). The common

principles listed by these resources can be divided into three broad categories: 1) knowledge of content and pedagogy; 2) the manner in which staff development is carried out; and 3) leadership skill development.

In order to teach reform mathematics, teachers need to feel comfortable with the subject and to be able to interact with it in new ways (Carpenter & Lehrer, 1999; Hajar, 1997; Heaton, 2000; NCTM, 1991; Schifter, 1993). Research on learning shows that most students cannot learn mathematics effectively by listening and imitating alone; nonetheless, this is the way most teachers teach mathematics. Most teachers teach as they were taught, and not as they were taught to teach (MSEB, 1989, p. 65). Elementary school teachers need to learn more mathematics and to develop the right sort of attitude towards mathematics (Ball, 1991; Ball, Lubienski & Mewborn, 2001; Graham & Ferrucci, 1998; MSEB, 2001; Sowder, Philipp, Armstrong & Schappelle, 1998). As stated in the *Partners in Change Handbook* (1997, p. 6), “Many mathematics teachers do not consider themselves to be mathematical problem solvers. They teach the required mathematics to their students but they do not find mathematics pleasurable or interesting; they certainly do not engage in thinking or playing with mathematical ideas!” In addition, as in many other countries, Israeli elementary school teachers tend to lack basic mathematical knowledge (Graeber & Tirosh, 1988; Hershkowitz, 1989; Linchevski, 1985; Ma, 1999; Simon, 1993). For these reasons, teaching mathematics to the teachers themselves needed to be a crucial element of my professional development courses. Teachers must experience learning mathematics in the manner that we want them to teach it (CBMS, 2001; Duckworth, 1987; Friel & Bright, 1997; Loucks-Horskey et al., 1998; Schifter, 1993; Silver, Kipatrick & Schlesinger, 1995). This means that teachers—as students—should work on mathematics through investigations. In addition, the verbalization and personal reflection that is required of students in reform mathematics also must be crucial elements in the process of learning that these teachers undergo in the professional development courses (Cobb, Boufi, McClain & Whitenack, 1997; Schoenfeld, 1994;

Schorr, 2000; Simon & Blume, 1994; Sowder et al., 1998; Thompson & Thompson, 1996).

My professional development courses also took into consideration various other factors that have been discussed in the literature as contributing factors in the success of such courses:

1. Teachers need to learn how to use assessment tasks: what is assessment, how to choose appropriate tasks for assessment and how to utilize assessment results in order to plan further teaching. In the classroom, teachers tend to look for feedback to confirm the success of their teaching. Using prepared assessment tasks, with performance levels provided, requires changes in teachers' practice. As written by Shepard (2000a, p. 10): "Our aim should be to change our cultural practices so that students and teachers look to assessment as a source of insight and help instead of an occasion for meting out rewards and punishments". (For further discussion see also ATM Working Group on Assessment, 1989; Black, 1998; Bryant & Driscoll, 1998; Bush, 2000; Bush & Leinwand, 2000; Falk, 2000; Herrington, Sparrow, Herrington & Oliver, 1997; Krulik & Rudnick, 1998; Lambdin, Kehle, & Preston, 1996; Madfes & Muench, 1999; Shafer & Romberg, 1999; Sheingold, Heller & Paulukonis, 1995; Wiggins, 1998; Wilcox & Lanier, 2000.)

In my courses these were central aims. The teachers learned to use the investigative tasks for assessment, analyzed their students' work and planned follow-up lessons.

2. Teachers must have classroom trials between workshop meetings. After learning something, teachers try it out in their classrooms, then return to the workshop and discuss their experiences (Clarke, 1992a; Dole, Nisbet, Warren & Cooper, 1999; Driscoll & Bryant, 1998; Ginsburg, Jacobs & Lopez, 1993b; Silver & Kenney, 1995; Silver & Kilpatrick, 1988).

All my courses were based on monthly meetings, which enabled the teachers to use the tasks in their classes and then report their findings at the following workshop.

3. I have always found that personal experience is a crucial element in success as a leader. Thus, I believe that leading teachers must experience new ways of teaching students before they try to lead other teachers in similar ways. Lampert (1991) used this method herself in her work as an elementary school teacher, where one of her goals was to be able to better teach future teachers.

This emphasis on direct experience was basic in all my courses. Teachers first experienced reform mathematics as learners, by investigating mathematics, writing reports and reflecting on their experiences. Then they were asked to use the investigation tasks with student—as teachers. Only after these experiences were they requested to lead other teachers.

4. Teachers need to be provided with good materials. They are sometimes asked to adapt materials, but not to create original ones (Ball & Cohen, 1996; Clarke, 1997; Cobb, Yackel & Wood, 1991; Hofstein & Even, 2001; Sullivan & Lilburn, 1997; Watson & Mason, 1998).

In my study teachers were provided with mathematical investigation tasks for assessment. They were not asked to create tasks.

5. Like for students, teachers need to learn by constructivism both in regard to mathematics and in their construction of pedagogical knowledge (Cobb et al., 1991; Hand, 1996; Henningsen & Stein, 1997; Schorr, 2000; Simon, 1995). The difference is that experienced teachers have many questions and are searching for answers, whereas students need to be motivated. Davis (1992) used the metaphor of “looking for pieces of a puzzle” as a way of determining when understanding occurs:

“Understanding” occurs when a new idea can be fitted into a larger framework of previously-assembled ideas. This is particularly important when the new piece is recognized as the “answer” to a question that had already been of interest, perhaps of burning interest. It is then a piece of the puzzle that one had really been seeking, and its inclusion helps to make sense of the entire picture. (p. 229)

In my courses I worked with teachers using a constructivist approach (see Section 1.1). This meant carefully listening to them and aiding them in their learning, exactly as we do with our pupils.

6. Teachers' learning needs to be conducted in an environment in which there are both social norms and socio-mathematical norms (Cobb et al., 1991; Yackel & Cobb, 1996). As an example of a social norm, the derogatory comment "that's easy" is not allowed in my classroom—I do not permit students or teachers to laugh at mistakes made by others. For true mathematical learning to take place, there must also be socio-mathematical norms.

socio-mathematical norms are intrinsic aspects of the classroom's mathematical microculture. ... they cut across areas of mathematical content by dealing with mathematical qualities of solutions, such as their similarities and differences, sophistication, and efficiency. Additionally, they encompass ways of judging what counts as an acceptable mathematical explanation (Yackel & Cobb, 1996, p. 474).

7. Another element crucial to effective professional development is the importance of relationships.

Any educational reform strategy that improves relationships has a chance of succeeding; any strategy that does not is doomed to fail. This is why formal policies and procedures themselves will never provide the answer. (Hargreaves & Fullan, 1998, p. 90)

This relationship building takes place on two levels: both the interpersonal relationships that the leader develops with the teachers and the professional modeling that he does. An effective leader must demonstrate by personal example all the attributes he would like his teachers to have and the actions he would wish them to undertake. In my work with teachers I tried always to act as an example.

In order to successfully implement these aspects, a course needs to be long-term in range and to include many workshops. In addition, many elements of the course need to be planned in an *ad hoc* fashion from workshop to workshop and not entirely in

advance. This is necessary in order to meet teachers at the point “where they are at” (according to the precepts of constructivism) and to build upon their actual knowledge in mathematics and in pedagogy. Testing the long-term effect of these elements is another aspect of my research. It is hoped that this will provide a contribution to the expanding knowledge of the mathematics community regarding the essential ingredients necessary for good professional development.

CHAPTER 3: RESEARCH METHODS

This chapter describes the purpose of my study, the research questions that were posed and the initial assertions that motivated these questions. This is followed by details of the investigative assessment tasks: their characteristics, the criteria used for evaluating the results and an explanation of how the tasks were developed. Data was collected from two different sources: on the one hand, from workshops given to select teachers throughout the country in the framework of national courses on assessment, and on the other hand, from an intensive study carried out in a small community known as Sharon, in which all teachers attended the workshops. Details of the data collected in each of these sources and of the progression of each of these studies are detailed in Chapters 4 and 5 respectively.

3.1 The Research Questions and Working Assertions

The reader will recall that the purpose of this study was to determine the relationship between good assessment practices and good teaching methods. I wanted to investigate whether making good assessment tasks available to elementary school mathematics teachers would have a positive effect on their teaching. To that end, the following research questions were posed:

Will the teachers' use of mathematical investigation tasks for assessment purposes influence their view of mathematics?

Will the teachers' use of mathematical investigation tasks for assessment purposes influence the way they teach and if so, in what ways?

Will the teachers' use of mathematical investigation tasks for assessment purposes influence the way they assess their students and if so, in what ways?

My plan was to observe teachers who were given the investigative tasks, and determine the effect this would have on their teaching and assessment practices. My

experience as a mathematics teacher and as a teacher trainer led me to develop several working observations. These observations led to the formulation of my research questions.

My experience as a mathematics advisor led me to believe that teachers taught routine procedures most of the time because they believed that this is what is involved in elementary mathematics teaching. Most elementary-school teachers in Israel have studied only one year of mathematics beyond high school; moreover, the level of mathematics to which they were exposed in high school was often low. Their concept of teaching mathematics is based on their own experiences at school, which for the most part involved learning by means of routine procedures. At no time were they exposed to problem solving, doing exploratory work in an effort to come up with underlying rules, verbalization, justification, or any other basic concepts of reform mathematics. This lack of exposure to such concepts makes it likely that this is not part of their teaching repertoire. Thus my first observation, based on my knowledge of the teachers' backgrounds, was that teachers invest most of their efforts in teaching routine procedures.

Since it appears that teachers spend most of their time teaching routine procedures, they prepare worksheets as practice drills that involve solving single-answer exercises. Most of the exercises are computations: e.g., "Multiply 345 by 8", "Add 5487 to 6902". These are usually written without any words, except for the heading "Compute" and simple single-step word problems. Thus, my second observation was that the worksheets that teachers prepare for the most part consist of single-answer exercises.

The tests that teachers compose assess achievement on the procedures they have taught. That is, test questions tend to be similar to the worksheets that teachers prepare. As previously mentioned, most teachers are not familiar with investigation tasks as a mathematical activity. Since they do not generally use mathematical investigation tasks in their classrooms, it is reasonable to assume that such items do

not appear on their tests. Again as previously mentioned, if the worksheets prepared by teachers involve, for the most part, single-answer exercises, then these are the same sort of questions that are included in their tests. As suggested by the literature (see Section 2.3), teachers teach and test in similar fashions. Thus my third observation was that most mathematics test questions require a single answer only. These are generally questions of knowledge and do not require higher-order thinking.

3.2 Investigative Assessment Tasks

The Israeli syllabus for the elementary school, written in 1987 on the basis of a previous syllabus written in 1971, is still quite traditional in nature. The focus is on topics to be covered and skills to be mastered. Although the syllabus contains frequent references to the need for understanding, “understanding” has traditionally been taken to refer to knowing the meaning of algorithms. Whenever “problems” are mentioned, the reference is to traditional word problems.

It is against this background that the School-Based Assessment Project has created a “bank” of investigative tasks which exhibit ideas of reform mathematics for elementary-school students (see Appendix A for examples). These tasks were intended to be used by teachers for assessment purposes and to serve as examples of the type of teaching tasks desired.

These mathematical investigation tasks are characterized by the following eight features:

1. Encountering authentic problem situations

One of the guiding principles motivating these mathematical investigation tasks is that they should be based on the students’ areas of interest and should relate to meaningful concepts; they should raise the students’ intellectual curiosity and their willingness to invest efforts in solving the problems. The task should be posed at an appropriate level: it should be sufficiently challenging so as to ensure satisfaction upon solving,

but should also contain some lower-level questions in order to prevent the frustration that would be generated by complete failure.

2. Encountering unfamiliar problem situations

Unlike the exercises designed to master a skill or a specific concept, a problem-solving activity should be based on a situation which is relatively new to the student. A new problem situation is also needed to assess an in-depth understanding of a topic or a relationship among several concepts.

3. Integrating domains

Interesting mathematical tasks frequently require the integration of several topics from within the field of mathematics or from outside its domain. The integration of domains is also important as a means to assess real understanding and to provide meaningful models for mathematical concepts.

4. Monitoring processes

Open questions allow students to choose their own ways of solving the problem; thus, teachers can monitor thinking processes, cognitive styles and the ability to apply learned concepts, in addition to the students' knowledge of formal mathematics and their ability to produce an answer.

5. Emphasizing thought processes

Tasks which require "reversed thinking"—i.e., the reversal of roles between the given and the variable in a frequently used algorithmic task— are considered to involve a higher level of thinking. It is believed (see for example, NCTM 2000) that activities involving higher-order thinking should replace a large part of the short algorithmic questions, both in teaching and in assessment.

6. Finding patterns

The discovery and justification of numeric or geometric patterns and principles lie at the core of any mathematical activity; this should therefore be an integral part not only of classroom activities, but of assessment tasks as well.

7. Ranking the level of difficulty

The sequence of questions presented should be ordered in a gradually increasing level of difficulty. This scaffolding of a complex activity allows most students to “get started” and to progress to a stage that corresponds to his level of understanding and ability.

8. Using manipulatives

Providing concrete materials in assessment situations is vital for two reasons. First, these situations are encountered by the student for the first time and the manipulatives enable him to cope with them. Second, assessment tasks should parallel the learning tasks.

3.3 Criteria for Scoring

In the opinion of many researchers and reports (for example, Clarke, 1992b; Greenwood, 1993; Lambdin Kroll et al., 1992; MSEB, 1993a; NCTM, 1989), the traditional evaluation method of attaching a number grade to students' performance on complex mathematical tasks cannot reflect the wide variety of aspects of their work. These sources recommend that student assessment be based on detailed profiles of the whole problem solving process, rather than relating solely to its final results. These profiles are based on a set of well-defined criteria of assessment, with a scale of from three (e.g., low / medium / high) to at most seven levels of performance on each criterion (Cai, Lane & Jakabsin, 1996; Schwartz et al., 1995). Thus, by analyzing their students' criteria-based profiles, teachers can obtain a rich diagnosis of mathematical abilities and dispositions, and determine what sort of intervention is needed.

The criteria for scoring used in the Assessment Task Bank was developed in several stages. In the Second National Course for Assessment (Section 4.3), we (Alex Friedlander, my colleague in the School-Based Assessment Project, and myself) established three criteria:

- Understanding the task
- Thought processes
- Computations.

By the Third National Course for Assessment (Section 4.4), after much work with teacher-leaders, and more research reading, we (Friedlander and myself) established six criteria:

- Understanding the task
- Understanding the concepts
- Mathematical thinking
- Calculations
- Communication
- Variety and creativity.

More detail about how these criteria were used with the teachers is provided in the descriptions of the national courses for assessment which follow (Sections 4.3 and 4.4).

3.4 The Development of the Investigative Assessment Tasks

The development of these tasks underwent several stages. The first draft of the tasks was developed by the project staff members on the basis of their pedagogical knowledge and on research findings (for example, Brown & Walter, 1990; Lange, 1992; Marshall, 1988; MSEB, 1993a; Walter, 1993). All tasks were then critiqued by

a committee of curriculum developers, mathematicians and supervisors (peer revision). In addition, they were revised by a language editor. Each revised task was then piloted with approximately 300 children from a variety of socio-economic levels and schools. These were students who had not necessarily encountered similar mathematical investigation tasks prior to the pilot study. The teachers also filled in a questionnaire (Appendix B – Questionnaire used for pilot of tasks) about class atmosphere, their opinion on the tasks and the tasks' suitability as tools for assessment. The students' answers to the tasks were graded according to answer keys and the data were analyzed.

The tasks were then revised in light of the results of the pilot study. Questions which had not been understood by most students and questions which seemed to be unnecessarily difficult were either corrected or omitted. The project staff often read and re-read student answers, in an effort to find patterns and to define different levels of student thinking and understanding of concepts.

The final versions of the tasks and the corresponding answer keys were again critiqued by mathematicians, curriculum developers, a language editor and an expert in evaluation and assessment.

3.5 The Research Paradigm

This research project was based on a "naturalistic inquiry" paradigm (Lincoln & Guba, 1985; Moschkovich & Brenner, 2000).

A paradigm may be viewed as a set of *basic beliefs* that deals with ultimates or first principles. It represents a *world view* that defines, for its holder, the nature of the "world", the individual's place in it, and the range of possible relationships to that world and its parts (Guba & Lincoln, 1994, p. 107).

Thus, in my study "knowledge" was created in interaction between myself, as researcher, and the respondents. My research was based not on statistical analyses of

teachers' responses to questions or other "objective" data, but rather, on a holistic approach to the teachers' reactions to the assessment workshops and the tasks made available to them through my open-form questionnaires and semi- or unstructured interviews.

The different data collection techniques allowed data collected according to one method to be used to validate the accuracy of data gathered according to another method (Goetz & LeCompte, 1984). These techniques allowed for triangulation, an essential process which safeguards the investigator against accepting the validity of initial impressions without sufficient evidence (Glaser & Strauss, 1967; Lincoln & Guba, 1985).

Guba and Lincoln (1989) established a set of standards necessary to ensure the internal validity of the data collected in such an inquiry. Among the techniques suggested are prolonged engagement and persistent observation. These, indeed, were techniques employed in my study, which ranged over a number of years and in which I was involved as an active participant.

In addition, the most crucial technique for establishing credibility, according to Guba and Lincoln (1989, pp. 238-239), is that of member checks: "the process of testing hypotheses, data, preliminary categories and interpretations with members of the stake-holding groups from whom the original constructions were collected." In other words, it was not sufficient to record and interpret my own analysis; the process was complete only when my findings were subsequently confirmed by the participating teachers themselves.

My research corresponded to Brown's "iterative conceptualization" (1998, p. 268), a process which consists of making a preliminary analysis of initial data and using that as a basis for further rounds of data collection and analysis, modifying one's methods all the while in accordance to the findings.

Goetz & LeCompte (1984) explain that an “ethnographic product” is evaluated by the extent to which it recapitulates the cultural scene studied, thus enabling readers to envision the scene as it was witnessed by the researcher. My research recounts the progression of the national courses on assessment and examines how these courses affected the concept of mathematics held by the participating teacher-leaders and the assessment practices they practice. Similarly, I recount the progression of the study in the Sharon community and its effect on the mathematics coordinators. In each case, I have detailed my experiences in order to provide the reader with a view of the entire picture.

In the course of the five years of my research, I used many strategies of data collection: unstructured participant-observations, interviews, questionnaires and content analysis of artifacts (tests and tasks written by teachers). All interviews were taped and subsequently transcribed. The teachers’ responses to each question on every questionnaire were translated from Hebrew into English, analyzed and classified into categories. Conclusions were then drawn from these classifications. The interview transcripts were similarly analyzed.

LeCompte (2000), writing about the analysis of qualitative data, compared the process to that of taking a puzzle apart and reassembling it. This was the process that I employed: reading and re-reading answers, determining categories, drawing conclusions, and then—upon seeing that the pieces did not “fit”—starting the process from the beginning, time and again, until all the pieces fell together. This process was also described by Glaser and Strauss (1967, p. 6): “generating a theory from data means that most hypotheses and concepts not only come from the data, but are systematically worked out in relation to the data during the course of the research”.

Maykut and Morehouse (1994) wrote about the importance of verbalization in this process. They explained that the qualitative researchers need to attempt to understand the situation as it is constructed by the participants. They need to attempt to capture what people say and do, that is, the products of how people interpret the world. This

task requires of the qualitative researchers the ability to reproduce in their own minds the feeling, motives and thoughts which underlie the actions of others.

Words are the way that most people come to understand their situations. We create our world with words. We explain ourselves with words. We defend and hide ourselves with words. The task of the qualitative researcher is to find patterns within those words (and actions) and to present those patterns for others to inspect while at the same time staying as close to the construction of the world as the participants originally experienced. (Maykut and Morehouse, 1994, p. 18)

3.6 The Background for the Collection of the Data

My research was divided into two parts: 1) a national study of the country; and 2) an intensive study of a small Israeli community, Sharon, in which all six elementary schools adopted the use of the School-Based Assessment Project for mathematics. I began by examining what was happening in the teaching and assessment of mathematics throughout the country, through the national courses on assessment. The first nationwide workshops began in the spring of 1995. However, I discovered that in order to determine how teachers were reacting to the use of reform mathematics and how this affected their teaching and assessment methods, it would be necessary to undertake a more intensive study. Thus, I embarked upon the intensive study of Sharon, a community in which I was already working and where teachers had already been exposed to the concepts and ideas of reform mathematics. School-based assessment in mathematics was introduced to the community of Sharon in December, 1996—a year and a half after the first national course on assessment.

The national study included three nationwide courses on assessment catering to teacher-leaders in elementary-school mathematics. In each course, the teachers were introduced to the mathematical investigation tasks and received training in their use for assessment purposes. This study also included a program called *Ofek*, a satellite

interactive on-line course which was run with the help of leading teachers who had graduated from the earlier national courses on assessment.

The description of the national study, Chapter 4, is followed by a description of my intensive study of the small community of Sharon, Chapter 5.

3.7 Teacher Quotes: Coding Used

The teachers in my study have been coded consecutively (teacher AA to teacher AZ, teacher BA to teacher BZ, CA to CZ, DA to DZ, etc.). The double lettering was necessary because there were more than 26 teachers involved. The coding of the teachers in the national study commences with an A, B or C. The coding of the teachers in the Sharon study commences with either a D or E.

The questionnaires I developed appear in Appendix B. Exemplars of the teachers' responses on the questionnaires appear in Appendix C. (All questionnaires were originally written in Hebrew and were translated by the researcher.)

Teachers' quotes are denoted as follows:

Teachers' code (see explanation above)

Course number (Roman numeral):

- I First National Course on Assessment;
- II Second National Course on Assessment;
- III Third National Course on Assessment;
- IV Sharon study.

Questionnaire number (Arabic numeral):

- 1 – first or only questionnaire
- 2 – second questionnaire

Page number – refers to Appendix C where the exemplars of teachers' responses are exhibited.

Teacher transcripts are coded as follows:

teacher CD.III.2; p.40.

Refers to teacher CD, who attended the Third National Course on Assessment. The quote is taken from her response to the second questionnaire. The full text of her response to the questionnaire appears on page 40 of Appendix C.

CHAPTER 4: DATA COLLECTION IN THE NATIONAL STUDY

4.1 The Method of Data Collection

The national study consisted of three courses on assessment for teacher-leaders in elementary-school mathematics. This was followed by additional work with some course graduates which included *Ofek*, interactive satellite broadcasts and advanced assessment workshops. In order to obtain data to help address my research questions, I used open-form questionnaires and unstructured interviews, as well as collecting sample tasks and tests.

The following table provides a summary of these activities.

4.1.1 Summary of Data Collected

	Activity	Date	Sample	Data collected
First National Assessment Course	4 workshops	Spring, 1995	29 teacher-leaders	1 questionnaire – 24 teachers responded
Second National Assessment Course	8 workshops	School year 1995–96	25 teacher-leaders	1 questionnaire - 23 teachers responded
Third National Assessment Course	8 workshops	June, 1996 until January, 1997	32 teacher-leaders	2 questionnaires- 27 teachers responded to the first, and 25 teachers to the second
<i>Ofek</i>	3 workshops/ meetings	January, 1997 October, 1997 January, 1998	20 graduates of the national assessment courses	Semi-structured group discussions
Additional in-service workshops	5 workshops	July, 1998 until March, 1999	30 graduates of the national assessment courses and	1 questionnaire - 12 teachers responded

			other leading teachers involved in assessment	
Interviews		3 in June, 1998 1 in January, 2001	4 graduates of the national assessment courses	4 unstructured interviews

4.1.2 Construction of the Questionnaires

All the questionnaires were of the open-form type, which allowed respondents to raise all the relevant issues and enabled me to learn how the teachers related not only to the explicit questions, but also to implicit issues.

The questions on the questionnaires were formulated in response to the needs of the particular situation. The initial questions dealt with the general use of the investigation assessment tasks; these were followed by questions about the use of the criteria; the questionnaires concluded with questions about the entire implementation process. The questions did not explicitly refer to my research questions; instead, the teachers' answers were analyzed and categorized, the analysis ultimately providing data relating to the three research questions.

The first brief questionnaire is presented below, followed by the rationale underlying its questions.

First Questionnaire – First National Assessment Course

1. In what type of school did you use the investigative tasks?
2. What were the results of your use of the tasks with the students?
3. Do you think these tasks can be used for assessment?

This was the first assessment course I had held and the first time I had presented the investigative tasks to teachers. Consequently, I was interested in obtaining as much

information as possible about their experiences in using them. The information about the types of schools in which they were used and how the students had reacted to them was invaluable. In addition, by their responses to the question whether they believed the tasks could be used for assessment purposes, I could gather information about their teaching.

The second questionnaire is presented below:

Second Questionnaire – Second National Assessment Course
Questionnaire for Math Leaders

Lately we have added criteria for evaluating the tasks, in addition to the answer guide which is attached to each task.

We would like to ask a number of questions which relate to the use of these criteria.

1. How many different tasks have you checked using these criteria? Which tasks?
2. What do you think, in general, about the use of criteria for checking tasks?
3. Did the use of criteria cause:

	very much	some	no difference	definitely not
extending the checking process				
deepening understanding of the pupil's performance				
emphasizing the difference between students				
acquiring more detailed information about the student's level				

4. Which criteria are not clear? That is, about which is it hard to make decisions?
5. Did checking by criteria help you plan your teaching? If so, provide an example.

These questions explicitly refer to the use of criteria in the evaluation of students' performance on investigative assessment tasks. Using criteria for evaluation purposes was a new experience for the teachers; thus, their reactions were important to me.

Again, the construction of open questions, rather than multiple-choice questions, enabled me to gather data that I could analyze and relate to my research questions. The specific issues raised in question three were chosen after many informal discussions with the teachers—these were the issues that continually arose. The third questionnaire, like the first, was brief:

Third Questionnaire – Third National Assessment Course

“Add-a-digit” – an estimation task

1. How did this task differ from the tasks that the students usually perform?
2. In your experience, does checking tasks by means of criteria provide you with information that conventional grading does not provide? yes / no
If so, describe this information.
3. Additional comments.

This questionnaire gave me valuable information regarding the teacher-leaders’ experiences in using an investigative task for assessment and regarding their feelings about the use of criteria for evaluation purposes. The first question provided data about the type of tasks that teachers were generally using, as opposed to the investigative ones that I provided. As with the previous questionnaire, the questions were formulated as open ones, in order to obtain rich responses which could then be considered in relation to my research questions.

Fourth Questionnaire – Third National Assessment Course

The fourth questionnaire was identical to the second question, but with the addition of question six:

What has changed in your class / in your supervision as a result of this course? Explain, providing examples (for example, different mathematical content, difference in teaching styles, difference in ways of supervising).

My aim was to examine how the teacher-leaders of the Third National Assessment Course were reacting to the use of criteria for evaluation purposes and in addition, to determine whether their responses were similar to those of the teacher-leaders of the

Second National Assessment Course, one year earlier. Since this questionnaire was administered in the final workshop, I added question six, which required the teachers to reflect upon what they had learned. The open form of the question provided me with rich results, which were relevant to all three research questions.

The fifth questionnaire was again brief:

Fifth Questionnaire – Additional In-Service Workshops

Questionnaire - 2.7.98

1. Record three important things that you learned in these two days.
2. Would you like to continue to meet during this year?
If so, which subjects would you be interested in?
3. Comments on the process of implementing the assessment task bank.

This questionnaire was administered to graduates of the teacher-leaders from the National Assessment Courses who were participating in additional workshops. The open questions provided me with valuable information regarding the issues the teacher-leaders were focusing on, the sort of language they were using, and the type of mathematics to which they were referring.

4.1.3 Interviews with Teachers

The unstructured nature of the four interviews allowed me to question the teacher-leaders closely about their use of the investigative assessment tasks, following each teacher in the direction she led. The four teachers were chosen as representative of various sectors of the Israeli school system. All interviews were taped and subsequently transcribed, thus facilitating my analysis.

4.1.4 Sample Tasks and Tests

I collected sample tasks and tests used by the teachers and teacher-leaders involved. These tests/tasks were analyzed by means of the criteria I had developed. Avital and

Shettleworth (1968) provide examples of mathematical questions on different levels of thinking, applying Bloom's taxonomy to mathematics. There are many more recent publications which analyze questions (e.g., Henningsen & Stein, 1997; Reys & Long, 1995; Stecher & Mitchell, 1995; Stein et al., 1996; Stein, Smith, Henningsen & Silver, 2000; Sullivan & Lilburn, 1997). My analyses are based on these ideas.

4.2 The First National Course on Assessment

During the spring of 1995, the first national course on alternative assessment was held. This course was geared towards leading-teachers of mathematics in the elementary schools from throughout the country, and a total of 29 teachers registered. The course included four day-long workshops, and between meetings the teachers tried out the new investigative tasks in their classrooms. These teachers were the first to use these tasks; thus, their reactions were of great importance to the study.

4.2.1 The First Workshop

The first meeting began with the teacher-leaders solving two investigative tasks: *Table Arrangements* and *Geoboard Quadrilaterals* (Appendix A). This was followed by a discussion of the tasks' attributes and the mathematics involved in the tasks. Professor Arye Levin (Tel Aviv University), the academic advisor for the School-Based Assessment Project, presented the philosophy underlying the project, helping the teachers to understand the connection between their work as mathematics leaders and alternative assessment in Israeli elementary schools. This was followed by a discussion of assessment in general and of alternative assessment in particular. The teachers were then required to solve two additional tasks, *Trains* and *Population* (Appendix A).

Dr. Alex Friedlander, my colleague in writing the investigative tasks, presented the underlying philosophy, discussed how the tasks are all based on situations—some

taken from everyday life and others from mathematics—and dealt with the relationship between the mathematics involved in the tasks and the national syllabus.

The final session of the workshop was devoted to further aspects of assessment: how one grades a student's work on a task and what a teacher is to do with the results of such assessments. Before the teachers left we decided which tasks they would try out in their classes before the next workshop, and they filled out a form providing personal details and information regarding their current professional obligations.

4.2.2 The Second Workshop

The second workshop, which took place three weeks later, focused on the teachers' impressions based on their experiences with their students. In this workshop the teachers were required to respond in writing to a questionnaire (Appendix B) that dealt with the use of the task for assessment purposes.

There were 24 teachers present at that meeting; five of the original teachers were absent. After they had completed the questionnaire, we held a group discussion. Professor Pearla Nesher, then Chief Scientist of the Ministry of Education in Israel and the person who had initiated the School-Based Assessment Project, participated in the discussion as she was very curious to hear how teachers and students were relating to the tasks. Both the questionnaire and the group discussion focused on the following questions: "What questions were hard for the students?" "What types of answers the students gave?" "Could the students verbalize their explanations?" "Could the students formulate the necessary generalizations?"

I continued to associate these investigative tasks with concepts of assessment: what information they were providing regarding the students' thought processes and their abilities to perform mathematical activities, such as generalization and justification. These concepts were difficult for some of the teachers who were accustomed to testing knowledge and skills alone.

I talked about the implications of Vygotskii's theories on teaching and learning—theories that were hitherto unfamiliar to some of the teachers.

The teachers then solved several additional tasks. At this stage, we (the project staff) were interested in their responses to different types of tasks and thus introduced them to a large variety. In the course of the four workshops they solved a total of 15 tasks of various lengths. (Nowadays I tend to use fewer tasks in workshops and to spend more time on each one: discussing the mathematics involved in more depth and working with the teachers on follow-up classes for those students who experience difficulties.)

The session concluded with a discussion of how they should work on the tasks with other teachers in their schools. My philosophy for teacher-educators is that they first need to experience new concepts as teachers; thus I had requested them to use the tasks in their own classes, or to be present in the classroom when the students were working on the tasks. I believe that one cannot guide others in something not personally experienced. For this reason I was encouraging them to be “just teachers” at this stage, rather than teacher-leaders. However, they were excited about these new types of tasks and wanted to share them with others.

4.2.3 The Third Workshop

At the third workshop, three weeks later, Professor Pearla Nesher presented the concepts underlying the School-Based Assessment Project, which included new directions for assessment. Professor Nesher stressed that those who really need assessment results are the classroom teachers, rather than the Ministry of Education; hence, school-based assessment is of far greater importance than national testing (see Chapter 1, Section 1.4, for the Israeli system of national testing). She talked about the connection between assessment and teaching. Since Professor Nesher is also the academic advisor for the mathematics textbooks used by almost all the elementary-

school pupils in Israel, her presence was important to all the teachers. She was able to help the teachers understand how these investigative tasks assess important mathematical activities, even though many of the teachers still felt that mathematics should be skill-based.

Professor Arye Levin also presented his ideas on the connection between learning and teaching, and stressed that there is no teaching without testing. A good teacher must therefore continually assess his or her students. He discussed the fact that good teachers have always used types of alternative assessments to collect information on their students' knowledge and thought processes, and noted that our current efforts are geared towards making the practice of alternative assessment available to **all** teachers.

As in previous workshops, the teachers solved additional investigative tasks and discussed them.

4.2.4 The Fourth Workshop

The fourth workshop, which was held two months after the third one, focused on the use of the investigative tasks as assessment tools. I continued to discuss with the course participants the students' reactions and the response of other teachers. As in the previous workshops, a recurrent theme was "What mathematics should be taught in elementary schools?" These teacher-leaders had been accustomed to putting most of their energies into perfecting their students' basic skills. They found it hard to accept that mathematics, even on the elementary-school level, is not merely about the acquisition of basic skills, but that it involves generalizations and justifications as well. For many teachers, the idea that students should have to write explanations was a foreign one. Many were still claiming that "writing is **not** mathematics!"

Change is a slow process (see literature review, Chapter 2), and some of these teacher-leaders were feeling challenged by this new concept of mathematics. For this reason, I and the teachers devoted much time in every workshop to discussing these issues.

4.3 The Second National Course on Assessment

The second course for in-school assessment based on mathematical investigative tasks was given during the 1995–96 school year. Like the previous course, this was a national course open to leading teachers recommended by supervisors from throughout the country. A total of 25 teachers registered for the course, which consisted of nine monthly meetings, each one seven hours long. (The first course had been too short. Its purposes had been to expose the community to the new tool. This second course was to enable teacher-leaders to implement the tool successfully. Some teacher-leaders from the first course participated also in the second one.)

The purpose of this course was to acquaint its participants with principles of alternative assessment, with the use of investigative tasks for assessment purposes and with methods of reform-based teaching. Each meeting was held in a workshop format and was planned in a flexible manner in order to be able to address teachers' concerns and problems stemming from their work as leading-teachers.

4.3.1 The First Workshop

The first meeting began with a discussion about assessment, based on David Nevo's article, "Assessment as a Tool for Improving Teaching and Learning" (1991), which is concerned with assessment in general, including the evaluation of schools, curricula, teaching methods and the purpose of these evaluations. The discussion included various tools for the assessment of student learning—the focus of our course. The debate about assessment led into a broader discussion of reform-based mathematics, and the ensuing changes in emphasis of the type of assignments given to students (no

longer just drill-and-practice). Thus I began the work of establishing the connection between instruction, learning and assessment—a continuous theme throughout the course.

The teachers analyzed students' work on the *Table Arrangements* task (Appendix A) and created rubrics for evaluation. We (Friedlander and myself) were using David Clarke's (1995) ideas, and asked the teacher-leaders to come up with their own criteria based on their analysis of student work. The criteria we established at this point were:

- Understanding the task
- Mathematical reasoning
- Calculation skills.

Three possible levels of performance were established for each criterion, as delineated below:

<i>Criteria for assessing students' work on Table Arrangements</i>	
Criterion	Level of performance
Understanding the task	<ol style="list-style-type: none"> 1. Does not understand the problem. 2. Forgets some of the constraints. 3. Understands the problem.
Mathematical reasoning	<ol style="list-style-type: none"> 1. Unsystematic work and no clear line of reasoning. 2. Clear, but incorrect line of reasoning. 3. Clear and correct line of reasoning.
Calculation skills	<ol style="list-style-type: none"> 1. Most calculations incorrect. 2. Some of the calculations incorrect. 3. Most calculations correct.

After lunch, I discussed the overall plan for the course and the teachers' responsibilities. They were assigned two tasks—*Table Arrangements* and *Table 2*—to use in their classes as tests, to check the students' work using the rubrics and to write their conclusions.

Before leaving, the teachers filled out a form providing personal details and information regarding their current professional obligations.

4.3.2 The Second Workshop

The second meeting began with a discussion of what it means for a task to be authentic. This is an issue that arose many times during the course and continues to concern teachers. The dilemma is between the concept of authenticity as applied to the student's environment versus authenticity as applied to the subject of mathematics. Elementary-school teachers, whose knowledge of mathematics is often quite limited and whose past experiences were usually based on learning procedures, have difficulty grasping the concept of a task's "authenticity to mathematics".

I discussed with the teachers the entire process of assessment, ranging from the planning of the learning activities to the actual testing and subsequent planning of additional "treatment" for unsuccessful students. We discussed how they could change their tests from containing twenty exercises and five word problems to consisting of five exercises, two word problems and one investigative task, or even the possibility of using an investigative task alone as an assessment tool.

The teachers discussed their own students' work on the *Table Arrangements* task, debated how to divide the class students into smaller groups for further instruction (based on their performances) and what instruction to provide them. We talked about the need for some students merely to do revisions, whereas other students need some teacher-time in order to learn how to generalize and other students need more time in order to grasp the concepts. Thus, we discussed how the assessment results they had obtained by grading the task by means of rubrics, or criteria, could—and should—influence their teaching.

We discussed how to work with teachers in their school and how to build in-service courses on assessment: what the crucial elements in such courses are and how to build appropriate activities.

The teachers solved another investigation task—*Bridges* (Appendix A)—and we discussed its use in assessment. They were asked to use it in their classrooms as a “homework assignment”. The workshops concluded with a questionnaire that required the teachers to address two questions in writing, requiring them to do some reflecting:

- | |
|---|
| <ol style="list-style-type: none">1. Today I learned ...2. Next time I would like to ... |
|---|

4.3.3 The Third Workshop

The third workshop began with the teachers working on a mathematical investigation. Knowing that in order for teachers to teach in investigative ways they need to learn in such ways, I had decided to start all further workshops with the teachers experiencing what I expected the students to experience. This meant presenting the teachers with a problem situation and having them explore it mathematically in small groups. They then had to present their conclusions to the whole group and discuss them. This process usually lasted approximately an hour and a half (see Appendix D).

This was followed by work on the *Bridges* task, which had been assigned as homework. What had they learned about their students’ thinking? Were the students able to generalize? Could the students verbalize their generalizations? The teachers then observed a video showing two pairs of students working on this task, and we discussed using the investigative tasks as performance assessment tasks, querying why two children often offer more information than only one, when to intervene and when to let the students proceed even though they are making mistakes.

The third activity for that workshop was based on Sullivan and Clarke’s article, “Catering to All Abilities through ‘Good’ Questions” (1991). The teachers solved the

problem of finding five numbers whose average is 17.2. This was followed by a discussion of the solutions they had offered and of other possibilities. I presented the characteristics of “good questions” and examples from the article. Their assignment then was to come up with a “good question” which deals with a number line.

4.3.4 The Fourth, Fifth, and Sixth Workshops

During the fourth, fifth and sixth meetings similar activities took place. The teachers continued to “do mathematics”, that is, to experience for themselves what it means to work together on mathematical investigations and to participate in mathematical discourse.

They continued to use the investigative tasks from the Assessment Task Bank with their students and to analyze the students’ work. Some of the tasks were checked using rubrics (criteria) and some using conventional assessment methods. We discussed ways of working with students who were having problems and ways of teaching in general.

We performed activities associated with assessment, such as to compare and contrast the attributes of an investigation task with those of a conventional test or to discuss the advantages of using a variety of tools for assessment.

The teachers planned activities and workshops for courses they were responsible for running. Some of these were based on the activities I did with them, and some were assignments I gave them: to build an activity for a group of teachers in order to help them acquire the concepts of alternative assessment

I also had teachers try to write their own tasks, after we discussed the attributes of a good investigation task.

At the sixth workshop, before doing more work with the teachers on the use of criteria, they filled in the following questionnaire.

Questionnaire for Math Leaders

Lately we have added criteria for evaluating the tasks, in addition to the answer guide which is attached to each task.

We would like to ask a number of questions which relate to the use of these criteria.

1. How many different tasks have you checked using these criteria? Which tasks?
2. What do you think, in general, about the use of criteria for checking tasks?
3. Did the use of criteria cause:

	very much	some	no difference	definitely not
extending the checking process?				
deepening understanding of the pupil's performance?				
emphasizing the difference between students?				
acquiring more detailed information about the student's level?				

4. Which criteria are not clear? That is, which are hard to make decisions about them?
5. Did checking by criteria help you plan your teaching? If yes, give an example.

4.3.5 The Final Workshops (the Seventh, Eighth and Ninth)

Three additional workshops were scheduled to take place in the course of the year. One was cancelled, and instead the teachers attended a one-day conference of the Israeli Mathematics Teachers' Organization.

In the course of the other two workshops I continued the activities mentioned above: the teachers' investigation of mathematics, discussion of their meetings with other teachers, and analysis of student work. I continued discussing assessment by means of rubrics (criteria); they then checked one task, both with its answer key and with criteria, and we discussed what additional information we may gain from the criteria.

In addition to these activities, our final workshop included planning the courses they would need to give to other teachers the following year.

4.4 The Third National Course on Assessment

This course, held during the 1996–97 school year, benefited from the experience that I had gained from the previous courses. This is the course with the most complete data because it was the third course, and so I knew better how to organize it, and also I had gotten seriously into my doctorate, and knew I needed to collect as many artifacts as possible. Also from the fact that most of the teacher leaders who participated were experienced teachers who had been trained by the mathematics group of the Center for Technological Education, which produced the Israeli elementary-school textbooks, and were serious about their professional training.

The course consisted of eight monthly meetings, held in Tel Aviv. A total of 32 teachers attended, from throughout the country.

Like the previous courses, the purpose of this one was to acquaint the teachers with the principles of alternative assessment, with the use of investigative tasks for assessment and with methods of reform-based teaching. Each meeting was held in a workshop format and was planned in a flexible manner so as to address the teachers' concerns and the problems arising from their work as teacher leaders.

The subjects covered in the course included: familiarization with varied types of assessment; the theoretical background underlying the establishment of assessment task banks; familiarization with the assessment tasks; the analysis of task results and

formulation of criteria for determining grades; the use of assessment task results to plan further teaching in school; the planning of workshops for teachers in this subject; and writing additional assessment tasks.

4.4.1 The First Workshop

The first workshop opened with the participants discussing the question “What is mathematics for you?” in small groups (see responses, Chapter 4). This was followed by a discussion of the methods they used to assess their students’ progress. The methods mentioned were tests, quizzes and checking homework. I pointed out the discrepancy between the goals they set for their teaching and the methods they used for assessment. The participants solved two investigative tasks, *Getting Results* and *Quadrilaterals* (Appendix A), and we discussed the tasks’ attributes. They watched a video of two pairs of girls solving the *Bridges* task and discussed how to assess their performance. The teachers were asked to use the *Getting Results* task in a fourth-grade class and to bring all the students’ papers to the subsequent workshop. Finally, each teacher was required to describe his expectations for the course in writing.

4.4.2 The Second Workshop

The second workshop opened with group discussions, dealing with the following questions: “What is alternative assessment?” “What is alternative assessment in mathematics?” “Should alternative assessment replace conventional assessment or supplement it?” and “What is the importance in mathematics of correct answers, of knowledge, of creativity?” After the discussion, the participants were required to solve an investigative task. Rather than giving them a task from the School-Based Assessment Project, this task was a problem pitched at their own level, which I felt would require them to “do mathematics”. As previously mentioned, it is my firm belief that teachers cannot teach students in ways which they have not experienced as

students themselves; thus, each workshop was planned to include a mathematical experience for them. They were required to think mathematically, to discuss their ideas within the small groups and to find ways of presenting ideas to the larger group. In other words, it required of these teachers the same kind of activities that I hoped they would demand of their students: formulating generalizations, justifications and explanations. As in the previous courses, for many of the teachers these were new experiences.

This activity was followed by a discussion of the merits and weaknesses of two different methods of evaluating students' work: using conventional grading as opposed to using rubrics for evaluation. In previous courses, we (Alex Friedlander, my colleague in the School-Based Assessment Project, and myself) had asked teachers and teacher-leaders to come up with criteria on the basis of their analysis of student work. By this time, we had already finalized the criteria we planned to use in future evaluation in the schools and the leaders were "given" the criteria ready-made. The six criteria we (Friedlander and myself) chose were the following:

- Understanding the task;
- Understanding the concepts (in each task the relevant concepts—such as the meaning of multiplication, or the order of operations—are specified);
- Mathematical reasoning (e.g. reversed thinking, generalization, finding proof);
- Calculation skills;
- Communication: written explanations, the way the student displays the solution;
- Variety and creativity.

Each group was given six student papers on the *Getting Results* tasks and checked each paper twice, each time using a different method. I found that checking the papers of "unknown" students initially was easier for the teachers. They were able to discuss the work freely, without worrying about insulting each other. (Comments such as "Who taught this student this subject, he doesn't know anything?" or "How could a student write such silly things?" arose frequently. Since the papers were not from their

own students this caused no problem.) For homework the teachers had to check ten of their own students' papers, using both forms of evaluation. Since this workshop was at the end of the summer, the next workshop took place on the consecutive day.

4.4.3 The Third Workshop

The third workshop began with the teachers solving a new task, which, like the previous one, was geared towards their own level of mathematics. The second activity was intensive work on the assessment tasks they had graded as a homework assignment. Working initially in small groups and later within the framework of the whole group, they developed plans for teaching on the basis of the results. They divided the students into small groups and planned what sort of learning activities they would use with each group. This was a point that I reiterated constantly in the workshops: after the evaluation of the student papers, what next?

The following activity was to create tasks using pattern blocks. This was quite difficult for the teacher leaders, and often the tasks they came up with were not very mathematical. The final activity was an exploration of "growing shapes". The task involved taking any pattern block and forming a series in which each element is similar to the first one, but larger.

4.4.4 The Fourth Workshop

The fourth workshop opened with the participants solving the problem of determining the total number of squares of all sizes on a chessboard. The teachers were required to write a report describing their processes of solution. These reports were then used as the focus in the ensuing discussion: what can one do with such written reports?

We discussed criteria for evaluating the written reports. I had discovered that the process of formulating such criteria requires the teachers to decide what are the crucial

elements in the activity in question. By formulating criteria for evaluating a written report, for example, they must first determine the important elements in a report.

This led to a discussion of how one can evaluate individual students when the report is written by a group of students. I described the three-part method created at the Weizmann Institute for our work with heterogeneous classes at the junior high school level. This method is based on a group report, a class synthesis and a subsequent individual report.

The next activity was geared towards a discussion of assessment. The teachers debated the question: "What does it mean to know (in relation to mathematics in elementary schools)?" This question was crucial because it forced the teachers once again to deal with the issue of what sorts of things we want to assess. I continually wanted them to focus on the "big ideas", encouraging them to reject the traditions of testing for many different individual skills.

Then we worked on the fourth-grade student textbook (as previously mentioned, over 90% of all Israeli elementary-school pupils study from the same mathematics textbook series). One chapter, "From a Thousand to a Million", was examined and the teachers tried to determine what it means to "know the material" in this chapter. They were asked to consider what skills or understanding we really expect the students to have acquired by the time they have reached the end of the chapter, taking into consideration the existence of calculators and computers.

The teachers then solved the *Population* task (Appendix A) and discussed what material from the chapter "From a Thousand to a Million" it assessed. We then looked at the general criteria for assessment and adapted them to the *Population* task.

The final activity of the day was to create an investigative task. The teachers were presented with four minor mathematics problems and were asked to turn one of them into an authentic task for assessment purposes.

4.4.5 The Fifth Workshop

Like the previous workshops, the fifth meeting began with the teachers solving a mathematical problem, geared towards their own level of mathematics. As usual, they were requested to write a group report of their process of solution and this was followed by a discussion about generalizations and justifications.

Then we discussed their work in the classrooms. As usual, the discussions at first took place in small groups, followed by the task of writing “group reports”, requiring them to formulate their thoughts with clarity. This initial stage of writing required them to focus on the crucial aspects of what it is they had to say on the subject, sifting out the trivial or non-essential elements. It was also an assessment tool for me, providing me with crucial information on what they were doing, and helping me plan future workshops in a more relevant way. In other words, teachers’ reflections served as formative assessment tools for the teacher-educator.

The teachers then solved the investigative task *Trains* (Appendix A). They evaluated five student papers using conventional grading methods and wrote their conclusions from this activity. Their conclusions included subjects which would need further work. The teachers then adapted the general criteria for assessment to the *Trains* task and they evaluated the same five student papers, using these criteria, and recorded their conclusions. The final part of this activity involved a comparison of these two methods of evaluation. These comparisons were written on transparencies, presented before the group, and discussed.

For homework they were given the following assignment:

1. Use the investigative task *"Add-a-Digit"* (Appendix A) in one fifth grade under test conditions, and make sure you are present while the students work.
2. Check the students' papers twice: by means of conventional grading and with the use of criteria.
3. Write your conclusions, making sure they include:
 - a. Thoughts about the performance of the class.
 - b. Ways of teaching further lessons to help those students who experienced difficulties.
 - c. Other comments.

4.4.6 The Sixth Workshop

The workshop began, as usual, with a mathematical problem for the teachers. Afterwards, the teachers were required to fill out a questionnaire concerning their use of tasks and especially about grading with the help of criteria.

"Add-a-digit" – an estimation task

1. How was this task different from the tasks that the students usually do?
2. In your experience, does checking tasks with the use of criteria provide you with information that conventional grading doesn't give you? yes / no
If yes, describe this information.
3. Additional comments.

The questionnaire was followed by a full group discussion.

In general, the teachers were reacting positively to the use of assessment tasks (see Chapter 4). Nevertheless, they tended to limit the use of these tasks to course assignments, while relying on conventional short-answer tests for building class profiles. This highlighted the need for activities such as the following one, which was motivated by a discussion I had with a teacher in an in-service course at my teachers'

college. She was working as a teacher-leader and had been exposed to alternative assessment and to the use of investigative tasks for assessment purposes. She told me about her need to write basic skills tests in order to create class profiles for the school principal and supervisor. We discussed the reasons she was writing such tests, the kind of information they would provide her with and the kind of mathematics she wanted her pupils to acquire. She expressed many conflicting thoughts and issues, which I decided should be presented to other teacher-leaders as well. Thus, I developed an activity which I anticipated would be appropriate for this purpose:

Second Grade Subtraction

1. Write all the mathematical activities you want students to experience, which are related to subtraction in second grade (that is, subtraction with numbers up to 100).
2. For each activity, provide an example of an appropriate assessment item.

This activity proved to be quite provocative. The teachers managed to list 18 different subtraction skills, and it became evident that many could **not** be assessed by means of the traditional paper-and-pencil short questions. Here again, we were discussing the different types of possible assessment tools and the type of information each tool would provide. I continued to emphasize that alternative assessment meant expanding the types of tools we used and not merely replacing standard tests with better tests. The teachers also saw that for many of the assessment tools, one would need some sort of mathematical problem to be solved—and thus the investigative tasks could be used in many different ways, and not solely as replacements for tests.

4.4.7 The Seventh Workshop

This workshop began with the problem entitled *1997* (the workshop took place immediately after the beginning of the new calendar year). The task requires writing

exercises with results from 1 to 100, using all the digits of the year, each one no more than once. This led to a discussion of what factors are important in a task.

The following activity was based on Sullivan and Clarke's article, "Catering to all Abilities through 'Good' Questions" (1991). I discussed how a teacher can still find time to use such questions despite the pressure she is under by the requirement to cover all the topics in the national syllabus. As is often the case, this generated a discussion about the length of time that is required for changes in teaching styles to take effect.

The activity of "opening up" questions from problems in a textbook was one that became increasingly central in their teaching. They were given a page from the textbook and asked to "open up" the questions. This is a skill that in time became basic to them and was one they found very helpful in their work with other teachers. In the past I have often been disillusioned when encouraging teachers to create their own tasks. I have often found the level of mathematics that they exhibit to be low, and the tasks they create often lack emphasis on important mathematical concepts. Thus, I gradually became convinced that my goal would be better served by aiming at a less ambitious goal: teaching these teachers to use their textbook in a better way. As previously mentioned, the textbook series used by most of the schools in the country was written by a highly professional staff, whose academic advisor is Professor Pearla Neshet. My goal therefore became to help them understand how to "open up" the questions and create questions that lead to investigations. In addition, I taught them how to use the questions in the text in order to develop mathematical discourse among the students.

This workshop concluded with a discussion of the tools the teachers felt were lacking in order to continue the process of expanding their tools for assessment use, and to aid their work with other teachers in this field.

As a homework assignment, they were asked to choose a new task, to try it out on an entire class, to check the students' papers and to write conclusions. The teachers were

asked to formulate five questions to present at the next workshop, arising from their desire to plan further work for the class.

4.4.8 The Eighth Workshop

This workshop began with the participants viewing a video in which I was seen teaching geometry to a group of students. In my courses I sometimes use videos of my own teaching in order to present participants with examples of practices I would like teachers to use. I find that teachers often have trouble visualizing the practices I describe and that seeing these practices in action makes it seem possible to them. If possible, I teach a participant's full class. I have found this to be a very powerful activity: the participants of the workshop know that the students in the class are unused to working with the methods I use, but nevertheless can be seen to be thinking, exploring and discussing mathematics. When it is impossible to use live students, then I use videos.

We continued with a discussion of the questions they had brought in response to the homework assignment. They then responded to a final questionnaire. This questionnaire was similar to that used in the Second Assessment Course (see above) with an additional question:

6. What is different in your class / in your supervision as a result of this course?

Explain with examples. (Different mathematical content, different in styles of teaching, different in ways of supervising, etc.)

4.5 Follow-up Work with Course Graduates

4.5.1 Ofek

The *Ofek* courses are in-service workshops given by satellite to approximately 20 centers throughout the country (the precise number depending on the registration for each course). The system includes a closed-circuit television broadcast, during which the participating teachers can telephone the lecturer, although he cannot see them. The telephone questions and the lecturer's response are broadcast to all the centers.

The *Ofek* courses on School-Based Assessment consisted of an interactive television broadcast, 1 ½ hours long, and a workshop, also 1 ½ hours long, at each center, given by local teacher-leaders who had met with the lecturer and had been involved in the planning of the course. Each course contained a number of general sessions on alternative assessment, with the participation of Ministry of Education representatives who explained the philosophy behind school-based assessment and sessions on each specific subject: mathematics, social studies and language. These *Ofek* courses were geared towards school principals and elementary school teachers. Principals were encouraged to attend together with two or three teachers who specialized in the various subjects, who would then aid the other teachers in the school in the use of school-based assessment tasks. My interest in the *Ofek* courses was two-fold: not only was I interested in the reactions of the participants, but the teacher leaders who were running the workshops were mostly graduates of the national courses on assessment; thus, the *Ofek* courses afforded me with a unique opportunity to observe the effect of these teacher leaders' continuous work with assessment tasks.

In January 1997, I met with a group of teacher-leaders hand-picked to represent different regions throughout the country, to plan the mathematics sessions of the first *Ofek* course. We had to plan three meetings, each including a broadcast by myself, to encourage active participation, and a workshop run by the teacher-leaders. I was initially opposed to the concept of *Ofek*, since I was concerned that the fact that I could not see the participants' reactions would render the course ineffective.

However, when I began relating to *Ofek* as a course given by the teacher-leaders, where I only gave input, the idea became attractive. The ability to reach over two hundred teachers simultaneously, knowing that the follow-up workshops would be on a high level, was appealing.

In order to help the teacher leaders plan the *Ofek* course, I asked them to consider two questions: "What do we want at the end of the process?" and "Given these goals, what sort of activities are necessary for the teachers who participate in *Ofek*?" One activity on which we worked was that of "opening up a problem".

Opening Up a Problem

Sample problem:

A number of people got on a bus. At the first stop, $\frac{2}{5}$ of the people got off and $\frac{3}{5}$ of the original number got on. At the second stop, $\frac{1}{2}$ of the people got off and $\frac{1}{3}$ of the number that was left on the bus got on. At the last stop, $\frac{3}{4}$ of the people got off, leaving 5 people on the bus. How many people were on the bus before the bus reached the first stop?

How can we open up this problem?

What other numbers, in addition to 5, could be used?

What do all these numbers have in common?

Write a similar problem with somewhat different constraints.

Using such activities, we constructed a plan for the mathematics sessions of the first *Ofek* course. This was a course of 14 meetings, each four hours long (a total of 56 hours), which started in February 1997. There were two general meetings on alternative assessment, followed by four meetings on mathematics, four meetings on language and four on social studies. In October 1997, we met again to plan the second *Ofek* course (see Chapter 6, Section 6.4.1). In January 1998, we met to sum up the first part of the second *Ofek* course and to plan the second part. I used this meeting as an opportunity to hear about the teacher-leaders' experiences using the

Assessment Task Bank in various parts of the country: to find out how teachers were relating to its use for assessment purposes. The discussion was taped and transcribed. The question I posed to the teachers was: "Tell us about a school or a classroom in which you see positive processes taking place, and tell us what this means."

4.5.2 Additional In-Service Workshops

For the school year of 1998–99, I directed in-service workshops for a select number of teacher-leaders from various regions of the country, who would then run workshops on assessment in their home areas. Participants were teacher-leaders who had completed a national course in assessment and who were recommended by the Ministry of Education's local supervisors. Teacher-leaders from the *Ofek* courses were included in these workshops.

The first workshops took place during the summer vacation. At the end of the workshop, the teachers responded to a questionnaire I had prepared:

Questionnaire - 2.7.98

1. Record three important things that you learned in these two days.
2. Would you like to continue to meet during this year?
If so, what subjects would you be interested in?
3. Comments on the process of implementing the Assessment Task Bank.

4.6 Interviews with Teachers

In June 1998, I interviewed three of the leading teachers who had been working with the Assessment Test Bank for several years. Each of these teacher-leaders was from a different sector of the Israeli school system: religious, secular and kibbutz. These were unstructured interviews undertaken in an effort to determine the influence of the

Assessment Task Bank on their work. Excerpts of these interviews are provided in Chapter 6 (Sections 6.5.1, 6.5.2 and 6.5.3).

Two and a half years later, in January 2001, I interviewed another teacher-leader. This was a teacher who had been working with the Assessment Task Bank since its inception. The questions I asked were geared towards finding out her feelings about assessment after so many years of use of the Assessment Task Bank. Excerpts from this interview are also provided in Chapter 6 (Section 6.5.4).

CHAPTER 5: DATA COLLECTION IN THE SHARON COMMUNITY

As previously mentioned, school-based assessment in mathematics was introduced to the community of Sharon in December, 1996—a year and a half after the first national course on assessment. The small community with six elementary schools, was chosen for the intensive study because the leading teachers of each school were already involved in an in-service project, which was geared towards changing their concept of mathematics. The Sharon project was an interdisciplinary intervention project, the purpose of which was to integrate computers into the teaching of mathematics, Hebrew language and science on the elementary and junior high school levels.

5.1 The Method of Data Collection

The data collected in the course of the Sharon study included extensive background questionnaires from all twenty-nine elementary school teachers who participated in the assessment course. In addition, at two stages during the project, I interviewed teachers using semi-structured interviews. On one occasion these were mathematics teachers; the other time these were mathematics school coordinators. In addition, I collected tests from all participating schools several times during the project.

The data collection process is summarized in the following table:

5.1.1 Summary of Data Collected

	Activity	Date	Sample	Data collected
Second Year	Assessment Course – 3 workshops	Spring, 1997	29 teachers	1 background questionnaire – 29 teachers responded
	Sample tests	Spring, 1997	Collected from all 6 schools	Sample tests
	Interviews	June, 1997	5 teachers	5 semi-structured interviews

Third Year	20 workshops	School year, 1997-98	14 teachers	Final assessment projects – projects of 6 coordinators analyzed
Fourth Year	8 workshops	School year, 1998-99	6 math coordinators	July, 1999 End-of-year Questionnaire – 3 coordinators responded
Interviews		Spring, 2001	4 math coordinators	4 semi-structured interviews

5.1.2 Construction of the Background Questionnaire

This questionnaire provided information about the teachers' education, their teaching experience, examples of typical tasks they used in their mathematics lessons and in their assessment practices, and information about the way their school worked together as a team. The questionnaire was developed with the assistance of an expert at the Weizmann Institute of Science, who specializes in writing questionnaires for teachers, particularly in science education. Because the sample size was small (only 29 teachers), it was decided that the questions would be open-ended.

Background Questionnaire

Questionnaire to Teachers in Sharon

I. Personal information.

A. Teaching Experience:

1. Number of years of teaching experience _____
2. Number of years of teaching mathematics _____
3. This year I teach mathematics in grades _____

B. Education:

1. Graduate of _____ teachers' college
_____ Teaching certificate
_____ B.Ed.
My major was _____
2. Graduate of _____ university
My major was _____

II. Choose one class in which you teach mathematics and answer the following questions:

Which grade is it?

1. What topics did you teach in mathematics in this class last week?
2. Give examples of two tasks which the students were given last week.

III. Ways of assessment:

1. What tools (e.g., tests) do you use to assess your students' mathematics achievements?
2. Why do you check your students' mathematics achievements?

IV. Is it usual in your school for the mathematics staff to work together?

If yes, give details:

1. How often does the staff meet?
___ once a week ___ once a month
___ once every two weeks ___ a few times a year
2. Who participates in the meetings?
___ all the math teachers of the school
___ all the math teachers of one or two grades
3. What are the subjects usually discussed?
4. What did you do at the last meeting?

V. Requests or expectations from the meetings about assessment.

5.1.3 The Follow-up Questionnaire

A follow-up questionnaire assumed the following form:

**Follow-up Questionnaire
Questionnaire (July 1999)**

1. Give examples of two tasks that the students performed as class work, from the last week of their studies.
2.
 - a) What tools do you use for student assessment?
 - b) Why do you check student achievement?
3. What was the contribution of our workshops regarding student assessment?
4.
 - a) Did you use a task from the assessment task bank in the last half of the year?
 - b) If so, which task and in which class?

The purpose of this questionnaire was to gather data relating to my research questions. The similarity of these questions to those on the background questionnaire enabled me to draw comparisons, thus obtaining information about changes over time.

5.1.4 Interviews with Teachers

In June, 1997, I interviewed five mathematics teachers, each from a different school in the Sharon community and none of whom was participating in the ongoing weekly workshops of the in-service project. The purpose of these interviews was to obtain data for my research questions from teachers who had not participated in extensive workshops. In other words, I wanted to neutralize the effect of other factors, gauging the effect of the three assessment workshops alone. I wanted to see whether three workshops could affect their views of mathematics, the way they taught and the way they assessed their students. The following questions formed the basis of the interview:

Interviews with Teachers

Did these workshops affect your teaching?

Did these workshops affect your methods of assessing students' achievements?

How do you create tests?

Have you used the mathematical investigation tasks in the Assessment Task Bank?

What did you learn from using them?

Does your department work as a team? Do you create tests together? Do you talk about students' achievements?

What do your report cards look like?

Are there things you will do differently next year when you have new classes?

5.1.5 Interviews with Teacher Coordinators

Two years after the end of the project, I interviewed four of the six mathematics coordinators. These were semi-structured interviews which enabled me to determine each coordinator's position vis-à-vis use of the investigation assessment tasks, and to evaluate how these tasks had influenced her teaching and her view of mathematics.

The following questions formed the basis of the interview. As in any semi-structured interview, the other questions asked depended upon the answers to these questions.

When you prepare a test today, what do you do?

What are the elements of a good test?

What would you say to a teacher who showed you this test? [I showed them a test I had received from a teacher recently, which I thought was not a good test.]

All interviews were taped and subsequently transcribed, enabling detailed analysis.

5.1.6 Sample Tasks and Tests

I collected sample tests and tasks used by the teachers and teacher-leaders involved, and analyzed them using criteria I had developed. (For further details and references, see Section 4.1.4.)

5.2 The First Year of the Sharon Project

During the 1995–96 school year, nine teachers of mathematics participated in this project, in which I served as academic advisor. They met each week for an eight-hour-long workshop, half of which was devoted to computer skills (Windows, Word, Powerpoint and Excel) and half of which was concerned with mathematics. In the course of the mathematics workshops, teachers were given problems to solve, with an emphasis on geometrical investigations. Through the activities they developed the realization that mathematics does not solely involve computational skills. During the first year of the project, the workshops did not deal with activities for students at all.

5.3 The Second Year of the Sharon Project

In the second year of the project, 1996–97, five additional teachers joined the original nine participants. That year I ran all the workshops. The six-hour-long weekly workshops were devoted to three main goals: 1) strengthening the mathematical knowledge of the teachers (teaching for example, topics such as probability, which has recently been introduced in the national curriculum for elementary schools); 2) working on teaching methods and developing new ways of organizing the classroom (for example, heterogeneous groups instead of homogeneous ones); and 3) working on mathematics investigations and challenging problems, which the teachers then tried out with their students one hour a week.

5.3.1 Assessment

Assessment did not constitute a part of these workshops. I had decided that in Sharon, school-based assessment would be introduced first to the principals (Nelson 1998) and then to all mathematics teachers—and not just the leading ones participating in the weekly workshops.

In December 1996 a workshop was held for the elementary-school principals and their supervisor from the Ministry of Education on the rationale of the School-Based Assessment Project. The principals solved a mathematics investigation task, discussed its characteristics and were introduced to other tasks and samples of student work. They were shown how these tasks enable teachers to assess the mathematical achievements of their students in a broader way than is possible with conventional single-answer tests, which place the emphasis on computation skills.

During the spring term, I ran three workshops for all fourth- through sixth-grade mathematics teachers from the Sharon community. Approximately 25 teachers participated in each workshop. In the workshops the teachers solved mathematics assessment tasks, and discussed their students' solutions to similar tasks. They also discussed how such tasks could be used for assessment purposes.

5.3.2 The First Workshop

This three-hour workshop began with the teachers filling out a background questionnaire (see Section 5.1.2). Then we held a discussion on what they considered important in teaching elementary school mathematics and on how they assessed their students. There were many discrepancies—evident to the teachers themselves—between what they considered to be important and their assessment practices. This led to an introduction of the purposes behind school-based assessment in general, and its practice in the field of mathematics in particular (see description of School-Based Assessment Project, Chapter 1, Section 1.4).

The teachers then solved the investigative task *Getting Results* (Appendix A). We then discussed what knowledge a student needed to possess in order to succeed in this task and what attributes the task had (what was “special” about it). These attributes were generalized and expanded to a presentation of the principles underlying all the tasks in the Assessment Task Bank.

Afterwards, the teachers checked samples of two students’ work on the task *Trains* (Appendix A), and attempted to draw conclusions from this exercise. Analyzing students’ work in an effort to discover the extent of their knowledge is a skill teachers need to acquire, rather like learning how to listen. As a homework assignment, the teachers were asked to use the task *Getting Results* as an assessment tool in their classes and to draw conclusions from this exercise.

5.3.3 The Second Workshop

This three-hour workshop took place three weeks later. The teachers first worked in small groups on the problem detailed below; the goal was to get them to experience for themselves what it meant to “investigate mathematics”. After working on the problem and discussing it in small groups, each group summed up its results in three generalizations, which they then wrote on overhead transparencies and presented to the full group. We discussed these generalizations from a mathematical point of view, with the teachers participating as students.

This led into a long discussion on the meanings of verbalization, generalization and justification and their distinctions. Finally, the groups were given an additional overhead transparency, and were asked to consider how this experience could influence their teaching.

The problem the teachers worked on is presented below:

$$15 - 14 + 13 - 12 + 11 - 10 + \dots + 3 - 2 + 1 = ?$$

$$101 - 100 + 99 - 98 + \dots + 3 - 2 + 1 = ?$$

$$1001 - 1000 + 999 - 998 + \dots + 3 - 2 + 1 = ?$$

$$14 - 13 + 12 - 11 + 10 - \dots - 3 + 2 - 1 = ?$$

$$100 - 99 + 98 - \dots - 3 + 2 - 1 = ?$$

$$1000 - 999 + 998 - \dots - 3 + 2 - 1 = ?$$

- a) Solve the above problems.
- b) Formulate a generalization.
- c) Explain why your generalization is correct (i.e. provide justification).

The teachers were divided into groups according to the classes they teach (fourth-grade, fifth-grade and sixth-grade teachers), and they discussed the conclusions they had drawn from using the *Getting Results* task in their classrooms and planned follow-up teaching on the basis of the results. This was followed by a discussion in the forum of the entire group, which focused on three issues: 1) What did I learn about my pupils? 2) How will having used this task affect my teaching? and 3) Is this task an assessment tool?

The workshop concluded with the teachers being asked to compare the work of two students on the *Animals* task (Appendix A) and to consider which student had had more success. The purpose of this exercise was to get the teachers them to analyze the abilities the task required of the students.

As a homework assignment, the teachers were asked to use the *Animals* task as an assessment tool in their classes and to write conclusions.

5.3.4 The Third Workshop

This workshop, 2 ½ hours long, took place a month later. The teachers first solved the *Bridges* task (Appendix A); then they watched a video of a pair of students solving *Bridges* and were asked what we could learn about the girls' mathematical abilities. This led to a discussion of additional criteria for assessment, their first introduction to such criteria. Afterwards they used these additional criteria for assessing another pair of students, who were working on the same task.

This was followed by their adapting the additional criteria to the *Animals* task, and working in groups; they then used these criteria to check their students' performance on this task.

Since this was the last workshop planned, the final discussion focused on the question of what was expected of them in regards to the use of the Assessment Task Bank. They were informed of the individual workshops that were planned for each school separately (see below), and thus there were no special expectations.

5.3.5 Workshops at the Individual Schools

During the final two months of the school year, I gave a separate workshop at each school for the entire staff: the teachers and the principal. The purpose of the meeting was to discuss assessment in general and to help all teachers—not just those who taught mathematics in grades four through six—to understand the background and the underlying concepts of the School-Based Assessment project.

Each of these workshops began with all the teachers working on a mathematics investigation task. This was a problem about placing coins on the sides of square, and required formulating a generalization. The first steps of the problem could be solved intuitively, using real coins, and this enabled teachers of subjects other than mathematics to get immediately involved. Nonetheless, the problem involves enough mathematics to capture the interest of mathematics teachers.

The problem led to a discussion about why this is a mathematics problem: after all, there are no formal exercises and none of the usual skills most teachers relate to. The mathematics is in fact necessary in order to formulate a justification of the generalized solution. For those teachers who had not attended any of my workshops, these were new concepts. We also discussed the method of working that they had chosen—small cooperative groups—and pondered why it was so successful.

The second item on the agenda was a discussion about assessment. We questioned what assessment is, what is it that we want to assess and what tools we have for assessment purposes. In most schools this discussion expanded into a general one, not restricted only to assessment in mathematics. Many teachers, especially those who were not mathematics specialists, found the questions very interesting, stimulating and relevant. They were concerned with assessing a wide range of abilities, not just basic knowledge, and looking for ways to accomplish this. Many of these teachers were not aware of these dilemmas in their mathematics teaching because they had viewed mathematics as a subject involving skills only.

We discussed the need for tests which assess skills and the ability of investigative tasks to assess additional abilities, thus expanding our picture of the student's achievements.

The workshop concluded with some practical matters on the availability of the Assessment Task Bank and its correct usage.

After each workshop, I held a meeting with the principal and the mathematics coordinator to make sure that they understood their roles in this project. The principal received the Assessment Task Bank for the fourth grade and assumed responsibility for using the tasks for assessment only and not as teaching worksheets.

5.3.6 Sample Tests

I collected sample tests that the schools were using during the spring term. Wishing to see what use the schools would make of the investigative assessment tasks, I purposely did not emphasize their use in the weekly workshops with the leading teachers.

5.3.7 Interviews

At the end of the school year, in June 1997, I interviewed five mathematics teachers, each from a different school in the Sharon community and none of whom was participating in the weekly workshops. I wanted to see what influence the three special assessment workshops exerted on the teachers. These teachers were chosen through the recommendation of the schools' mathematics coordinators. (Two of the teachers joined the workshops the following year.)

The interviews were semi-structured: they began with the same questions, but the subsequent questions varied, depending upon the teachers' responses (Section 5.1.3). I read, re-read and summarized each interview, categorizing responses and writing main points.

5.4 The Third Year of the Sharon Project

It was decided that during the following school year, the workshops for the leading teachers would be held every fortnight, each one six hours long. Again there were fourteen participating teachers (eight from the previous year and six new ones). (As happens with all in-service projects, some teachers drop out due to other commitments, such as their academic studies, and others are encouraged to drop out when it is felt that they are unsuitable to lead other teachers in their schools.) I ran all the workshops during that year, with the exception of several workshops on using computers in mathematics. The focus of the workshops changed somewhat in that the

teachers requested that we address the problem of integrating the investigation tasks and challenging problems with the standard syllabus required by the national curriculum. Emphasis was placed, therefore, on methods for using the textbooks in innovative ways. Such methods included, for example, the use of the challenge question at the end of each chapter as an investigation to “open up” a lesson with a heterogeneous group of students, leaving the drill to come as follow-up practice. I continued to work on enriching the mathematics of the teachers and to provide them with additional teaching methods, such as the incorporation of games.

I continued to refrain from mentioning the use of the school-based mathematics assessment tasks. I wanted to observe whether teachers whose teaching styles were gradually changing would encounter a need for new types of assessment and whether, in this case, these tasks would help.

In March 1998, as I was disappointed to find that they were not using the tasks (data acquired from the teacher coordinators), I changed my tactics and required each teacher to use one, as a reminder that they existed. For their final project in June, they were required to collect samples of questions that would enable them to determine for each student in their class whether he rated “high”, “average”, or “low” on each of the criteria (understanding the task, understanding the concepts, mathematical reasoning, computation skills, communication and creativity). I collected these class profiles along with the sample questions.

5.5 The Fourth Year of the Sharon Project

During the fourth year, I held monthly meetings with the department coordinators. We continued to solve challenge problems together as a way of keeping them involved in investigating mathematics. Assessment was not a major topic discussed. This was the last year I worked in the Sharon community.

5.6 The Follow-up Questionnaire and Interviews

At the end of the fourth year of the project I distributed a questionnaire (Section 5.1.3) to all mathematics teachers in Sharon.

Three of the six department coordinators filled out this questionnaire. Almost two years later, in the spring of 2001, I interviewed four of the six department coordinators. The fifth had left her school for a new city, and the sixth refused to be interviewed. This teacher had continued to resist reform mathematics and to insist that the emphasis in tests should be on computational skills.

CHAPTER 6: THE RESULTS OF THE NATIONAL STUDY

In this chapter the results of the national study are described. As explained in Chapter 3, the national study is based on three national courses on assessment, in which I instructed teachers how to use the investigative tasks of the Assessment Task Bank.

Rather than presenting the results of the courses sequentially, the findings were broken down into three main categories which correlate to my research questions. Within each research question, the results of each course are given separately, in sequential order, since I feel that this way one can see the progression of the changes that the teachers underwent. Afterwards, I have described my experiences with the *Ofek* satellite courses where the teacher-leaders were involved in the planning and running of the workshops. In addition, there are four follow-up interviews and an analysis of tests which were collected throughout my research.

6.1 The Teachers' View of Mathematics

The first question that the study attempted to address was:

Will the teachers' use of mathematical investigation tasks for assessment purposes influence their view of mathematics?

In the transcripts that follow, teacher responses are coded as follows:

teacher CD.III.2; p.40.

Refers to teacher CD, who attended the Third National Course on Assessment. The quote is taken from her response to the second questionnaire. The full text of her response to the questionnaire appears on page 40 of Appendix C.

6.1.1 The First National Course on Assessment

As mentioned in Chapter 3, the teachers were required to fill out a questionnaire during the second workshop of this course. (There were 24 teachers present at this workshop.) The fact that the questionnaire was given at a relatively early stage in the course meant that the results were not influenced significantly by workshop and/or field experiences. One of the questions posed on the questionnaire was: “*What were the results of your use of the tasks with the students?*” By this point, the teachers had solved five investigative tasks (*Table Arrangements, Geoboard Quadrilaterals, Trains, Population* and *Add-a-Digit* – Appendix A) and had tried some (or all) of them out with their students.

Most of these teacher-leaders had originally viewed mathematics as involving solely the acquisition of basic skills. Thus they reacted strongly to the new requirements that students explain and formulate generalizations. More than half of the teachers’ comments (14 out of 24) related to mathematical activities in the questionnaire.

The students’ ability to explain verbally is quite low. They are not able to summarize their explanations in writing, even when their thinking is correct. ... The hardest task was on estimation. Very few children used estimation; instead, they calculated everything. (teacher AV.I.1; p.8)

Teachers AX, AT, and AQ all pointed out that generalizations and rule formulating were difficult for the students, thus reflecting their views that these abilities were elements of mathematics.

The levels of difficulty were appropriate—most of the children succeeded on the easy questions, but the generalizations were impossible for them. One of the tasks, *Estimation*, I used in a good school, and some children actually answered the hard parts and managed to generalize. (teacher AO.I.1)

The tasks in general were not easy for the students. They performed the calculations, but didn’t use methods like estimation. Generalization—mostly didn’t get to this level, and the verbalization was very hard (only a few succeeded). (teacher AM.I.1)

Trains—in one school the justifications were correct and well written, yet in the other school no student could explain what they had done—they had all worked only by trial and error. (teacher AJ.I.1)

The students, in general, gave good explanations—the teacher has been working for the past few months on oral and written explanations. (teacher AH.I.1)

All the students had difficulty with the *Population* task. ... Most of the pupils had trouble with verbalization – although they performed correctly they didn't know how to explain. I saw weak students who succeeded by using concrete materials—even though their teachers said they were very weak. (teacher AF.I.1; p.1)

Trains—10 students in a heterogeneous group. Most of the students had trouble making generalizations—only four mentioned even numbers, multiples of 4. Most of the students counted until they found the right number. (teacher AE.I.1)

As is evident in these quotations, a recurrent response is that the students were unable to perform mathematically in the manner required in the tasks. What I found in the workshops was that these activities were in fact difficult for the teachers as well. They themselves were unaccustomed to formulating generalizations and justifications, but preferred to search for the right answer. Some even complained about the need to explain in writing, (e.g., “I chose to teach mathematics because I do not like to write!”). As can be seen from the comment by teacher AH.I.1 (above), when the students are taught these skills they are likely to succeed at them. It became increasingly evident as time passed that if students apply higher-level mathematical thinking, they will be able to perform at that level. As discussed in Chapter 2, the problem in mathematics teaching at the elementary-school level in Israel, like in other countries, is the lack of mathematics knowledge of the teachers themselves.

After the questionnaires were filled out, we held a group discussion. One teacher talked about the new sort of mathematics being proposed: “The tasks give another focus beyond ‘the meaning of multiplication’ and skills.” It was thus evident that after

even a very limited exposure to the investigative tasks teachers began to use a new kind of language to describe mathematical activities. This lends credence to one of the claims being examined in this study, namely that instructing teachers in the use of mathematical investigative tasks would influence their view of mathematics.

6.1.2 The Second National Course on Assessment

At the sixth workshop, the teachers filled out a questionnaire which related to the use of criteria for evaluating student performance on the investigative tasks. (For descriptions of workshops one to five, see Chapter 3.) In the First National Course on Assessment the teacher-leaders were only exposed to investigative tasks, whereas in the Second National Course they were taught about additional criteria for assessment. (For a definition of the criteria used for evaluation purposes and how the teachers were taught about them, see Chapter 3, page 55.) Following are some of the teachers' responses regarding the way in which the use of criteria had influenced their choice of tasks to give their students.

This sort of checking contributes to building tasks which emphasize the process of solving the problem and do not only have one answer.
(teacher BH.II.1)

- a) The students work more with manipulatives, even on tests.
- b) I use more open problems.
- c) I use less activities which are only exercises—and more do the drill on the computer / and emphasize more activities in which process is important.
- d) I require verbalization (and generalization) in the students' work.
(teacher AY.II.1; p.9)

It can be seen that after approximately six months of exposure to the use of criteria in the assessment of investigative tasks, these math-leaders were themselves influenced by them. They put emphasis on different aspects of the students' work, and not merely on whether their answers are correct or not. Furthermore, they assimilated some of the

criteria—such as communication (verbalization or justification)—as aspects of mathematics that require emphasis.

6.1.3 The Third National Course on Assessment

At the fourth workshop of this course, the teachers held a round-table discussion on the question: “What does it mean to ‘know’ (in relation to mathematics in elementary schools)?” This question was crucial because it once again required the teachers to deal with the issue of what sorts of things we want to assess and what is important in mathematics in the elementary school (i.e., “What is mathematics for you?”). The table below presents the reports of the various groups:

What Does It Mean to “Know”? in learning mathematics on the elementary-school level
Active construction of the knowledge, at every stage that the student is at. To understand that the rules of mathematics do not change, but rather expand to further fields. To know the algorithms of arithmetic operations. To understand the meanings.
To be able to represent a concept with a concrete model. To use concepts. To evaluate the knowledge—both of understanding and of automatic responses, memory.
To know is a state of “creating and unraveling”. It is an unstable state.
To know the connections between subjects.
To be able to open the “correct drawer” for each given problem.
To be able to deal with authentic problems using one’s prior knowledge.
Understanding and skills.
Knowledge of skills leading to interpretations, leading to the ability to think critically.
Understanding and control of skills.
The ability to use the skills for dealing with new situations, and new problems.
To know how to explain in various ways—orally, written, etc.
Control of skills. Thinking strategies.
The ability to use concepts learned in daily problems, and in problems which require generalizations. Checking. Critical thinking of processes and products.
Understanding meanings.

Evidently, the teachers' view of elementary-school mathematics had expanded: it was no longer perceived merely as a collection of arithmetic skills.

By the sixth workshop the teachers were expressing these ideas even more explicitly:

In mathematics, we want to work with the central ideas on which mathematics is based, and not with technical computations. This task checks whether the student can draw the connections between estimation, multiplication and place value. Until now we checked whether he could compute estimated answers and round numbers. We never drew the connections between them. (teacher CO.III.1)

On the questionnaire distributed in the final workshop, the teachers continued to express these revised views of mathematics. The total change in outlook which occurred in many teachers is encapsulated in one teacher's words:

- The use of verbalization on the tasks. For example, on most of the tasks I ask the teachers to verbalize and arrive at a generalization. It is most important for me to emphasize with the teachers the process rather than the final answer.
- Work with heterogeneous groups. The work meets the needs of most of the pupils, how to work, how to solve, that's an important focus I take from the course.
- Investigative tasks. To start workshops with investigative tasks and to give them to the pupils—how to assess such tasks.
- How to work with teachers on solved tasks—that is, teachers should assess students' work. To see the difficulties in assessing.
- The mathematical content. I've changed somewhat in that I have learned to open tasks from the book, and not just to accept the book as a given authority. (teacher CB.III.2)

The changes in teaching induced by the teachers' exposure to this revised concept of mathematics can broadly be divided into four areas:

1. Types of activities: how the characteristics of the investigative tasks affect classroom activities;
2. The use of additional investigative tasks for teaching;
3. The emphasis placed on verbalization;
4. Changes in the mathematical content of the tasks.

1. Types of Activities

Teachers related to the way in which the attributes of the investigative tasks used for assessment affected the types of activities they planned for the classroom.

The features of the assessment tasks have shaped the activities in the subjects which are taught in class. (teacher BO.III.2; p.22)

The tasks I give students and teachers have changed. The emphasis is now on things which were missing, such as creativity in mathematics, justification. Also, the mathematical discussions in class are more open. (teacher BR.III.2)

One teacher summed it up as follows:

- 1) The main difference is that we now have more work in groups with emphasis on dialogue between the pupils.
 - 2) We plan longer tasks which are scaffolded.
 - 3) We put more emphasis on having the students explain their thinking.
 - 4) We plan tasks which include different math topics—more complicated tasks.
 - 5) I have the students write on transparencies; this encourages more precise work, more focused and more aesthetic.
- I am generally pressured to accomplish the material on the curriculum, and now I feel less pressured... (teacher CJ.III.2; p.47)

2. The Use of Additional Investigative Tasks for Teaching

Whereas mathematics previously consisted of only computations and skills, it is now clear that investigations are an integral part.

Because of the course I began using open investigative tasks, which are not directly connected to the subject being studied. (teacher BP.III.2; p.25)

Investigative tasks—at least one each week, accompanied by follow-up work.
(teacher BU.III.2; p.29)

I work with [the teachers] on the need to give investigative tasks of different levels, and to have all students in the class work on such tasks in heterogeneous groups, so that they can help one another. (teacher BW.III.2)

Once a week we spend an hour doing investigative tasks.
(teacher BZ.III.2)

3. The Emphasis Placed on Verbalization

Verbalization which includes justification has become an integral part of mathematics learning. Previously teachers tended to pass out worksheets and only check the correctness of the answers.

In my class as well as in my work with teachers, I emphasize several aspects: [among them] mathematical discussions - pupil with teacher, pupil with pupil, group with group, or the entire class...
(teacher BU.III.2; p.29)

In working with teachers, I emphasize meta-cognitive processes—the great need for verbalization and mathematical discussions with the students. (teacher BW.III.2)

I emphasize verbal explanations and justification in the solution to problems. (teacher BZ.III.2)

As a result of this course, I emphasize more verbalization—to describe the process of your work and the solution, to describe your original plan and hypothesis. (teacherCBF.III.2; p.44)

4. Changes in the Mathematical Content of the Tasks

For some teachers, the needed changes were clear:

less emphasis on basic mathematical skills and more emphasis on finding rules and generalizations (the same subjects but tackling them from a different angle) (teacher BV.III.2; p.32)

It is evident from the quotes above that the teachers were now expressing themselves in a different language than before. They were writing about a much richer concept of mathematics: one which includes generalizations and justifications and which emphasizes the thinking process and not just the computations and the final solutions.

The process of verbalization and communication became particularly important to the teachers, perhaps because they could immediately perceive its benefits. Yet it remained a problem area, as students were not accustomed to explaining themselves in writing. As one teacher (CC) wrote: “all the students have problems in written explanations” (III.1; p.36).

I found that the area most problematic for the students was communication, and thus it requires work. (teacher CF.III.1; p.42)

The process of verbalization is very important because the student needs to know how to deliver information, and we need to teach him how to do this. (teacher BY.III.1)

The process of explaining enables us to understand how the student worked on his solution. Was it by trial and error? Or did he use the tools of estimation? (teacher BX.III.1)

Difficulties: The children lack the ability to explain. This task confirmed my need to train the students to express their thinking processes on paper after learning tasks. Most of the children write too briefly. They need to get used to writing in more detail, to write as though the reader doesn't know anything about the subject.

(teacher BS.III.1)

When assessing through the use of criteria I saw that most of the children have trouble explaining. (teacher CL.III.1; p.48)

In sum, an examination of the results of the three national courses on assessment, as reflected in the questionnaires filled out by the teachers at various points during the courses, suggests that the teachers' use of mathematical investigative tasks for assessment purposes indeed influences their concept of mathematics.

6.2 The Way Teachers Teach

The second question that this study addressed was:

Will the teachers' use of mathematical investigation tasks for assessment purposes influence the way they teach, and if so, in what ways?

Those findings of the three National Courses on Assessment that pertain to this question are presented below.

6.2.1 The First National Course on Assessment

As mentioned in "The Teachers' View of Mathematics", the questionnaire filled out by teachers at the beginning of the course (during the second workshop) included the question: "*What were the results of your use of the tasks with the students?*"

Two teachers mentioned that the use of these investigative tasks for assessment would influence the way they taught:

The tasks can serve as examples for the type of activities we should prepare for the students in order to work on mathematical thinking.

The tasks are aids for the teachers—first they will use them in their classes, and eventually the tasks will change their teaching, and then the teachers will build different tests. (teacher AD.I.1)

One teacher however, was not sure how to reconcile the assessment tasks with the regular curriculum.

It wasn't clear to the students or to the teachers the connection between the tasks and what was currently being studied.

(teacher AL.I.1; p.5)

Of the 18 teachers who felt that the investigative tasks could indeed be used for assessment, 12 wrote about the need to make appropriate changes in teaching methods. Some of their responses are cited below:

In my opinion, these tasks are very good for assessment, on condition that the teacher actually works this way in her lessons.

(teacher AP.I.1)

These tasks can be used for assessment, but first one needs to use similar tasks for teaching in order to familiarize the students with this unusual way of thinking and also to give them the opportunities to explain themselves in writing. (teacher AV.I.1; p.8)

Students need to work on such tasks as in class before these tasks are used for assessment. (teacher AF.I.1; p.1)

Three teachers wrote about the influence the tasks would have on teaching (their own teaching or other teachers'). They felt that their own teaching was being assessed as much as their students' achievements.

The tasks are good for assessment, and and they were even more useful for myself to evaluate my own teaching than to assess the students achievements. (teacher AN.I.1)

These tasks are very important for assessing students, as well as for assessing the work of the teachers. The tasks can serve as examples for the type of activities we should prepare for the students in order to work on mathematical thinking. (teacher AD.I.1)

a) It was hard for the teachers to accept these tasks for assessing knowledge and mathematical thinking. They felt that the tasks criticized their teaching—and that if they were for assessment, then the curriculum needs to be changed!

b) It needs to be made clear that these tasks give new messages, and thus teachers need to do more than teach skills! (teacher AL.I.1; p.5)

These comments were very encouraging to me, because this is exactly what I was trying to do: to influence the way teachers teach by providing them with examples of investigative tasks for assessment purposes.

6.2.2 The Second National Course on Assessment

As discussed in Chapter 2, the importance of assessment lies in its impact on teaching. Thus, the teachers' responses to the question, "*Did checking with criteria help you plan your teaching?*", were significant to my study. Half of the teachers (12 out of 23) wrote that the use of criteria helped them to understand the extent of the students' knowledge, thus enabling them to form "study groups" on the basis of common problem areas.

The criteria helped me to plan my teaching in relation to the students—I could continue to work with each student in relation to his level. (teacher AM.II.1)

After analyzing the student profiles which are generated, one can easily plan for future work—one sees in which areas the students are weak and in which areas average. (teacher BA.II.1; p.12)

...points clearly to the problems and makes it possible to work with groups according to the problems discovered. (teacher BI.II.1)

...to build work groups according to the students' levels in relation to the criteria. (teacher BJ.II.1; p.19)

After checking by criteria, I group the students and work with them to improve their levels. (teacher AW.II.1)

Assessing through the use of criteria helps to divide students into work groups according to their different levels. (teacher BN.II.1)

Eight additional teachers gave specific examples of the ways in which they needed to plan their teaching methods. Three of these mentioned the need for more work with concrete materials. For example:

A pupil who is diagnosed by the criteria as not understanding the task—I give him concrete materials which may help him understand, and often with guiding questions I can bring him to understand. (teacher AQ.II.1)

Some teachers gave examples of the type of remedial work they would undertake with the students:

If I discovered a student who forgot some of the conditions in the problem, then I could teach him how to mark them and then go back and check that he fulfilled all the conditions. (teacher BD.II.1; p.14)

For example, after using the *Multiples of 2* task, I brought the teachers other charts to explore, to look for additional generalizations and to delve deeper into their investigations of numbers. (teacher BE.II.1)

...to teach the advantages of systematic work, to teach how to carry out a plan, to teach how to make sure you have used all the conditions. (teacher BF.II.1; p.15)

If the problem is the understanding of the task, then I would work on reading comprehension, and if the problem is mistakes in computation, then I would check what was causing the problems.

(teacher BG.II.1; p.17)

6.2.3 The Third National Course on Assessment

As previously mentioned, the teachers who attended this course filled out a group questionnaire during the fourth workshop, relating to their field work. The final question was concerned with schools that teach in innovative ways. There were six groups of teacher-leaders, each with four or five members and each teacher working with between seven to 14 schools. Most of the teacher-leaders wrote about innovative teaching methods, including tasks with concrete materials, authentic “real-life” tasks, weekly puzzles, mathematical discussions and “assignments appropriate to the student’s level”. The reports revealed a great variety within the groups regarding their perceptions of innovation. For some teachers the idea of puzzle problems was an innovation, whereas others were already talking about student writing and journals. This highlighted the importance of exchanging ideas between teachers and brought home the crucial need to continue to allow time in our assessment courses for the

teacher-leaders to learn from each other. By discussing their own experiences they could aid and influence each other. Learning that not all the answers should come from the “experts”, but that they could come up with solutions themselves became a critical factor in their professional development.

During the second meeting of the course the teacher-leaders had been introduced to using criteria for evaluating the investigative tasks of students. The teachers were then given a homework assignment: to grade a task first by using the marking guide and then by using the criteria, and to write their conclusions from this exercise. In subsequent meetings, the teacher-leaders were given other tasks to grade by means of the criteria. The two ways of analyzing student work were compared and participants had to consider what could be learned from “conventional grading”, as opposed to what could be learned from “grading by means of criteria”.

At the sixth workshop, the teachers were given the following questionnaire:

<p style="text-align: center;">Questionnaire - 21.11.96</p> <p style="text-align: center;"><i>Add-a-digit: a task involving estimation</i></p> <p>1. How does this task differ from the tasks that the students usually carry out?</p> <p>2. In your experience, does grading tasks through the use of criteria provide you with information is not obtained through conventional grading? Yes / No</p> <p> If so, describe the information it provides.</p> <p>3. Additional comments.</p>

Twenty-seven teachers (out of the 32 registered in the course) responded to this questionnaire. The responses of two of these teachers are pertinent to the question of the use of criteria in assessment:

Checking by criteria gives a clear picture of factors which are important to us in relation to mathematics—for example, understanding concepts, understanding instructions ... in relation to the whole task and not just on individual items. (teacher CD.III.1; p.39)

Using criteria for checking gives us additional information. In this the student needs to explain, and this helps us to understand the process he uses, where he gets stuck, enabling us to help move him forward.

Conventional grading gives us information about what he knows, and whether he got the final solution or not. What should interest us as teachers is what important things aren't clear to him.

(teacher CO.III.1)

Four teacher-leaders reflected on the way in which the use of criteria in assessment could help improve teaching. By making it clear to the teachers what the crucial areas for evaluation are, the criteria may offer new directions for teaching. This is evident in their responses below:

We need to encourage the teachers to check with criteria—this will help their teaching, it will improve their testing skills and ultimately, it will improve student achievement. (teacher BO.III.1; p.20)

Checking by criteria enables the teachers to follow the students' thinking processes and forces the teachers to teach in a different manner, not just solving exercises. (teacher BX.III.1)

I very much enjoyed using the tasks with the children. I sat with the teachers and we went over the students' work. The teacher discovered certain things about her students, got a different picture of her class, and from here we can start changing the methods of teaching. No longer only closed questions, but tasks which require thinking, explanations and drawing conclusions. (teacher CB.III.1)

This task and the teachers' exposure to it opens a door for rethinking our teaching and our testing methods. Assessment which requires the student to use a variety of thinking skills. The quality of the task/test versus the qualities required in learning—and the thinking required of the teacher as a trainer of processes and a creator of tests.

(teacher BT.III.1)

At the eighth workshop, the teacher-leaders were required to fill out a questionnaire. As in the Second National Course on Assessment, the questionnaire contained the question, "*Did checking by criteria help you plan your teaching?*" Since, as previously

discussed, the importance of assessment lies in its impact on teaching, the responses to this question were of particular interest.

The teachers' responses to this question were broken down into three main categories:

1. Student assessment;
2. Types of intervention;
3. Changes in teaching style.

1. Student Assessment

Approximately one quarter of the teachers (6 out of 25) wrote about the impact of the use of criteria on their assessment of their students. Some of these responses are given below:

I now know how to differentiate between the different levels of the pupils. The criteria directed me to pay attention to verbalization and justification, and to try to understand what the student meant, and thus to determine his level of thinking. (teacher BQ.III.2)

Assessment has turned into a **real working tool** for continuing treatment of the pupil. The criteria enabled me to see the pupil from different angles, and to make further assessments. (teacher BS.III.2)

If until now, checking was to determine which exercise the pupil had trouble with, checking with criteria enables us to get a full picture of his abilities. (teacher BT.III.2)

Beforehand, we used to check the final answers and now we check his way of thinking. In looking at his way of thinking, we can decide how to progress and work with him so that he can understand and apply his knowledge in further tasks. (teacher CH.III.2)

2. Types of Intervention

Nine teachers gave specific examples of the areas in which they needed to plan special intervention to overcome problems that were uncovered through assessment by use of criteria. Five of them mentioned verbalization as one of these areas:

The criteria directed me to areas in which the child needed more work—for example, verbalization, to write about the way you worked. The criterion of following instructions trained me to give students the instructions in various ways, when these raised the level of difficulty of the task. (teacher BP.III.2; p.25)

If I discover that students are not verbalizing their solutions we will work together in class on verbalization skills and have many more mathematical discussions. (teacher BV.III.2; p.32)

If a large number of students had trouble with verbalization, I would plan activities which require verbalizing and emphasize that. (teacher CE.III.2)

After using criteria for checking, I have been looking for ways to improve the mathematical language of the students and their ability to express themselves. (teacher CD.III.2; p.41)

If, for example, I discovered difficulties in understanding concepts: such as approximate answer, then I need to check individually with each student if the problem lies in written verbalization or in a lack of understanding of the concept. (teacher CK.III.2)

Six teachers wrote about the need for intervention related to mathematical thinking (e.g., generalizations and justifications).

Evaluating by criteria enables me to determine which areas need to be emphasized in class and with which students. For example, if I noticed difficulties with formulating generalizations or justifications, I would need to emphasize tasks which require these. (teacher CL.III.2; p.50)

For example, if I discovered students who are not managing to reach the right conclusions, then I would give them an easier task and meanwhile try to work with them and little by little bring them to higher levels. (teacher BV.III.2; p.32)

For example, the patterns in the *Table Arrangements* task gave me tools to teach the teacher how to work with the pupils on finding patterns. (teacher BY.III.2)

Checking with criteria enables me to know what things to stress with the learner and how to teach. For example, in *Table Arrangements*, the children arranged tables with manipulatives, but didn't know how to calculate how many tables without them (by using tests of divisibility). Thus I emphasize making generalizations and understanding the relationship between area and perimeter. (teacher CB.III.2)

Similarly, ways to improve their mathematical thinking—to make it more organized, more systematic—and how to order their prior knowledge. (teacher CD.III.2; p.41)

After checking the *Population* task, I concluded that I needed sorting activities in mathematics, as well as activities on determining sorting categories. I also decided that I needed to work on understanding. (teacher CF.III.2; p.44)

3. Changes in Teaching Style

Almost half of the teacher-leaders (12 out of 25) wrote of the ways in which the use of criteria for assessment purposes had influenced their teaching.

In every math topic I try to give expression to as many criteria as I can. In planning each topic I arrange that each child can have an opportunity to show his creativity, his verbalization, his ability to generalize, so that assessment will reflect the process of teaching. (teacher BO.III.2; p.22)

Checking with criteria gives me feedback on my work as a teacher, what needs to be changed. For example, if the students have trouble with verbalization then I conclude that I need to train them to express themselves orally and in writing about the processes they use in working. (teacher BR.III.2)

It gives a very clear picture of what I need to emphasize in teaching, what is missing in my math lessons. The thing that stands out is the children's inability to justify. (teacher BW.III.2)

Using the criteria taught me the value of mathematical discussions, to work on one's ability to describe one's math doings, to be exact in verbalizing the processes and the conclusions. (teacher CC.III.2; p.38)

I try to bring my class tasks which require investigations or discovering rules. For example, "If the area of different rectangles is equal, are their perimeters equal?" I asked the students to create different rectangles, to find their perimeters, to create a table, and to find interesting relationships.

As a result of today's workshop, I will try an additional task: "How can I know if I've found all the rectangles—is there any rule?"

(teacher CI.III.2)

In another question, the teacher-leaders were required to describe what had changed in their classes, or/and in their supervision as a consequence of the course. Although this was a very open question, it was clearly influenced by the other questions. Almost three quarters of the teachers (18 out of 25) acknowledged that their exposure to the course had led to changes in their classes. For discussion of such changes, see Section 6.1, "The Teachers' View of Mathematics."

In conclusion, it can be seen that the teachers' use of investigative tasks had influenced their teaching. They were discussing teaching issues, like "inability to justify" and "discovering rules", issues which never arose before.

6.3 The Teachers' View of Student Assessment

The third question that this research addressed was:

Will the teachers' use of mathematical investigative tasks for assessment purposes influence the way they assess their students, and if so, in what ways?

In this section I include the teacher-leaders' attitudes towards assessment and their feelings about the use of criteria for assessment purposes (see Chapter 3 for the criteria chosen and the way they were introduced to the teachers).

6.3.1 The First National Course on Assessment

As mentioned in “The Teachers’ View of Mathematics”, the questionnaire filled out by teachers at the beginning of the course (during the second workshop) included the question: “*What were the results of your use of the tasks with the students?*” This, together with the subsequent question, “*Do you think these tasks can be used for assessment?*”, provided information regarding the teachers’ feelings about and attitudes towards assessment of students. In answer to the first of these questions, two teachers wrote about the difference in student performance depending upon the extent of their knowledge of the mathematics involved in the task:

One can see the difference between their performance on a task which deals with material which has already been learned in class and which has already been “digested”) and a task which deals with material which they have just begun learning, or haven’t learned at all.

(teacher AC.I.1)

There was a significant difference between the results of a task given immediately after the subject had been studied and a task on a subject which hadn’t been studied or whose study had just begun. (teacher AB.I.1)

Ten teachers referred to assessment aspects of the tasks. Some of these commented on the possible discrepancy between teachers’ assessments of students and their achievements on the tasks. Teacher AN, for example, noted that the results did not always agree with her evaluation of the students, and teacher AG (I.1; p.3) also mentioned the lack of correlation between teachers’ evaluation of their students and their success on the tasks. Teacher AW pointed out that some students whom she had considered weak had surprised her with their correct answers. One teacher (AX) disagreed with this evaluation, claiming that the students’ success on such tests was consistent with their achievements on regular tests. Other teachers wrote that assessment by means of criteria enabled them to gain “true assessment information”. For example:

The results of the tasks illustrate clearly what students know well and what I need to work on more, even though the form of the tasks was unfamiliar to them. The grading of the questions enabled me to gain a better understanding of where each child's difficulties lie.

(teacher AW.I.1)

There were however, teachers who were convinced that the tasks placed unreasonable demands on the pupils:

The questions were very hard for the students. Even after explanations, not everyone knew what was expected of him.

For most of the open questions where they were supposed to write an explanation, they didn't answer, and we had to sit with each student individually and encourage him to recount what he had done and why.

(teacher AR.I.1)

As previously mentioned, the subsequent question asked whether the teacher-leader thought that these tasks could be used for assessment. Four teachers did not respond to this question, and of the remaining 20 teachers, 18 responded in the positive. Only two responded negatively, and even these answers were not categorical. One of these teachers (AU) suggests that the tasks would be more beneficial for developing mathematical thinking than for assessment purposes, and notes that their use required too many explanations. The other one (teacher AA) raised doubts about using the tasks for assessment, pointing out that some of the children had worked together on the tasks. A number of issues were raised by the teacher-leaders with regard to the use of investigative tasks for assessment:

1. Necessary changes in teaching methods;
2. The use of investigative tasks to supplement conventional testing methods, rather than replace them;
3. Miscellaneous comments.

1. Necessary Changes in Teaching Methods

Of the 18 teachers who felt that the investigative tasks could be used for assessment, 12 wrote about the need to make appropriate changes in teaching methods to accommodate them. For these comments, see above, "The Way Teachers Teach."

2. The Use of Investigative Tasks to Supplement Conventional Testing Methods

Three teachers mentioned that the tasks were useful for assessment, but suggested that they should be used as a supplement to—rather than as a replacement of—existing testing methods.

These tasks are good for assessment on condition that they are used in addition to (conventional) tests and not instead of tests, because for students who perform on a level of knowledge and skills these tasks are very hard. (teacher AQ.I.1)

Comments such as that of teacher AQ are a matter of concern for us, as they reflect an underlying belief that "thinking skills" are for the brighter students alone.

3. Miscellaneous Comments

Two teachers mentioned additional issues in connection with the evaluation of student work. One (teacher AG.I.1; p.3) pointed out that it takes a long time to check student performance using the investigative tasks, a factor that may constitute a potential problem for some teachers. Another teacher offered a comment which we found to be quite helpful in the planning of future tasks:

You should request explanations on the easier questions as well—and not just on the hard ones at the end. (teacher AI.I.1)

In the ensuing group discussion, several teachers mentioned assessment "benefits" from the tasks. That is, they learned additional information about their students, although this knowledge was often discouraging to the teachers.

In the teacher-leader's class (where she was the teacher herself) she was discouraged by the phenomenon of students who "know, but don't know how to explain". She was

taken by surprise by some children's success and by the fact that one child did not succeed, although she had thought he would.

In a class on a kibbutz, the teacher was discouraged by tasks that required arithmetical and geometrical skills which had not yet been acquired. They managed to solve those questions that could be resolved in a standard manner, but those that required higher-level thinking were difficult for them and led to mistakes.

A differentiation must be drawn between classes that have worked on such types of tasks in the past, and those which have no prior experience with them. One class, in which the teacher is accustomed to working on "thinking skills", had exceptionally outstanding results.

Professor Pearla Nesher, who attended this discussion, noted that the advantage inherent in using investigative tasks for assessment purposes is that the teacher learns more about the students. These tasks provide additional assessment information and can inform the teacher about whether the child gets into difficulty on the simple examples or on harder ones. The levels of the tasks provide guidance on what to work on with the child. However, she pointed out, if the teachers use the tasks as learning material, they can no longer be used for assessment. She gave an example of the use of an open question:

In second grade there are 30 students. There are more boys than girls.

How many boys and how many girls are there in this second grade?

Possible student answers:

Yossi: There are 20 boys and 10 girls.

Ya'acov: There 20 boys and 10 girls and also 16 boys and 14 girls.

Ruth: There are many answers and I know how to find all of them.

According to Vygotskii, each child has a spread of knowledge where he is situated, and thus one needs to talk with each one in his spread and bring him up higher. In this specific example, we need to ask Yossi whether there are any more possible answers. We must ask Ya'acov whether there are any other

possible answers. And Ruth, how does she know that she can find all the answers?

Dr. Alex Friedlander contributed to the discussion by pointing out that the learning and assessment processes should not be detached from one another. A large part of assessment in fact takes place in the course of learning; the learning process should thus be organized in such a way as to enable the teacher to constantly observe and assess. Most assessment needs to be **formative and not summative** (Section 2.3).

There is, however, an inherent problem in creating similar investigative tasks for assessment: if one wants to test real problem-solving skills, then the tasks provided must involve hitherto unknown, unconventional and new problems. This, of course, is an unrealistic expectation on classroom teachers.

6.3.2 The Second National Course on Assessment

Since the Second National Course on Assessment was the first one in which additional criteria for assessment, beyond the traditional methods, were introduced (see Chapter 3 for details), it was important to me to gauge the teachers' response to them. Of the 23 teacher-leaders present at the sixth workshop, in which the questionnaire (see below) was filled out, the majority had used criteria for grading two or three tasks.

Questionnaire for Math Leaders

Lately we have added criteria for evaluating the tasks, in addition to the answer guide which is attached to each task.

We would like to ask a number of questions which relate to the use of these criteria.

1. How many different tasks have you checked using these criteria? Which tasks?
2. What do you think, in general, about the use of criteria for checking tasks?
3. Did the use of criteria cause:

	very much	some	no difference	definitely not
extending the checking process?				
deepening understanding of the pupil's performance?				
emphasizing the difference between students?				
acquiring more detailed information about the student's level?				

4. Which criteria are not clear? That is, which are hard to make decisions about them?
5. Did checking by criteria help you plan your teaching? If yes, give an example.

Question 2: On the Use of Criteria for Checking Tasks

In general, the teachers' responses to the use of criteria for grading tasks were positive. Although six teachers noted difficulties adjusting to this system, all of them

agreed that the end result was worth the effort, noting that it “helps focus” (teacher BB) or “helps give direction” (teacher AM). The responses can be divided into two areas: 1) the effect of the use of assessment criteria on instruction; and 2) the effect of the use of criteria on student assessment.

1. Effect on Instruction

Five teachers wrote that using criteria for checking tasks helped focus their instruction. One teacher (BA.II.1; p.11), for example, wrote that “using criteria helps and focuses, adds a framework”.

2. Effect on Student Assessment

The majority of the teachers (15 out of 23) responded that the criteria improved their assessment of the students’ abilities. They noted that the criteria provided a framework for evaluating the students’ work by adding new dimensions to the teacher’s assessment and generating a better understanding of the students’ thought processes. Following are a selection of the teachers’ comments:

. . . possible to check if (the) student understands the task, using makes it possible to check different areas of ability (teacher AY.II.1; p.9)

Using criteria deepens the knowledge of the student, differentiates between (the) student’s lack of knowledge in computation and his way of thinking (teacher BF.II.1; p.15)

I relate very positively to the use of criteria. They enable me to pay attention to elements that previously I didn’t grade. (teacher BG.II.1; p.16)

I think that the use of criteria is important, because then one doesn’t just relate to computations—it helps us gain a better understanding of the student’s thinking, through his mistakes, why he performed as he did. (teacher BJ.II.1; p.18)

Very helpful—helps focus the assessment beyond looking at each item individually. [refers to looking at the test in a holistic manner] (teacher BM.II.1)

- a) The criteria help the teacher understand what the child needs to understand and know;
- b) The emphasis moves from product to process;
- c) The teacher gets information on the student’s way of thinking and performing and how to work with him. (teacher AL.II.1; p.6)

It is very agreeable to check in this way—makes equitable assessment possible. This assessment is ‘real’ and not based on the ‘feelings’ of the teacher or on being too ‘strict’ or too ‘lenient’.” (teacher BB.II.1)

Question 3: Grading Various Effects of the Use of Criteria

Participants were required to evaluate a number of parameters pertaining to the effects of the use of criteria in assessment. They graded their responses on a scale of 1 to 4 (from “very much” to “definitely not”). A statistical breakdown of their responses follows:

Question 3. Did the use of criteria cause:

	Number of Teachers			
	Very much	Some	No difference	Definitely not
Extending the checking process?	12 (52%)	4 (17%)	5 (22%)	2 (9%)
Deepening understanding of the pupil’s performance?	20 (87%)	3 (13%)	0 (0%)	0 (0%)
Emphasizing the difference between students?	15 (65%)	6 (26%)	2 (9%)	0 (0%)
Acquiring more detailed information about the student’s level?	14 (61%)	7 (30%)	2 (9%)	0 (0%)

As is evident from their responses, the majority of teachers felt that using the criteria provided them with a deeper understanding of a pupil's performance. In addition, they felt that using the criteria helped to highlight the difference between students and provided them with more detailed information about a student's level. Half of the teachers (12 out of 23) felt that the use of criteria significantly lengthened the checking process, whereas almost a third (7 out of 23) responded either that it made no difference to the length or that it definitely did not extend the checking process.

Question 4: Unclear Criteria

The teachers were asked to determine which of the criteria were unclear to them, based upon a consideration of how difficult it was to make a decision about the student's level on the basis of these criteria. A breakdown of their responses to specific criteria is presented below:

Number of teachers who consider each criterion to be unclear

Criterion	Number of teachers
Understanding the task's instructions, the situation, the givens and the constraints	10 (43%)
Understanding the concepts	5 (22%)
Mathematical thinking, generalizations, justifications	10 (43%)
Computation	1 (4%)
Communication, explanations, presentation of the solution	3 (13%)
Creativity	5 (22%)

As is evident from the breakdown of responses, almost half of the teachers (10 out of 23) felt that it was hard to evaluate the student's understanding of the task. The reason for this became evident in the course of our discussions. The teachers assumed that the criteria were intended to assess the students' written work, whereas the students' "understanding of the task" often had to be assessed on the basis of the questions they asked while they were working. The idea that what was being assessed was the totality of the student's performance—and that questions raised in the course of his ongoing work were as valid a component of the assessment as the final paper that he hands in—was a difficult concept for some teachers to accept. The other criterion deemed "unclear" or "hard to evaluate" by a significant proportion of the teachers (10 out of 23) was "mathematical reasoning". Since this criterion included areas that were new for some of the teachers themselves, it was difficult for them to gauge the presence of "mathematical reasoning" in the work of their students.

6.3.3 The Third National Course on Assessment

The teacher-leaders came to the course with some preconceived notions—some based on other courses in which they had participated and some from hearing about our investigative tasks from colleagues. At the second workshop, we held a full group discussion about alternative assessment, following their work in smaller groups. The ideas about the components of alternative assessment which arose in the course of the discussion can be separated into three areas: 1) what type of tasks to use, 2) the need for correlation between teaching and assessment, and 3) crucial element needed in the learning process. The sub-areas are listed here:

- 1) Multi-stage problems, Scaffolded tasks, Challenge problems, Authentic problems, Room for creativity;
- 2) Correlation between the subjects undergoing evaluation and the subject being studied, How to teach and what to teach—learning about the child's thinking

processes, Learning based on those questions, rather than the composition of tests after the teaching of the subject;

- 3) Students' reflections, Group work, Use of manipulatives, Mathematical discussions and discourse.

As is evident from these suggestions, these teacher-leaders formed a heterogeneous group, with some possessing prior knowledge about reform-based teaching while others were exposed to these ideas for the first time.

At the fourth workshop, the teachers compared conventional grading with grading by means of criteria and responded to two questions:

What did you learn from the conventional grading?

- 1.
- 2.
- 3.

What did you learn from grading by means of criteria?

- 1.
- 2.
- 3.

The responses were presented to the group and discussed. What emerged from the discussion was that each type of grading held its own advantages for the teachers. While "conventional grading" presented a clear picture of each student's performance on every question, and together provided a class profile for each question, "grading with criteria" offered a view of broader domains and included areas that had not previously been considered, such as communication and mathematical reasoning. The general feeling among the participants was that conventional grading was more objective, while grading with criteria was more subjective. Despite this drawback, it was clear to them that grading by means of criteria helped pinpoint areas that require specific treatment.

One group of teacher-leaders summed up the advantages of grading with criteria as follows:

1. One can learn in which sort of questions the students made mistakes or knew the material.
2. One can sort questions according to their level of difficulty and the level of thinking required.
3. Accordingly, one can learn about the child's thought processes.
4. By means of (2) and (3), one can develop ways of treatment for each individual child.

The questionnaire to which the teachers responded at the sixth workshop provided information about the teachers' reactions to using criteria for evaluating tasks. Of the 27 teachers present on the day the questionnaire was distributed, 26 responded positively to the use of criteria for assessment purposes. Even the sole teacher who expressed negative feelings felt that she had learned something about her students—but their inability to explain their thought processes bothered her so much that she blamed the type of grading, rather than the students' performances. Her reaction is presented below:

On this task, there isn't much room for qualitative checking—only in question 4 which requires the children to explain. And even there the wording is so problematic and the writing errors are so many that none of the fourth grade answers could be said to be really correct. Thus the picture we got was catastrophic—what should we do? (teacher CN.III.1)

The teachers' responses to the use of criteria for assessment purposes deal with a number of aspects: 1) the new dimensions that such grading methods provide to the assessment picture; 2) assessment of the mathematical ability of students; and 3) the drawbacks of conventional grading methods.

1. Providing New Dimensions to Student Assessment

Eighteen teachers (two-thirds of the 27 who responded) mentioned explicitly that using the criteria gave them a more in-depth picture of their students' performances—

the criteria provided new dimensions to the teachers' understanding of the students' strengths and weaknesses.

Using the criteria give further information. Criteria tell us about communication and explanations and what work to do with whom.
(teacher CM.III.1)

Using criteria helps us see if the student understood the task, and if he worked on the basis of understanding or worked purely by trial and error. It gives us a broader picture than just his computations, and helps us understand if the student really understood or just happened to succeed. (teacher CL.III.1; p.48)

Using criteria enables us to check how the student got his answers, instead of just knowing how many correct answers he had.
(teacher CH.III.1)

Certainly, using criteria gives us different information. It enables us to point out where each individual student is having difficulty. How to help him from the point where it's hard for him. Which subjects need more emphasis with the students. (teacher CG.III.1)

Conventional grading generally gives us information about proficiency in mathematical skills and techniques (how many computations were correct). In using criteria for checking we get a profile of the student on different dimensions. Thus, we can give specific advice on how to improve his performance. (This advice is helpful to students and to teachers.) It also helps us observe the student's thought processes.
(teacher CP.III.1)

Checking by criteria forces me to pay attention to the different criteria, it makes it possible to assess a student from different aspects, and makes it possible to get information about process which I wasn't previously aware of. (teacher CI.III.1)

One teacher explicitly compared the two styles of assessment—conventional grading and assessment by means of criteria—and reached the conclusion that the two styles complement each other:

The criteria enable the student to demonstrate his different abilities, not just facts and basic skills. I think that the two ways of assessing

complement each other, and that checking by criteria shouldn't come in place of conventional grading. (teacher CQ.III.1)

2. The Mathematical Ability of Students

Eighteen teachers wrote that the criteria provide an understanding of the students' thinking and mathematical ability. They felt that this information would be helpful in planning future lessons to deal with problem areas.

The criteria tell us exactly on which points each pupil needs work.
(teacher BQ.III.1)

Conventional grading tells us on which questions the students succeeded/failed, and how many students succeeded, whereas using criteria provides details on what is needed to help each student improve his performance. (teacher BP.III.1; p.23)

There were a few children whose conventional grades were low in computation, but who made correct conclusions in their explanations. Checking by criteria helped to understand thinking processes throughout the whole task, and not just on single items.
(teacher CD.III.1; p.39)

Rather than conventional grading which gives information about numbers—how many questions were answered correctly/incorrectly, and what were the final grades—and information about the situation of the class, checking by criteria enables us to find the reasons for the mistakes. (teacher BT.III.1)

3. The Drawbacks of Conventional Grading Methods

Five teachers wrote about the drawbacks of conventional grading:

Conventional grading only tells us how many answers were correct.
(teacher BS.III.1)

Conventional grading gives us a grade for the student, but doesn't point out problems areas or give us information about why these are problem areas. The criteria give us indications of the student's level of thinking and way of thinking. (teacher BZ.III.1)

Conventional grading gives us information about what he knows and whether he got the final solution or not. What should interest us as teachers however, is what important things aren't clear to him.
(teacher CO.III.1)

Checking by criteria looks at different things than conventional grading. It helps us to understand where the problems begin.
(teacher BX.III.1)

Conventional grading gives the student a grade and rates him on the class profile. Checking by criteria places the emphasis on what needs to be improved in order to get a better grade, and on the level of knowledge of the student. (teacher CC.III.1; p.36)

At the eighth workshop the teacher-leaders answered the questionnaire that I used in the sixth workshop of the Second National Course on Assessment (see above, "Questionnaire for Math Leaders"), with the addition of another question:

What has changed in your class/in your supervision as a result of this course? Explain and provide examples (different mathematical content, difference in styles of teaching, difference in ways of supervising, etc.)

Question 2: On the Use of Criteria for Checking Tasks

Although seven teachers still felt that there were problems in using criteria for assessing tasks, none rejected them altogether. They all felt that the use of criteria contributed to their understanding of the students' ways of thinking.

I think there is a need for qualitative criteria, but there is still confusion about some of them. The clear ones are: understanding the task, understanding the concepts, mathematical thinking and computation. It's very hard for me to evaluate creativity.
(teacher BU.III.2; p.28)

I haven't used criteria grading in classes, but just under experimental conditions—but for teachers who have to give grades on report cards, it seems that there will be problems. (teacher BV.III.2; p.31)

The criteria are good, the "ratings" are hard. The criteria help greatly in knowing what to emphasize and enable making definite conclusions for work with children. (teacher BW.III.2)

The teachers were so accustomed to conventional grading, where the exercise is perceived as either right or wrong, that the task of assessing a written piece of work seemed to them quite subjective, one with which they had great difficulty. This is a skill that takes time to acquire. The other problem that they encountered was the discrepancy between the use of criteria for assessing tasks and the necessity of filling in conventional report cards. This problem frequently arose in discussions, and is gradually being alleviated due to changes that are being implemented in the system of report cards (see separate section below). The teachers' responses broadly relate to two main spheres: 1) the effect of the use on assessment criteria on instruction; and 2) their use for student assessment.

1. Improved Instruction

Six teachers dealt with the way the use of criteria influenced their instruction.

This direction, in general, seems to me to be correct. Turning attention from computation to the way of thinking is critical, and leads to a way of teaching which seems to me to be better. (teacher CJ.III.2; p.46)

the pupil learned to relate differently to the subject when he received compliments which he was previously not used to getting. Thus we all benefited. (teacher CC.III.2; p.37)

2. Better Assessment of Students

The majority of the teachers (22 out of 25) responded that the criteria helped them obtain more information on their students and contributed new dimensions to this information.

The use of different criteria gave us the ability to relate in another way to each pupil. To differentiate between the pupils and to discover new faces of the pupil (that we hadn't seen in frontal teaching).

(teacher CC.III.2; p.37)

The use of criteria adds dimensions which weren't visible in conventional grading, and they are important—for example, creativity, the ability to verbalize, the ability to generalize, etc.

(teacher CBD.III.2; p.40)

The criteria give the teacher a different way of assessing the pupil, additional information about the pupil's knowledge or about what he doesn't know. (teacher CE.III.2)

The criteria give the teacher a fuller picture of the understandings and performance of the pupils. (teacher CF.III.2; p.43)

Several teachers noted how much verbalization contributed to an understanding of the extent of the pupil's knowledge.

Verbalization makes transparent the learner's thinking and points out mistakes in thinking processes. (teacher BO.III.2; p.21)

The benefit of knowing more about the pupil is translated into being better able to plan follow-up work. That is, the more I understand my pupils, the better I can plan my lessons.

The criteria help very much in knowing what to emphasize and enable drawing definite conclusions for work with children.

(teacher BW.III.2)

The criteria focused me in evaluating the tasks, and enabled me to make a more accurate and exact check of the task, and thus I can know how to plan treatment for the pupil. (teacher CG.III.2)

The criteria help us—teachers and teacher-leaders—teach more and more about the thinking of the pupil, get a more exact profile of the pupil, understand what the pupil can do (and not what he can not do!) and begin from his positive side. The criteria give us information about what to treat, who to treat, and how—that is, they help us plan our teaching. (teacher CM.III.2)

Question 3: Grading Various Effects of the Use of Criteria

Participants were required to evaluate a number of parameters pertaining to the effects of the use of criteria in assessment. They graded their responses on a scale of 1 to 4 (from “very much” to “definitely not”). A statistical breakdown of their responses is given below.

Question 3. Did the use of criteria cause:

	Number of Teachers			
	Very much	Some	No difference	Definitely not
Extending the checking process?	16 (64%)	3 (12%)	1 (4%)	3 (12%)
Deepening understanding of the pupil's performance?	25 (100%)	0 (0%)	0 (0%)	0 (0%)
Emphasizing the difference between students?	23 (92%)	1 (4%)	1 (4%)	0 (0%)
Acquiring more detailed information about the student's level?	25 (100%)	0 (0%)	0 (0%)	0 (0%)

As is evident from the responses, all the teacher-leaders felt that the use of criteria generated a deeper understanding of the pupil's performance and that they acquired more detailed information about the student's level. Almost all of the participants (23 out of 25) felt that the criteria helped to differentiate between the students. Almost two-thirds of the teachers (16 out of 25) felt that the use of criteria lengthened the checking process significantly.

A comparison of the teachers' responses to this question with those of the teachers in the Second Assessment Course indicates that they were feeling much more positive about using criteria at this later stage.

Question 4: Unclear Criteria

The teachers were asked to determine which of the criteria were unclear to them, based upon a consideration of how difficult it was to make a decision about the student's level on the basis of these criteria. A breakdown of their responses to specific criteria is presented below:

Number of Teachers who Feel Each Criteria is Unclear

Criterion	Number of teachers
Understanding the task instructions, the situation, the givens and the constraints	0 (0%)
Understanding the concepts	2 (8%)
Mathematical thinking, generalizations, justifications	12 (48%)
Computation	2 (8%)
Communication, explanations, presentation of the solution	15 (60%)
Creativity	24 (96%)

The criterion with which the teacher-leaders had most difficulty was creativity. Their problems were two-fold: evaluating creativity and deciding whether a specific factor was evidence of creativity. Again, the problems stem from what they perceived of as a lack of objectivity and the "newness" of their goal. For most of the teacher-leaders, the idea that creativity is something to be encouraged in a mathematics lesson was a completely innovative concept. Mathematics had always been a matter of learning how

to do something in the most efficient manner and obtaining the correct answer. Encouraging divergent thinking was a new goal for the teachers.

Approximately half of the teachers wrote that communication and/or mathematical thinking were problematic criteria. Like creativity, these are both relatively new goals for elementary-school mathematics. In addition to the difficulty of evaluating students' work in these realms, the main problem for the teachers was knowing how to teach these skills. Over and over again, in the workshops, we talked about what it meant to get students thinking in mathematics. We discussed ways in which a teacher could take almost any exercise and, by asking additional questions, turn it into an exercise which required higher-order thinking. We continuously discussed the importance of communication and the validity of assessing communication in a mathematics lesson.

Question 6: Ensuing Changes

In this question the teacher-leaders were asked to describe what changes had taken place in their classes and in their supervision as a result of the course. This was a very open question, but the teachers' answers were certainly influenced by the other questions of the questionnaire.

Assessment

Two-thirds of the teacher-leaders (16 out of 25) wrote that they used new types of assessment and changed their tests:

Because of the course and the assessment tasks, I have begun to assess the different pupils with different tools and not only with the standard test. (teacher BP.III.2; p.25)

Because of this course I have begun to learn to assess differently. I've started to relate to the process and not the final answer, to vary the thinking, to emphasize the creativity of the students.

In tests, I've been opening tasks and try to vary their levels.

(teacher BQ.III.2)

I am still working on myself and the teachers on the need for alternative assessment in addition to conventional assessment. (teacher BW.III.2)

In my work with teachers, I place more emphasis on ways of assessing the pupil and connected with this, I emphasize planning for teaching in alternative ways. I emphasize things that I wasn't aware of previously. (teacher BY.III.2)

My way of assessing has changed to the use of criteria, and also to other assessment tools. The tests for checking knowledge that I "still" use have changed their format, and now include items which the pupil needs to justify, to explain why he thinks something or to reach a generalization. (teacher CF.III.2; p.44)

In my class and also in my work with teachers, I emphasize several aspects: ... building tests which are scaffolded (not just on one topic) which have not only one correct answer, which require the need to explain, to justify or to judge. (teacher BU.III.2; p.29)

One teacher sums it up, noting that they have included reflection as well:

In this course I received encouragement for other ways of ... assessment—process versus product. Assessment of mathematical topics plus assessment of thinking, discovering rules, etc. Computations—what causes the problems?

The course opened up for me many different ways to assess: quizzes, worksheets, performance assessment, dialogues, peer assessments, portfolios. All together, these can give a fuller picture of the pupil.

Assessing the pupil on his work: we've started requiring a page for reflection after a task. For example, what was the subject of the task? What did you learn that was new? What materials did you use? Did you have any difficulties?

The course for me marked the beginning of a different way of thinking about assessment. (teacher CA.III.2)

In conclusion, I can see that the teachers' views about assessment have widened. It has become a much broader issue, and no longer just a matter of grading short-answer questions on a conventional test.

6.4 Follow-up Work with Course Graduates

6.4.1 Ofek

As described in Chapter 3, the *Ofek* satellite courses on School-Based Assessment consisted of an interactive television broadcast and a workshop at each center. The workshops were run by local teacher-leaders who had planned the course together with the lecturer. The teacher-leaders of the *Ofek* courses were graduates of the national courses on assessment; thus, my prime interest was in how these leaders' continued work with the assessment tasks would influence them. In order to plan the workshops, these teacher-leaders were asked to reflect on their own experiences in the national courses on assessment and determine to what sort of information and activities the participating teachers should be exposed. They came up with the following list of topics:

1. The need for alternative assessment;
2. The use of tasks in the classroom;
3. The way the use of tasks influences teaching methods;
4. Verbalization;
5. The way to report results;
6. Reflection;
7. Talking in a different way about students.

In order to achieve these goals, they devised the following list of activities:

1. Experience with the tasks;
2. Analyzing the students' work;
3. Analyzing the tasks;
4. Comparing a task with a conventional test;
5. Changing a page in the textbook to be in line with the new goals;
6. Writing tasks (time permitting).

The following protocol is the course we formulated:

The Outline for the First *Ofek* Course

	TV Broadcast	Workshop
First meeting	<ol style="list-style-type: none"> 1. What are the goals of teaching elementary school mathematics today? (collect ideas) 2. An explanation about reform mathematics. 3. How do you assess students today? (collect ideas) 4. Demonstrate the conflict between their goals and their assessment methods. 5. Have them solve the <i>Getting Results</i> task. 6. Consider what a student needs to know in order to succeed on this task. What is special about this task? (collect ideas) 7. Discuss the ideas that arise, along with a presentation of the attributes of the tasks, including examples. 	<ol style="list-style-type: none"> 1. Solve the task <i>Trains</i>. 2. Analyze its attributes. 3. Check students' work using the answer key. 4. Discuss what conclusions can be drawn from their work and what additional teaching is required. 5. Solve the task <i>Polygons on a Geoboard</i>. 6. <u>Homework</u>: use <i>Polygons on a Geoboard</i> in a class, correct the students' work using the answer key, and write conclusions about what the students understood and what the follow-up lessons would need to be.
Second meeting	<ol style="list-style-type: none"> 1. Discuss their conclusions regarding their students' work on <i>Polygons on a Geoboard</i>. (collect ideas) 2. Their ideas for additional teaching on problems that arose, along with my own ideas. 3. Ten-minute video of exemplary geometry lesson using geoboards, with me as teacher. 4. Discussion on the lesson. 5. Task for participants: Find five numbers whose average is seventeen. 6. Present attributes of good questions. 	<ol style="list-style-type: none"> 1. Math investigation - the coin problem - including the follow-up mathematical discussion. 2. Textbook page to "open", i.e., make the questions more open, including talking about how to have follow-up mathematical discussions.

<p>Third meeting</p>	<ol style="list-style-type: none"> 1. What additional tools do we have for assessment, in addition to standard tests? 2. Have participants solve the <i>Bridges</i> task. 3. Video showing a pair of girls solving <i>Bridges</i>. 4. Discussion: what additional criteria would help us assess the girls' work. 5. Use of these additional criteria to assess another pair of girls solving <i>Bridges</i>. 	<ol style="list-style-type: none"> 1. Have participants solve the <i>Animals</i> task. 2. Discussion: what does this task assess? 3. Check students' work on this task using rubrics—the additional criterion presented in the broadcast. 4. Discuss what conclusions could be drawn from their work, and what additional teaching is required. 5. <u>Homework</u>: Observe two pairs of students solving <i>Getting Results</i>, check their work using rubrics and write conclusions.
<p>Fourth meeting</p>	<ol style="list-style-type: none"> 1. What does it mean to make a profile for a class? If we wanted to determine what a second-grade class knew about subtraction, what would we want to determine and what assessment tools could we use? (collect ideas) 2. The advantages to using varied ways of assessment. 3. Correct use of the Assessment Task Bank. 	<ol style="list-style-type: none"> 1. Plan how to assess a fourth- or fifth-grade class on their knowledge of quadrilaterals.

In October 1997 we met to plan the first part of the second *Ofek* course. The conclusions we derived from the first experience was to allocate more time for discussion by reducing the number of tasks the participants were required to solve. Thus, in the second course, the teachers solved fewer tasks, but analyzed each one to a greater extent than before.

The Plan for the Second *Ofek* Course – Semester One

	TV Broadcast	Workshop
First meeting	<ol style="list-style-type: none"> 1. What are the goals of teaching elementary school mathematics today? (collect ideas) 2. An explanation about reform mathematics. 3. How do you assess students today? (collect ideas) 4. Demonstrate the conflict between their goals and their assessment methods. 5. Have them solve the <i>Division by Three</i> task. 6. Consider what a student needs to know in order to succeed on this task. What is special about this task? (collect ideas) 7. Discuss the ideas that arise, along with a presentation of the attributes of the tasks, including examples. 	<ol style="list-style-type: none"> 1. Solve the task <i>Trains</i>. 2. Analyze its attributes. Discuss how this task differs from a conventional test. 3. Check students' work using the answer key. 4. Discuss what conclusions can be drawn from their work and what additional teaching is required, concentrating on verbalization. 6. <u>Homework</u>: Use <i>Trains</i> in a class, correct the students' work using the answer key, and write conclusions (if time permits, use the task 01/01/01 on Estimation as well).
Second meeting	<ol style="list-style-type: none"> 1. Discuss their conclusions regarding their students' work on <i>Trains</i>. (collect ideas) 2. Ten-minute video of exemplary geometry lesson using geoboards, with me as teacher. 3. Discussion on the lesson. 4. Task for participants: Find five numbers whose average is seventeen. 5. Present attributes of good questions. 	<ol style="list-style-type: none"> 1. Some investigative work using a geoboard. 2. Solve the task <i>Polygons on a Geoboard</i>. 3. Analyze its attributes. 4. Check students' work, using the answer key. 5. Discuss what conclusions can be drawn from their work, and what additional teaching is required 6. <u>Homework</u>: Use <i>Polygons on a Geoboard</i> in a class, correct the students' work using the answer key and write conclusions.

<p>Third meeting</p>	<ol style="list-style-type: none"> 1. Discuss their conclusions regarding their students' work on <i>Polygons on a Geoboard</i> (collect ideas). 2. What additional tools do we have for assessment, in addition to standard tests? 3. Have participants solve the <i>Bridges</i> task. 4. Observe a video showing a pair of girls solving <i>Bridges</i>. 5. Discussion: what additional criteria would help us assess the girls' work. 6. Use of these additional criteria to assess another pair of girls solving <i>Bridges</i>. 	<ol style="list-style-type: none"> 1. An investigation. <p>Or</p> <p>Check their students' work on <i>Polygons on a Geoboard</i> using additional criteria. In order to do this, the criteria first need to be adapted to this task.</p> <ol style="list-style-type: none"> 2. Textbook page to "open", i.e., make the questions more open, including discussion about how to have follow-up mathematical discussions.
<p>Fourth meeting</p>	<ol style="list-style-type: none"> 1. What does it mean to make a profile for a class? If we wanted to determine what a second-grade class knew about subtraction, what would we want to determine and what assessment tools could we use? (collect ideas) 2. The advantages to using varied ways of assessment. 3. Correct use of the Assessment Task Bank. 	<ol style="list-style-type: none"> 1. Plan how to assess understanding of a fraction. 2. What teaching goals can we assess with conventional tests and which goals require other assessment tools? 3. Analyze the fraction tasks from the Task Bank to determine what teaching goals each assesses.

In January 1998, we met to summarize the first part of the second *Ofek* course and to plan the second part. I used this meeting as an opportunity to hear about their experiences using the task bank in various parts of the country and to find out how teachers were relating to its use for assessment. I first asked the teachers to tell us about a school or a classroom where they had observed positive experiences taking place: the kind of teaching or assessment that we would like to have. The teacher-

leaders who responded had all participated in one of the national courses on assessment and who were major mathematical leaders in their local school systems.

Five teachers discussed explicitly the influence of the Assessment Task Bank on the types of questions now being asked in mathematics lessons. They referred primarily to open problems with more than one possible solution and the requirement that students explain their work. One teacher told us about a school whose pupils have a very low self-image that had begun weekly workshops with open problems. At the beginning these were challenges for the teachers themselves. However, after using such open problems with the students on a weekly basis, the teachers became interested in doing assessment with the task bank. Another teacher also recounted her experiences in a problem school in which they had begun working intensively on verbalization in mathematics lessons, requiring students on a weekly basis to explain their solution to one problem. This concentration of effort on verbalization is completely new for Israeli mathematics teachers—previously they would have argued that mathematics is about exercises, and not about writing!

Another teacher recounted her experiences in a school which had originally resisted reform mathematics until the day of the regional exam, which included a mathematics investigation task. That very day the school principal recanted and now accepts her ideas completely.

Yet another teacher discussed the problem of conflicting messages being sent out by the Ministry of Education. On the one hand, the use of the task bank for assessment was extolled, but on the other hand, a highly conventional test was being used to identify students whose performance was low, in order to determine the allocation of extra teaching hours to the school.

The consensus among us was that the second *Ofek* course was an improvement on the first: our roles had been clearer and we had a better idea of our goals and the means of attaining them, through the use of this new medium.

The teachers raised a number of problems: the heterogeneity of the participants, the absence of feedback regarding classroom experiences, the lack of space for good workshops and mechanical difficulties. These problems, however, were beyond our control. I told them how significant their feedback was to me, and emphasized the importance of their role: the success of the *Ofek* in-service workshops, to my mind, lays squarely on their shoulders.

One of the teacher-leaders asked whether, in order to be considered investigative, a task must lead to a generalization (see Chapter 1, Section 1.5 for a description of investigative tasks). My response was that what characterizes investigative tasks is that we do not have an algorithm for their solution. This question highlights the basic lack of knowledge with which we, teacher educators, must contend, evident even in these local “experts” who have attended a year-long course on the use of investigative tasks for alternative assessment.

I have found that there is a discrepancy between the stage I am at today and the ideas that the teachers need to hear about, which may relate to issues I was dealing with a number of years ago. For example, teachers need to be exposed to the difference between sorting questions according to cognitive level and sorting according to level of difficulty. Slower children can solve thinking problems (at a higher cognitive level) if the level of difficulty of the problem is lowered (e.g. by smaller numbers). This activity is one the teacher-leaders had all been exposed to many years earlier.

We considered other activities to carry out in the *Ofek* course, since the structure of the course prevents us from observing actual classes. One teacher-leader highlighted the importance of talking with the teachers about their classroom activities: doing investigative work and bringing the students to a higher level of understanding. The recurring problem of the lack of appropriate teaching materials arose—and it is exacerbated by the fact that the tasks of the Assessment Task Bank are not intended for daily teaching use.

With these goals in mind, we compiled a list of the issues to be dealt with in the second part of the *Ofek* course:

1. Cognitive levels vs. level of difficulty;
2. Treatment following a task-test—short term vs. long term (verbalization);
3. Concentrating on the big ideas of mathematics, with connections between lessons, rather than each small topic individually;
4. Teaching methods;
5. Comparing a conventional test with the use of a task-test;
6. Compiling a profile using criteria;
7. Familiarity with more assessment tasks;
8. Textbook and assessment tasks—their connections.

Incidentally, it became clear to me, in the course of this process, that the teacher-leaders involved had become a community, with common goals and a common body of knowledge—at times even developing our own language and coining new phrases.

This list formed the basis for the three broadcasts and three workshops that comprised the second semester of the second *Ofek* course, outlined below:

The Plan for the Second Ofek Course—Semester Two

	TV broadcast	Workshop
First meeting	<ol style="list-style-type: none"> 1. New government document on the organization of elementary school and the goals of mathematics teaching. 2. Mathematics needs to be centralized! How we can study mathematics and assess it according to these new directions and not lose the essence of the subject. 3. Solve the <i>Area 1, 2</i> task. 4. Analyze this task according to the new goals. 5. Solve and analyze a test from the textbook on area. Compare it to the <i>Area 1, 2</i> task. 6. Further explanation on the differences between verbalization, generalization and justification. 	<ol style="list-style-type: none"> 1. Solve a task requiring “writing”—<i>Is It Possible?</i> When does this writing constitute verbalization, when generalization, and when justification? 2. Solve the task <i>Tower of Cubes</i>, analyze it and discuss the use of criteria for assessment, as presented in the introduction to the Task-Test Bank. 3. <u>Homework</u>: Read the introduction to the Task-Test Bank for Grades Five and Six. Use the <i>Tower of Cubes</i> task in a class, and analyze the results like in the introduction.
Second meeting	<ol style="list-style-type: none"> 1. Discuss participants’ conclusions regarding their students’ work on <i>Tower of Cubes</i>. 2. The difference between short-term and long-term help for students. 3. Read together statistics of a class’s work on <i>Area 1, 2</i> and relate the results to the types of intervention required. 4. Talk about Bloom’s Taxonomy and compare cognitive levels of questions with their level of difficulty. 5. Analyze a textbook test on fractions and compare it to the <i>Difference and Inequalities</i> task. 	<ol style="list-style-type: none"> 1. Discuss student help for the <i>Tower of Cubes</i> task. 2. Add questions to the textbook test which require higher-order thinking. 3. Analyze a summary assignment from the textbook—cognitive levels of questions vs. their level of difficulty. 4. <u>Homework</u>: Read the article “A Constructivist Mode”.

Third meeting	<ol style="list-style-type: none"> 1. Compare and contrast the constructivist mode with the new government document. 2. How we can make profiles of students using the criteria, with examples from the <i>Area 1, 2</i> task? 3. The importance of creativity in mathematics and how one can teach for it. 4. What now? 	<ol style="list-style-type: none"> 1. Check students' work on the <i>Difference and Inequalities</i> task and make a profile using the criteria. 2. Read and discuss statistics from this task. 3. Plan treatment based on the profile and the statistics, including teacher time, class organization, etc.
----------------------	--	--

6.4.2 Additional In-Service Workshops

During the 1998-99 school year, I directed in-service workshops for prominent teacher-leaders from throughout the country, who could then run workshops on assessment in their home areas. Teacher-leaders who had completed a national course in assessment and who were recommended by local Ministry of Education supervisors were chosen. The teacher-leaders from the *Ofek* courses were also included in these workshops.

The first workshops took place in the summer vacation and were two days long. At the end of the workshop, the teachers were required to fill in a questionnaire. Fourteen teachers responded.

Questionnaire – 2.7.98

1. Record three important things that you learned in these two days.
2. Would you like to continue to meet during this year?

If so, what subjects would you be interested in?
3. Comments on the process of implementing the task bank for assessment.

Question 1: Eleven teachers mentioned activities related to the article on discourse as being very important to them. They wrote not only about the value of the article itself, but also noted that the didactic approach I had used in presenting it had been very helpful. The other main topics mentioned were: the discussions about their work as leading teachers; their sense that a support group was developing; and comparisons between the different tasks.

Question 2: All the teachers responded positively to the question about future meetings. Eight of them were interesting in gaining experience in writing tasks, including tasks intended for teaching, and not only those intended for assessment. Six teachers wanted to discuss further how to work with a school which is just embarking upon the process of school-based assessment. Four teachers were interested in reading more articles like the one we had read. The other topics they requested were: becoming familiar with additional assessment tasks from the bank; learning about new report cards; building math centers; encountering new materials; and learning about learning problems in mathematics.

Question 3: Answers to this question were mostly related to technical problems, such as the Ministry of Education's need to convince school principals of the importance of school-based assessment. The teacher-leaders felt that they were meeting resistance from principals: the latter were unaware of the value of the task banks and due to an overload of commitments, did not have the time to familiarize themselves with the issue. These teacher-leaders had the impression that participation in *Ofek* had a positive influence on the teachers, but that in general, the process of implementing this new type of assessment was slow.

6.5 Interviews with Teachers

I interviewed four of the teacher-leaders who had been working with the investigative Assessment Task Bank for several years and had been involved in the *Ofek* course. These unstructured interviews were part of an effort to determine the influence of the Assessment Task Bank on their work. Three of the teacher-leaders represented the different streams of Israeli elementary school – religious, secular and kibbutz. The fourth teacher is at the stage where I would like all teacher-leaders to be. She understands mathematics well and can always see the big picture, without getting bogged down on minor issues.

6.5.1 Interview with Teacher BC

Teacher BC attended the Second National Course on Assessment and was active in *Ofek*. She organizes the religious teacher-leaders of her area. She is always looking for ways to increase her professional knowledge – i.e., attends all possible in-service courses including one on games (not just related to mathematics) and is now studying computer use.

Five main points arose in this interview:

1. She said she had always wanted to know what was important or necessary to teach and she felt that the Assessment Task Bank gave her a place to turn to for advice.

For example, the question about one-fourth [the student is given six squares and is asked to draw different examples of one-fourth]: “I like it very much because you can give it and really see if the pupil understands or not”.

2. The way I ran the in-service courses impressed her very much:

I can't separate the workshops from the Assessment Task Bank, but the workshops were really different. I had always felt that in a workshop one needed to complete, to cover a lot of material, but your relaxed approach—and we did do a lot. ... another thing I learned from you was reflection, what did we learn? And lots of other things, really.

3. “The criteria help us to judge the students in a much better way—instead of only using basic skills, computation, now we have wider goals, like verbalization and generalization.” She also discussed how often the teachers discovered students were capable of succeeding, and this influenced their teaching.

4. Using the tasks as tests results in teachers coming and asking for help—“teach me...”.

(See discussion in the Chapter 2 of the teachers’ need to learn by constructivistic methods, which means including self-motivated learning.)

5. She talked about her experiences with the teachers, and their use/lack of use of the textbook:

I try to get them to use the textbook correctly and not all sorts of complicated worksheets—so I haven’t gotten to the stage you’re talking about.

It’s a process. Five years ago, the students didn’t have textbooks or manipulative materials, and now all my schools do. Even in the good communities, five years ago they didn’t have manipulatives.

I’m still “fighting” with the teachers to use the tests of the textbooks, because they think they are too easy, yet they are unwilling to use the tasks of the task-test bank. [They want to use tests which have much more complicated calculations on them.]

First and foremost, I would like things to be orderly, with proper textbooks, manipulatives. In this good school, in third grade, they still haven’t used the manipulatives because “they aren’t necessary—they’re only for play!”

Their worksheets have lots and lots of dry exercises, no thinking—only skills.

6.5.2 Interview with Teacher AG

Teacher AG attended the First National Course on Assessment and an additional course on the School-Based Assessment Project for leaders of regional areas. This

second course was run by Professor Arye Levin, the academic advisor of the whole project (see Chapter 1), and not just the mathematical part. After this course, Teacher AG worked in the Department of Evaluation, Ministry of Education, in her region, although her job tended to be mostly administrative.

The interview I held with her focused on the way her school works in mathematics. This is a secular school where Teacher AG has taught in the past and where she has worked as a teacher-leader for many years. In addition to the usual textbook-based lessons, there are math centers situated in the corridors, which contain authentic problems.

Look, the mathematical centers include open tasks that the children work on using thinking and verbalization and all that stuff. Generalizations where needed. Discovery and investigations and such things. ... For example, here is a restaurant ... The kitchen workers need to prepare themselves for the guests and to cook two meat meals. Please help them prepare this request - ingredients, how much to buy, the costs, etc.

She talked about how students work together in groups, discussing their tasks, and mentioned that the fear and trepidation that the subject of mathematics used to instill in the pupils is now a thing of the past.

Twice a year, however, the teachers continued to use summative assessment tests, prepared by this teacher-leader. These tests contained a variety of types of questions, and although there were some open ones requiring more than one solution, these investigative type questions constituted only a small part of the test as a whole. This was due to the fact that all the topics studied in the course of six months—including geometry—are included on a single test, and the student's proficiency in each topic is assessed. Consequently, although she professed to be directing change in the school's math education towards open questions, the message she was conveying to the teachers was an ambivalent one.

6.5.3 Interview with Teacher CJ

Teacher CJ is a kibbutz teacher who attended the Third National Course on Assessment. Teacher CJ is a teacher-leader in her school. The fact that her school is in a kibbutz means that the teachers devote more time to working together, since they are in no hurry to rush home at the end of the school day like teachers in the city.

Her school held workshops on alternative assessment, not specifically in mathematics, and she was adamant that alternative assessment should encompass all fields, and not solely mathematics. However, the ideas presented in the workshops were not relevant to her math teachers, who wished to modify their assessment practices in line with their changing style of teaching and learning. Thus she worked with the school staff on alternative assessment to help them apply such assessment methods to the field of mathematics.

Teacher CJ noted that she had always strived to teach in depth, and consequently did not manage to complete as much material as the other teachers she met. She had previously interpreted this as meaning that she was not a good teacher. Nonetheless, she had always felt it was her job “not to teach arithmetic, but rather to teach thinking, mathematical thinking”, and consequently, she had attended many in-service courses, becoming a specialist math teacher and a teacher for gifted children.

She discussed how she had found it necessary to work on the textbook with the teachers, for fear that they would otherwise “turn it into a collection of exercises” and view their purpose as “to progress through it”. She explained how she had learned to work with heterogeneous groups of students, with them writing their conclusions on overhead transparencies and then running a whole-class discussion, and how she had taught her teachers to open exercises, to look at the mathematics in an innovative way, including the crucial discussion after the investigation to raise the level of thinking. (These practices were integral parts of my in-service courses.)

She described her view of mathematics as being “not only a subject that deals with the real world, but which also has an internal structure that needs to be taught”. She

contrasted this view with views that arose at a discussion about instructors for special education at her local teachers' college, who teach with emphasis on computation.

We discussed the use of Cuisenaire rods, a controversial subject in Israel today, and this led into a discussion of how children solve simple computation exercises (like $8 + 7$):

One thing I do with such exercises is have everyone tell how he did it, and we see that the children really do use different methods.

She explained that the tests accompanying the textbook graded the children's ability to carry out exercises exactly as had been learned, but did not evaluate their understanding. This was not enough for her teachers.

I started to sit with the teachers of each grade level, and to consider how we could add to the textbook tests two or three questions of a different sort. But we said that we couldn't give questions like these on the tests if we didn't work like this in the lessons.

She explained how she worked with the teachers on the tasks from the Assessment Task Bank

We discussed how we check them, what they teach us about the students. ... and we also talked afterwards with about the tasks: what sorts of questions these are, is there something different here, do you like them, what did you like? It was a very open discussion. Then we went in another direction, to a very broad area, and asked the students what they thought should be checked in mathematical thinking. I had already told them we would call this not an arithmetic lesson, but rather a lesson in mathematical thinking, and they were always reminding me of this because I always forgot. What should we assess in that lesson, and they wrote topics and afterwards we compiled them and reached ten criteria, and each student choose for himself two that he wanted to improve on. It was a long process.

You did this and you had ten criteria, did it affect your teaching?

Teacher CJ: Yes. For example, I put more emphasis on individual work this year.

Describe what they had to do.

Teacher CJ: The student had to choose a subject, he had to collect data and turn it into a diagram. I said they could do the diagram either by hand or by computer, depending on their proficiency. Then he had to write at least three things that one could learn from the diagram. Then he presented it to the class, and the other students had to tell him what they could learn from the diagram, or to ask him questions.

Do you think this project led the discussion where you wanted it to go?

Teacher CJ: I think so. I'll tell you what I wanted. I wanted to show them that diagrams are a way of demonstrating for us graphically, in an interesting and attractive way, numerical data that as numbers are less clear.

However, when questioned about the mathematics they learned through these projects, she realized that they were deficient in this respect.

Teacher CJ: I also didn't really do that, because now I remember they measured rainfall all week in the different towns, they measured temperature, they calculated the average, they saw where it rains more and where less, and where the temperature is higher, things like that. To tell you that we worked on averages from a mathematical point of view—not on these projects. I now understand what you mean—if I add another measurement or two, what will happen to the average? Will the average rise or fall? ... I can tell you that in the topic of averages we dealt with the mathematical aspects, but not in these projects.

Why?

Teacher CJ: I just didn't think of it.

This is an example of a teacher who knows mathematics, yet still might “forget” to include it in her projects. This was a common occurrence among Israeli elementary school mathematics teachers: the problem of incorporating the essential mathematical

ideas into their projects. They tended either to work on straightforward mathematics, often from the textbook, or to work on large projects, such as planning a party or a trip—but not to be able to integrate the two aspects.

6.5.4 Interview with Teacher BD

This teacher attended the Second National Course on Assessment and was a local leader in her area. In the last two years, she has written three booklets with investigations for elementary school pupils. Her mathematics knowledge is good and she understands better than most of the teacher-leaders the importance of the different topics, and the relationships between them. My interview with her centered on assessment, and the use of the Assessment Task Bank.

When you prepare a test today, what do you do?

Teacher BD: I work with the teachers ... we think:

- a. What mathematical topics?
- b. If the test contains thinking—what percentage of the test will involve thinking?
- c. Then, with some of the teachers, we talk about types of thinking, thinking skills: direct thinking, reversed thinking, find all the possibilities, is there a generalization, a pattern.
- d. Will there be one test for everyone, or different tests. Nowadays I prefer one test with graded questions, ranging from easy to difficult.
- e. Another thing on which I work with them is to ensure that there should be a task and not lots of unconnected questions.

We sit and talk, they start writing, and bring their proposals to me for further discussion. It's a process.

We do this while they are teaching the topic, not before they start. It's a long process and the discussion about the test influences their teaching.

I work with Teacher AO and Teacher BF [teacher-leaders from Tel Aviv who had been in the Third National Course on Assessment], and we want to create an assessment task bank with different tests for the major topics, geared toward the first to sixth grade. But for some topics, like averages, we used assessment tasks from the Assessment Task Bank. Mostly we need assessment tasks for second and third grades, where there is no task bank.

What are the elements of a good test?

Teacher BD: I think a test needs to have an outline “story”, rather than being composed of a collection of short questions. [All the questions on her test relate to fractions in a “cloud”.]

I look at the types of thinking that the student needs to do.

I look to see if there is scaffolding: to make sure that the test doesn’t start immediately with hard questions.

I also look to see what specific topics the test includes, not just the major subject like percentages, fractions.

What about writing?

Teacher AD: That’s naturally included, it’s part of the “story”. There is no test today that doesn’t require students to explain and justify. It’s already an integral part of the teaching too. I can go into a class that I haven’t visited for many two months, and I will see how the teacher asks the students to explain, to verbalize and to write. In the beginning we asked for oral verbalization. Later I required them to have the students write, because writing requires each student to sit by himself and do personal reflection. Verbalization is such an integral part of our work that I didn’t even think to mention it.

In relation to the test that I showed her:

Teacher AD: All these things, they are things we can check through our usual work. We use small quizzes to find out what students know. I don’t think a test needs to be built from lots of unrelated items, because I believe that a good test needs to show me a higher-level of thinking. On this test, all these questions reflect skills—that’s all.

In relation to the process:

Teacher AD: I think that it's not right that we should always take a task from the Assessment Task Bank. I think that there needs to be a combination. Teachers need to feel how hard it is to write a good task so that when they receive a good task they can appreciate it.

Where does the change in teaching enter into the picture?

Teacher AD: It begins with the teacher-leader. The main thing that influences the teachers is the model I present—they see the tasks I bring the students, how the students talk differently, think differently.

Then we build lessons together. ... They need to write each word the students say, and then when we analyze the lesson I slowly mix in the new elements. For example, we talk about the use of language, what language the students used, whether they used mathematical concepts. Then I ask the teachers how we can make the language become more mathematical. Thus they begin to be aware of the students' language.

This way, each time I introduce another goal, without the teachers being aware of my ultimate purpose. They tell me that **they are not the same people they were before.**

How does the Assessment Task Bank enter into all of this?

Teacher AD: With a new teacher I bring a task (like Tables), and when she sees the results of her classes' work she is surprised because she thought it would be easy for her students. Then we analyze the results, using the criteria. And they say, "ah, the students aren't used to that type of test." ... they begin to understand that it's not a matter of the student knowing or not knowing, that the emphasis is in another place. There are other things besides skills in mathematics.

... It helps that the Assessment Task Bank is from the Ministry of Education – it comes from an authority, other schools use them too. It's not like a test that I wrote ... it's official!

I always taught this way—the Assessment Task Bank only reinforced this and I was glad that now everyone would be teaching this way.

She then gave an example of a lesson she had:

I had a student, for example, we were talking about the sum of the angles in a triangle, and someone asked: if I know that the angles are such and such, then do I know what sort of a triangle I have? [isosceles, scalene] They investigated and found that they could determine the type of triangle. Then they asked if the same thing happened with quadrilaterals. We had a quadrilateral with angles 120, 60, 70 and 110, and one student said it would be a trapezoid. How did he know? We had not learned about the angles adding to 180. I asked him and he said he *felt* it would be a trapezoid! It was wonderful! Fifth grade, I had never taught such stuff. What a wonderful discussion we had—because of their questions.

6.6 Test/Task Analysis

In the course of my research, I collected sample tests and tasks used by the teachers and teacher-leaders involved. The changes in the teachers' views of mathematics were reflected by changes on these tests and tasks.

The tests used by teachers before the beginning of the In-School Assessment Project tended to be composed of many short-answer computations and several standard word problems. Students were not required to justify their solutions, nor were they asked to solve open problems which had more than one possible answer. The following three tests illustrate these points.

Test – Factors and Prime Numbers

1. Factor into prime numbers:

15, 70, 18, 21, 63

2. Circle the prime numbers:

15, 11, 3, 19, 27, 30

3. Give examples of three additional prime numbers: _____, _____, _____

4. Mark in the table which numbers are divisible by 2, by 3, and by 9:

	divisible by 2	divisible by 3	divisible by 9
534			
621			
15			
1200			

5. Circle the numbers divisible by 6:

218, 33, 24, 413, 412

TEST – WHOLE NUMBERS

1. Solve the following exercises (copy in vertical format):

a. $1278 + 459 + 12 =$

b. $3800 - 2389 =$

c. $450780 - 24897 =$

d. $306 \times 7 =$

e. $235 \times 47 =$

f. $345 : 5 =$

g. $1268 : 25 =$

2. Find the difference between 5001 and 4899.

3. Find two numbers whose difference is 38.

4. Write down the number “eight million, five hundred thousand and one”.

5. Circle the number “twenty-one thousand sixty-five”:

21650 2165 21065

6. Write a word problem to the exercise $250 \times 24 =$

7. The price of 60 pencils is 240 shekels.

What is the price of 78 pencils of the same kind?

8. Write all the multiples of 9 between the numbers 162 and 200.

TEST – PERCENTAGES

1. Complete the table:

percent	decimal	fraction with denominator a power of 10	reduced fraction
5 %			
	0.7		
			$\frac{1}{4}$
			$\frac{1}{2}$

2. 12 % of 240 =

3. 129 students, who comprise 25 % of all the school's students, registered for camp.

How many students are there in the school?

4. The apples in the orchard were damaged by the frost.

The farmer checked one crate which held 35 kg. of apples and found 7 kg. damaged apples.

In the second crate which held 40 kg. apples, 14 kg. were damaged.

a. What percent of damaged apples was there in each crate?

b. What percent of damaged apples is there altogether?

A context-based task, Four Cups, shows similar features and illustrates quite well the low-level cognitive demands that were being required.

FOUR CUPS

A cup of wine contains 150 ml.

An adult drinks from a cup

A small glass of wine contains 80 ml.

A child drinks from a small glass

A bottle of wine contains 750 ml.

A bottle of wine costs 10.70 shekel.

As the Passover Seder there were 7 children and 8 adults.

Each guest drank 4 cups or small glasses of wine, as required.

Compute:

1. How much wine did the children drink?
2. How much wine did the adults drink?
3. How much wine did all the guests drink at the Seder?
4. Using these numbers, how many bottles of wine needed to be bought?
5. How much did the wine cost?
6. It is customary to have a glass of wine for Eliahu the Prophet.

Does this require buying another bottle of wine? Explain.

An example of the changes which occurred after exposure to the investigative assessment tasks can be seen in an analysis of two fourth grade tests written by teacher AH (a participant in the First National Course on Assessment). These were end-of-the-year tests used by approximately ten schools each time.

	Test 1, 1996	Test 2, 1999
Number of items	38	24
Multiple choice	13 %	8 %
Geometry	13 %	17 %
Questions which required an explanation	8 %	17 %
Questions which required more than one answer	46 %	50 %
Investigation-type questions	13 %	25 %

It is evident from this data that teacher AH had changed her approach toward preparing a test – fewer multiple-choice questions, more verbalization, and more thinking.

Another example is a test written by teacher CL (a participant in the Third National Course on Assessment). This was a test for multiplication of decimals in which the students were given twelve computation exercises, but asked not to solve them. These exercises were followed by six questions which required investigating the exercises, mostly by doing estimations. This was a unique type of test, which had not been used before.

Multiplication of Decimals – Teacher CL

Here are some multiplication exercises.

Answer the questions **without** solving the exercises.

3.2×10

$g) 3.14 \times 2.4$

0.32×100

$h) 6.7 \times 0.6$

1.25×0.2

$i) 1.45 \times 25.3$

42×0.5

$j) 132 \times 2.4$

12.3×0.02

$k) 4.12 \times 6.4$

1.51×0.4

$l) 3.5 \times 0.1$

Which exercises have an answer larger than 100?

Which exercises have an answer smaller than 50?

Which exercises have an answer larger than 50 but smaller than 100?

Which exercises have an answer smaller than 10?

Which exercises have a product larger than either factor?

Which exercises have a product smaller than either factor?

Even tests which focus primarily on computation had changed. This multiplication test is one example of such changes. It requires explanations and includes some questions which call for more than one solution.

Test – Multiplication

1. Solve:

a. $10 \times 300 =$

b. $1000 \times 4 =$

c. $706 \times 10 =$

2. Solve:

a. $40 \times 3 =$

b. $5 \times 600 =$

c. $300 \times 50 =$

3. Use the distributive property to solve:

a. $5 \times 24 =$

b. $6 \times 203 =$

4. Circle “correct” or “not correct” for each:

$43 \times 3 =$ a. $40 \times 3 + 3 \times 3$ correct / not correct

b. $4 \times 3 + 3 \times 3$ correct / not correct

c. $42 \times 3 + 1 \times 3$ correct / not correct

d. $41 \times 3 + 42 \times 3$ correct / not correct

For those exercises to which you responded “not correct”, explain.

5. What is the missing factor? $102 \times \underline{\quad} = 1020$

Explain your answer: _____

6. Write 3 multiplication exercises whose product is 1200:

_____ \times _____ = 1200

_____ \times _____ = 1200

_____ \times _____ = 1200

Are there any more such exercises? yes / no

If so, write 3 more:

Class tasks had also changed their format, often requiring higher-order thinking:

WITH 1, 2, 3, 4

Use the digits 1, 2, 3, 4.

1. Write 8 exercises (2 exercises for each arithmetic operation – addition, subtraction, multiplication, and division) so that the answers will be:
 - a. the largest possible
 - b. the smallest possible
2. Describe how you worked
3. Write generalizations about choosing the numbers for each operation

Division

1. Solve the exercises. Choose one exercise and explain your solution.

Formulate a rule.

- a. $5 : \underline{\quad} < 1$
- b. $\underline{\quad} : 5 < 1$
- c. $\underline{\quad} : \underline{\quad} < 1$

2. Solve the exercises. Choose one exercise and explain your solution.

Formulate a rule.

- a. $5 : \underline{\quad} > 1$
- b. $\underline{\quad} : 5 > 1$
- c. $\underline{\quad} : \underline{\quad} > 1$

3. Explain how exercises 1 and 2 are similar and how they are different.

In addition, the teachers were writing what they called “authentic” tasks. These placed heavy emphasis on realistic situations, but often involved only procedures and not investigative mathematics. For example, they would give the students a floor plan for an apartment and ask them to calculate the area of each room and what percentage it would be of the total apartment. Basically, this type of problem is no different from

standard textbook exercises—the only difference is that the floor plan is a photocopy of an advertisement from the newspaper.

Another example is of a two-page exercise that deals with averages and long-distance jumping. Half of the task merely requires the procedure of computing the average. The second half requires understanding averages and their properties—but all four questions in fact relate to the same property: that the average cannot be larger than the largest number involved.

The teachers were clearly making an effort, but they were having difficulty writing good tasks, often due to their lack of mathematical knowledge. Observing their difficulties convinced me that curriculum materials need to be prepared by experts and that the job of the teacher is to know how to use these materials and how to adapt them to the needs of her students.

6.7 Conclusions of the National Study

In conclusion, I would like to return to my three research questions. The first research question was: *Will teachers' use of mathematical investigation tasks influence the way the teachers see mathematics?*

When I look back over the progression that the teachers have made since the establishment of the Assessment Task Bank to today, it is clear that their view of mathematics has definitely been influenced. Israeli elementary school teachers tended to look upon mathematics as a “skill” subject, mostly based on arithmetic computations. Looking over the teachers' quotes one sees that this opinion has changed. Now they consider mathematics to include additional aspects, such as generalizations and justifications, verbalizations and creativity as well.

My second research question was: *In what ways will teachers' use of mathematical investigation tasks for assessment influence the way they teach?*

The teachers' new views of mathematics (see above) have certainly affected their teaching. They claim that their lessons have changed (although I have not had the opportunity to observe this first-hand). They use more group tasks, have more discussions with their students, and place more emphasis on explanations. In the textbook most schools use there are challenge problems which tended to be ignored, or used only by the gifted pupils. Now teachers are using them for all the students and often making these problems the focus of lessons – i.e., turning these problems into investigations.

My third research question was: *Do teachers' use of mathematical investigation tasks influence the way they assess their students?*

The teachers' assessment practices are indeed undergoing change. They are learning to use additional tools for assessment, not just conventional tests. In addition, the tests themselves have changed from requiring only short answers to computation problems to requiring explanations and involving questions with more than one answer. Some teachers are including performance on investigations as an element in their evaluation of their students.

CHAPTER 7: THE RESULTS OF THE SHARON STUDY

Chapter 6 was concerned with the study of a selection of teacher-leaders from throughout the country who were exposed to reform mathematics through national courses on assessment. The present chapter is devoted to the study of a small Israeli community in which all six elementary schools adopted the use of school-based assessment in mathematics. This intensive study was, in effect, a simulation of an ideal situation, in which all teachers and teacher-leaders were exposed to reform mathematics. (See Section 2.2 for definition of "reform mathematics".)

As explained in Chapter 5, this community was chosen because the leading teachers of each school were already involved in an in-service project which placed emphasis on reform mathematics in an effort to transform their concept and vision of mathematics. This prior exposure to reform mathematics made the Sharon community ideal for my purpose. As described in Chapter 5, during the first year of the Sharon Project I was the academic advisor. My research started in the second year of the Sharon Project in which I directed all the weekly workshops, and gave a course on assessment to all the Sharon elementary school mathematics teachers.

7.1 The Second Year of the Sharon Project

During the 1996–97 school year, the second year of the in-service project, a total of 14 teachers participated in weekly workshops. During the spring term of that year I ran three additional workshops on assessment practices. All fourth- to sixth-grade teachers of mathematics from the Sharon community were expected to attend these workshops. Approximately 25 teachers participated in each workshop.

7.1.1 Background Questionnaire

At the first of these three workshops the teachers filled out a questionnaire, giving data about their education and teaching experience, as well as providing examples of typical

tasks they give in mathematics classes, along with information regarding their assessment practices (see Appendix B for the full questionnaire). My purpose was to learn more about the background of the teachers, and more importantly, to obtain information such as: the kind of assignments they were giving their students, and the type of assessment practices they were using. This information gave me a baseline against which to gauge the effect of exposure to the views on assessment that I was promoting.

A total of 29 teachers, ranging from fourth to sixth grade, attended this workshop. Almost half of these (13) had also attended the workshops on the in-service project. This questionnaire therefore provided a useful distinction between teachers who had already attended workshops on reform mathematics, and those who had no prior experience with such concepts. It was thought that a breakdown of the results on the basis of this distinction would provide useful information on the effect of prior exposure to such concepts on the teaching and assessment practices of the teachers.

An analysis of the personal data provided by the teachers shows that both groups were quite similar in composition. The only significant difference is that four of the teachers who also attended the in-service project had an university education, whereas all the other teachers were educated in teachers' colleges. (The one teacher with a university education who was not in the project joined it the following year.) The following summary presents the information regarding teacher education, classified according to this distinction.

Where did you obtain your formal education?

	Teachers in in-service project	Teachers <u>not</u> in in-service project
Teachers' college	9	15
University	4	1

However, despite the fact that more teachers in the ongoing project had a university education, few had actually studied mathematics towards a higher degree. Only six teachers—three from each workshop—had in fact majored in mathematics, whereas most of the teachers had majored in "humanistic" subjects, such as language, literature, Biblical studies, or education.

The following summaries present the statistics regarding the participants' years of experience in teaching in general and in teaching mathematics in particular. It is evident from these statistics that both groups were quite similar in terms of the teaching experience of their members.

How many years of teaching experience do you have?

General teaching experience	Teachers in in-service project	Teachers <u>not</u> in in-service project
0–2 years	1	3
3–10 years	4	2
At least 11 years	8	11

How many years of experience do you have in teaching mathematics?

Experience in teaching mathematics	Teachers in in-service project	Teachers <u>not</u> in in-service project
0–2 years	1	3
3–10 years	5	2
At least 11 years	7	11

In order to obtain information on the kinds of assignments teachers were giving, the following questions were asked:

II. Choose one class in which you teach mathematics and answer the following questions:

What grade did you choose?

1. What topics did you teach in mathematics last week?
2. Give two examples of tasks that the students were given last week.

Since a perusal of the subject matter alone was not sufficient to indicate whether or not they were teaching "reform mathematics", they were required to give examples of the tasks they were giving their students. These examples were classified into different "types of task", and the following breakdown was obtained:

Type of task	Number of times this type of task was described	Number of times this type of task was described by teachers participating in in-service project	Number of times this type of task was described by teachers <u>not</u> participating in in-service project
Practice, textbook work	31	8	23
Thinking puzzles	4	2	2
Understanding operations	6	5	1
Creating a story	4	3	1
Investigating geometry	8	7	1

The results indicate conclusively that those teachers who also participated in the in-service project were teaching more in the direction of reform mathematics. It should be pointed out that since these results are based on the teachers' own self-evaluation, rather than on an objective observation of their teaching habits, they may not reflect their actual teaching methods. Nevertheless, they point to the fact that those teachers with more exposure to reform mathematics at the very least have assimilated what is expected of them. Of the teachers who referred to tasks involving straightforward

practice drills (as opposed to tasks requiring understanding or higher-order thinking), more than twice were those who had not attended the in-service project workshops. The following are examples of tasks that were classified as "Practice, textbook work":

A worksheet on the subject being learned: addition and subtraction of fractions and mixed numbers with similar denominators.

Exercises and problems on the blackboard. (teacher DN.IV.1)

Exercises. Practice in math exercise books. (teacher DO.IV.1)

A worksheet, with word problems. (teacher EC.IV.1; p.65)

In many cases, even when these teachers used varied teaching methods, the mathematics tasks tended to be very conservative, for example:

A worksheet practicing the addition being learned, which, if solved correctly, spells out a proverb. Work on work cards using fraction rulers, expanding and reducing fractions. (teacher DS.IV.1)

The teachers who had previously attended workshops, on the other hand, gave examples which demonstrated that they understood the importance of student comprehension:

Mark which of the exercises require borrowing:

42 - 35, 68 - 24, 49 - 7, 81 - 3, etc. (teacher DD.IV.1)

Place the following fractions on the number line:

$\frac{1}{2}$, $\frac{1}{3}$, $\frac{5}{5}$, $\frac{13}{12}$, $\frac{0}{4}$ and explain how you decided where to place them. (teacher DH.IV.1; p.56)

Verbalization of problems without giving their solutions.

Weekly thinking puzzle in arithmetic. (teacher DA.IV.1; p.51)

Another distinction between the two groups of teachers that emerges from the classification of their assignments into types of tasks is that the teachers from the in-service project gave more tasks involving investigative geometry. This is not

surprising, since geometry was the subject on which they worked in the in-service project workshops during the previous year. For example:

Write a dialogue between a square and a rectangle. (teacher DY.IV.1)

The examples of assignments cited by one teacher who had attended the in-service project were interesting, since her two examples—one in arithmetic and the other in geometry—exemplify diametrically opposed principles.

a. $1\frac{1}{2} + 3\frac{1}{3} =$

b. An investigative problem about perimeters and areas of rectangles and squares. (teacher DL.IV.1; p.58)

Her example of a task in arithmetic is a purely conventional one, involving straightforward practice drill, while her geometry task involves the use of investigative problems. Clearly, she understood the use of investigative problems in geometry, whereas in arithmetic she was still using standard exercises. This, once again, can be attributed to the fact that the in-service project workshops dealt in depth with investigative geometry during the previous year.

In order to evaluate their assessment practices, the following questions were asked:

III. Ways of assessment

1. What tools (e.g., tests) do you use to assess your students' achievements in mathematics?
2. Why do you check your students' achievements in mathematics?

The summary below presents the types of assessment tools that were used by the teachers.

School (No. of teachers)	Quizzes, tests	Investigative tasks, games	Homework, classwork	Participation	Group work with teacher	Computer drill
K (3)	2	2	1	--	1	--
G (3)	3	--	1	1	1	--
A (5)	5	--	5	2	--	4
O (4)	3	--	2	--	1	--
H (4)	4	--	1	1	2	--
U (10)	10	2	7	2	2	1
29 teachers Total	27	4	17	6	7	5

This summary reveals the following findings:

1. With the exception of two teachers, all the teachers mentioned tests or quizzes.
2. Only four teachers mentioned using an investigative task or a game for assessment purposes.
3. One school, A, was still using the printouts of a computer drill program for assessment.
4. More than half of the teachers (17 out of 29 = 59%) included the students' homework and/or class work in their assessments.
5. A fourth of the teachers (7 out of 29 = 24%) based their assessment on matters that were learned in the course of teaching a small group of students during a lesson.

The results were then broken down according to the participation of the teachers in the in-service project. On the left-hand side of each square: the teachers who had participated in the project; on the right-hand side: those who were participating only in the assessment workshops.

School (No. of teachers)	Quizzes, tests		Investigative tasks, games		Homework, classwork		Participation		Group work with teacher		Computer drill	
K (3)	1	1	2	--	1	--	--	--	--	1	--	--
G (3)	2	1	--	--	1	--	1	--	1	--	--	--
A (5)	2	3	--	--	2	3	1	1	--	--	2	2
O (4)	--	3	--	--	1	1	--	--	1	--	--	--
H (4)	4	--	--	--	1		1	--	2	--	--	--
U (10)	2	8	1	1	--	7	--	2	2	--	--	1
29 teachers Total	11	16	3	1	6	11	3	3	6 1	2	3	

Two interesting findings emerged:

1. At school U, none of the teachers who were participating in the in-service project included the students' homework and/or class work in their assessments.
2. All of the teachers who based their assessment on matters learned when teaching a small group of students were those who participated in the in-service project.

These findings indicate that participation in the in-service project, where teachers were exposed to reform mathematics, affected the teachers' approach to assessing their students. Rather than merely collecting papers and grading them, they were beginning to base their assessment of their pupils' abilities and achievements on listening to what they had to say.

The various responses that the teachers had given regarding purpose of assessment (Question 2) are presented in the following summary:

Purpose of assessment	No. of teachers
To discover what the students know	28
To give a grade on the report card or a report to parents	5
To evaluate oneself as teacher	7
To plan future appropriate work	14
To divide the students into homogeneous groups	1

As can be seen, the main reason these teachers listed for checking their students' achievements in mathematics was to discover what the students knew. This was expressed by the teachers in various ways:

Where the class is at in the syllabus. The amount of understanding of each pupil. (teacher DA.IV.1; p.51)

In order to help each pupil with his weak points, and to give enrichment activities to pupils who need them. (teacher DE.IV.1)

Checking knowledge to enable each child to progress from the point where he is at. (teacher DH.IV.1; p.57)

To know on which topics each child has trouble or is proficient.
(teacher DN.IV.1)

The one teacher who did not refer directly to the students' achievements in this question (teacher DF.IV.1; p.54) responded: "Mainly in order to plan".

A total of 19 teachers used either positive or neutral words when describing student achievement—words such as "understand", "absorb", "comprehend", "be proficient", or "internalize". Only eight teachers (four out of those 19, as well as four others) used negative words, such as "trouble spots", "difficulties", "not understanding", or "stuck". This general trend in attitude is a positive one in itself; however, as shown above, these responses almost exclusively refer to mathematical skills, rather than to higher-order

thinking. (See discussion above on the sort of tasks they were using in their classrooms.)

This emphasis on the acquisition of skills also influenced the sort of responses offered regarding the planning of future work.

Planning in order to identify trouble spots, to plan further work, according to the individual needs of each pupil. (teacher DG.IV.1)

...choose further work appropriate to the individual pupils and the curriculum.
(teacher DU.IV.1)

...check what my pupils have absorbed and understood, what needs to be strengthened and reviewed. (teacher DM.IV.1)

One teacher summed up the purpose of assessment as follows:

- 1) In order to know if they understood and absorbed the material, and are able to use what I taught.
- 2) To discover trouble-points and where the students need help.
- 3) To check each individual's progress in relation to himself and in relation to the class.
- 4) To find difficulties in putting across the material. (teacher DO.IV.1)

The final questions of this first questionnaire dealt with staff collaboration. Reform mathematics teaching and alternative assessment are difficult for teachers to cope with individually, and thus I wanted to know what staff collaboration was occurring.

Question 1. *How often does the staff meet?*

School (No. of teachers)	Once a week	Once every two weeks	Once a month	A few times a year	Does not meet
K (3)	3				
G (3)			1	1	1
A (5)	4	1			
O (4)	3				1
H (4)	2	2			
U (10)	3	4		2	1

At first glance, these results—in which different teachers in the same school appear to disagree about the frequency of their meetings—seem somewhat enigmatic. However, when correlated with the answers to question 2 (*Who participates in the meetings?*) the apparent contradiction disappears, since the teachers responded that these were meetings of the math teachers of one or two grades, and not of the entire school. In other words, the teachers of some grades met more frequently than others. Another reason for the discrepancy might be a difference in opinion between teachers regarding the definition of a “meeting”. The teachers did confer regularly with the other teachers of the same grade level. At times this might have been a meeting of only two teachers, who did not necessarily discuss mathematics. This, for example, may have been considered to qualify as a meeting by one teacher, but rejected as not qualifying as a meeting by another.

The results of the two subsequent questions (3: *What are the subjects usually discussed?* and 4: *What did you do at the last meeting?*) were analyzed in conjunction.

	Question 3: No. of teachers	Question 4: No. of teachers
Planning lessons	22	16
Preparing materials, worksheets	11	11
Preparing tests	10	7
Discussing teaching methods	9	0
Talking mathematics	1	3

It can be seen that most of the time the teachers spent in such meetings was devoted to planning, with an emphasis on preparing materials, especially worksheets and tests. Some teachers, however, wrote about discussing teaching methods and even about having observed another teacher's class and subsequently discussing it. Some of these comments are exemplified below:

Ways of teaching. I always give some tips. (teacher DD.IV.1)

Ways to teach new subjects, types of tasks to give to students.
(teacher DQ.IV.1)

Different approaches to teach a subject. Comments on a lesson which was given. (teacher DH.IV.1; p.57)

I observe a lesson of the other teacher at my grade level, and afterwards we meet and talk about what happened in the class and the whole subject being studied. (teacher DX.IV.1)

One surprising find is the discrepancy between the number of teachers who mentioned "discussing teaching methods" as frequently occurring at staff meetings (nine teachers) and the number who mentioned this as taking place at the most recent meeting (zero).

One possible interpretation of this discrepancy is that such discussions indeed took place, but only rarely, and thus were etched on the teachers' minds.

An even rarer event involved a discussion about the mathematics itself. One teacher referred to helping another teacher:

A teacher who has trouble understanding some specific material, then another teacher helps her. (teacher DC.IV.1)

Three teachers who had participated in the in-service project workshops, with their emphasis on challenge problems, referred to solving such challenge problems with the other teachers.

We solved mathematical problems, challenges, that we are going to use in the class. (teacher DA.IV.1; p.52)

This is not surprising, since the solving of mathematical problems was an integral part of our workshops.

The last part of the questionnaire was open, to allow the teachers to express requests for the assessment meetings. The responses showed that the teachers' expectations were consistent with the plan for the assessment workshops.

	No. of teachers from ongoing workshop	No. of teachers not in ongoing workshop
To learn new, better ways of assessing	6	9
To learn to prepare better tests	1	1
To learn if the way I assess today is good	2	1
No requests	5	4

7.1.2 Sample Tests

As mentioned in Chapter 3, I collected sample tests used by the schools during the spring term in 1997. Most of these tests were quite conventional—that is, they consisted primarily of computation exercises, with one or two standard word problems. There were a number of questions that required the students to explain their work, i.e. verbalization, some non-algorithmic thinking, like estimating or reversed thinking (where the answer is given and the student has to find the missing element), and some problems where students needed to find multiple solutions.

Only one school (K) included samples of investigations that had been used for assessment purposes. Yet the test included was very conventional with many, many computation problems (like $298,365 + 198,642 =$), standard word problems, and fifth-three items to answer!

One teacher (from School O) included a variety of types of assessment: tests, individual reports (one problem with an explanation) and weekly summary sheets.

Weekly Summary Sheet

1. Choose $<$, $>$, or $=$

$$298 \times \frac{99}{100} \bigcirc 298 \times \frac{191}{211}$$

Explain your choice:

2. Solve and show a diagram:

a. $\frac{1}{2} \times \frac{1}{4} =$

b. $\frac{2}{3} \times \frac{1}{5} =$

3. Complete the inequality so that the phrase will be correct:

$$193 \times \frac{14}{30} > 193 \times \square$$

Explain your answer:

From this example of a weekly summary sheet it can be seen that the students are being asked to think mathematically. They are required to produce “procedures with connections” (Stein, Smith, Henningsen & Silver, 2000, p. 12).

School H had a conventional test too, which included an investigation:

8. Check if the two equalities are correct, and complete the third:

$$\frac{1}{2} + \frac{1}{4} = 1 - \frac{1}{4}$$

$$\frac{1}{2} + \frac{1}{4} + \frac{1}{8} = 1 - \frac{1}{8}$$

$$\frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} = \underline{\hspace{2cm}}$$

Write three additional equalities like these.

Can you make a generalization from these equalities? If so, what is it?

As can be seen here, students are being asked to find a pattern and to produce additional examples. Then they need to verbalize it as a generalization. Certainly these are higher-order thinking demands.

Thus, I can say that the teachers are using some of the principles of the investigation assessment tasks in their work. One notable exception is the complete lack of any contextual situations.

7.1.3 Interviews

At the end of the school year, in June 1997, I interviewed five mathematics teachers, each from a different school in the Sharon community and none of whom was participating in the ongoing weekly workshops of the in-service project. Their school math coordinators recommended the teachers to be interviewed. With one exception, all had been present in the workshops given on assessment to all Sharon math teachers. What emerged from these interviews is that attendance in the three workshops alone had little effect on the views of teachers and on their teaching and assessment

practices. Although the teachers I interviewed talked about placing more emphasis on thinking, rather than on computations, their teaching methods had not changed and they still felt that the conventional tests with their emphasis on skills were sufficient. They all stated that they had not personally chosen a task from the Assessment Task Bank to use for assessment in their classes, and had just tried the tasks assigned in the workshops.

Each question that was posed in the interviews is presented below, accompanied by a summary of the responses, with examples of some of the teachers' statements. This is followed by a number of specific issues raised by individual teachers.

Did these workshops affect your teaching?

The responses suggest that while the teachers' way of thinking may have been modified somewhat by the workshops, this intellectual change found little expression in their teaching and assessment practices. The following quotations exemplify the effect of the workshops on these teachers' views:

I think that this whole thing is a process, it's not instant, not immediate. A serious change needs to occur to do such things. If you want an immediate change it won't be effective. It needs to work through a process. Clearly it does something in the head. That is, it's enough that it does something to my way of thinking and then I make steps to make it occur. And that is exactly what happened. My thinking has slowly changed, and as a consequence all sorts of processes in my class are beginning to change. (teacher EA)

It opened the door for me. ... but in these meetings, first of all, I very much enjoyed the investigations. Problems like these are the "spices" of learning and I think they also sharpen one's thinking. What I learned this year with your help and with Teacher AO's help [this teacher participated in a bi-weekly mathematics workshop given by a teacher-leader who was a previous student of mine and had attended the Third National Course on Assessment], was to think mathematically and this was very important to me. And very important to the children. (teacher EE)

I used a lot of what we discussed, I worked with the students especially on verbalization, which is quite hard for them, it's something they are not used to doing. (teacher ED)

Did these workshops affect your methods of assessing students' achievements?

Three of the five teachers mentioned specific ways in which their assessment practices had changed.

I pay more attention to the fact that I need to make the test material correspond to the material I teach. That means not to let three months pass and then to give some sort of test on material which the students have long forgotten...

... [This teacher also mentioned a specific practice which she had begun to implement.] Like the video we saw with two girls [a video using an investigation from the Assessment Task Bank as the focus for an assessment observation], there are children whose knowledge I now check this way. We sit together, and if they have trouble writing, then I give them all sorts of manipulatives and they have to show how they understand. For me, that is the most important part of learning. (teacher EC)

More questions which require mathematical thinking—not just on the material being learned, but in which the students can demonstrate their mathematical thinking and not just arithmetic skills. This is very important because what I discovered using this method [assessing with mathematical investigations] were students whom I had not known had mathematical thinking.

(teacher EA)

What was good was they worked with data [in the *Animals* task], they had data in a table and they needed to use that. And then we started to give them data and some ten related questions, and to check to see how they manage.

(teacher EB)

Like their effect on teaching methods (discussed above), the effect of the workshops on assessment is still in the realm of theoretical ideas, rather than actual changes in practices. These teachers have been given some ideas and they are beginning to think about them. It is impossible as of yet to determine whether these ideas will have a lasting effect on teaching and assessment practices.

How do you create tests?

Most said that they wrote tests by themselves, although if another teacher in the same grade was teaching the same material, she might also use the test. One teacher said that she worked closely with another teacher:

All the time together. She teaches two fifth grade classes and I teach the third fifth grade. So I sit with her and we create tests and assignments together. And not only that, periodically we work on the syllabus together—deciding what we want, how to build the lessons. (teacher EE)

One teacher (EA) described the kinds of conventional tests she used: “I give five to ten exercises and two word problems”. Since she did not express any reservations or doubts about her conservative approach, I interpreted her response as an indication that the workshops had not affected her in the least. The coordinator of her school confirmed this interpretation.

*Have you used the mathematical investigation tasks in the Assessment Task Bank?
What did you learn from using them?*

All the teachers responded that they had used investigation tasks, but when probed further it emerged that they had only used the ones assigned in the course of the workshops, but no others. Although they were aware that there was an Assessment Task Bank in the school and some teachers had even taken a look at the tasks in it, they responded that they would use these tasks only the following year.

One teacher could not even recall using the tasks from the course, as is evident in the following transcript of our discussion:

- T: I used *Animals*.
- I: So let’s talk about that for a minute. What did you learn from it?
- T: I don’t remember, I still haven’t analyzed it completely. I used another task but I don’t remember.
- I: Maybe *Table Arrangements*?
- T: Certainly! That one I brought to the following meeting.

I: But you don't remember?

T: It was a while ago. (T: teacher EA; I: interviewer)

Does your department work as a team? Do you create tests together? Do you talk about students' achievements?

The teachers talked about some work with colleagues who taught parallel classes. One teacher said that she wrote tests together with her colleague (see above). Discussions focused on the difficulty of the questions, rather than on students' achievements. They did not mention using the results to help them plan future lessons.

T: We build the test together.

I: And afterwards do you look at the results together?

T: Not really, not always. More or less she [the other teacher] says that in her class it was hard and I said that in my class also, and what was hard, etc. But not very much. (T: teacher ED; I: interviewer)

Teacher EB described the process as follows:

I usually approach [the other teacher] and tell her what was hard. Often if she uses the test a day before me, she tells me which question needs further explanation, etc. In addition, I tell her what I thought, like: I had thought it would be easy, but the question was very hard. Thus we sometimes change the grade or assessment.

What do your report cards look like?

All the teachers described report cards which related to students' skills. Some report cards had room for teachers to add comments, but none had any specific criteria to which teachers related.

We have grades and comments. For example, I write "very good" and follow with something like: "you can think mathematically". For the slower students, I don't write anything. (teacher EC)

Arithmetic and geometry are separate grades. The grade is not a number, but is given in words: "very good", "excellent", etc. And then two appropriate comments. That means, if I relate to topics, if I know that the child has mathematical thinking, then I write that. If I recently stressed decimals then I

write that he worked well on decimals, or that he needs to review the material on decimals, or that he has only partial competence in exercises, although he can solve word problems very well. (teacher EB)

Not grades. It goes by subject, for example, beginning fractions: “mastery”, “complete mastery”, or “has trouble”. Thus it’s difficult for me to be exact. (teacher ED)

In our school we don’t give grades on a report card, we just move the students up a grade. But we prepare a “sheet”, and for each subject we write details. For example, in mathematics, I have a notebook in which I jot down to myself what I write to them. ... And I write things like: “he manages nicely in addition and subtraction of fractions”, “he has trouble with borrowing”, etc. So the student gets that information when he goes home. I can’t write things like “good” or “very good”. What does “good” mean? What does “sufficient” mean? I give details. (teacher EE)

What emerges here is that even in those schools that do not distribute formal report cards and grades, the details are almost entirely related to competence in performing certain mathematical skills. This gives us yet another indication that for these teachers, mathematics is equated with skills.

Are there things you will do differently next year when you have new classes?

While some of the teachers indeed claimed that they would implement changes in their teaching and assessment practices, most had only begun the thinking process. I believe that without serious input—in the form of workshops and assistance within their classrooms—real change will not occur.

Special issues arising in the interviews

Each teacher raised in her interview issues that were particularly pertinent to herself. These concerns are described here, as they are pertinent to the subject matter of this research, i.e. assessment and reform mathematics teaching. They are presented here according to teacher.

Teacher EA

This teacher was still very concerned about teaching skills. She did not know what to do with the more advanced students: she was concerned that letting them progress in

the syllabus and even to the material of the following year would generate problems at a later date. She still had no sense of how to deepen their understanding of a topic. She continued to insist on dividing the students into groups according to their levels and on working with them on skills.

They learn to understand better the material being studied. For example, I don't give them all the material that I give to the whole class, and thus they are less frustrated.

In other words, she was the one who decided what each child would learn and how much he was capable of learning.

Teacher EB

This teacher did not take part in the assessment workshops. She was part of the on-going interdisciplinary project in Sharon as a language teacher for her school. This intensive language study was influencing her teaching in all areas, not just language.

Since I come from the "language" side, certainly something happened to me in relation to assessment, not just in mathematics but in everything. If we are talking about alternative assessment, while once I use to stress only tests or quizzes, well, now my methods are much more open. There are now reports, tasks, building games and student lectures, and in general, it is much, much more open. ...

In mathematics I built worksheets, for example, instead of a test; worksheets which I wrote and passed out to the students of the class. ...

That's it, certainly to check the child—not only his final result but the process by which he worked throughout the whole year. ... It comes from the fact that I'm in the language study group and also that I have an excellent team partner who is in the mathematics study group.

Teacher EC

This teacher was confused about whether children could think mathematically and what sorts of questions to use. On the one hand, she reported that she had tried some of the things I had suggested and was amazed that the students had succeeded:

It was an amazing thing with the numbers and the coins. I couldn't believe that the students were able to think! ... They made all sorts of conclusions, unbelievable, very very good.

On the other hand, her statements suggest that she was confused about what it means to pose "thinking questions" to children:

It certainly was interesting, the *Animals* page, because the strong students had trouble with it. They thought that there was a trick involved, they looked for the trick and didn't do what was asked. Afterwards, when I returned the papers to them and we talked about the task, I said to them: "Look how much you worked and there were instructions, why did you look for a trick?" They said that on my pages they always need to find the trick! But the task was very good, and one class, the fourth grade I think, succeeded on it quite well. There is no doubt that we will use the Assessment Task Bank.

Teacher ED

This was a teacher in her first year, who also participated in a workshop which met once every two weeks in the neighboring city. She received much support from this course and managed to implement many of the reform ideas she had heard about.

It's not the mathematics that we used to talk about where there was only one right answer. Today we look at thinking, and at the students' way of working.

Teacher EE

Although a very experienced teacher, teacher EE liked learning new methods.

Certainly I need more, I always need more, there's never enough. It strengthens me and what's most important is that you let us try alone and when one tries and experiences it then it's different. I'm a pretty standard teacher, I write everything. I've been a teacher for 19 years and I still write each lesson. Why do I write them? Because I return home and think ...

She participated in the same workshops as Teacher ED, and as a consequence used mathematical journals with her students.

Each child has a copy-book, and for the final five minutes of each math lesson I ask them a question or a problem which is connected to the lesson and which requires the child to reflect on what he learned in the course of the lesson. I collect the copy-books and return them the same day, immediate feedback. Thus, first of all, I know if something was wrong—it's feedback for me and also feedback for the child. And additionally, their writing improves over the course of time. We also talk about how to write in order to convince someone, because often I write that they haven't convinced me or that I didn't understand.

7.2 The Third Year of the Sharon Project

The following year 1997–1998, there were again 14 teachers participating in the in-service course—eight of whom were from the previous year. An analysis of their final project, which they were required to hand in during June 1998 provided a unique opportunity to determine the influence of the mathematical investigations on their assessment practices. As usually happens, different teachers were influenced in different ways. As mentioned in Chapter 3, this final project involved the collection of samples of questions that would enable the teachers to determine whether each student was high, average, or low on each of the criteria (understanding the task, understanding the concepts, mathematical thinking, calculations, communication, and creativity). I examined the projects submitted by each of the six mathematics coordinators.

Teacher ED, Math Coordinator at School O

This teacher knew what it meant to investigate worthwhile mathematical problems, and used such examples on her assessment tools. She required the students to explain their work by finding generalizations and justifying them. The one negative aspect of her assessments was the complete lack of contexts—all she had were numeral examples.

Teacher DA, Math Coordinator at School K

This teacher included four mathematical investigation tasks from the assessment task bank as her assessment tools. She analyzed the choice and use of each investigation task and the resulting students' work in a way which showed that she knew the difference between the various thinking skills and between mathematical thinking and computations. She chose to use investigation tasks prepared by experts, rather than writing her own tests.

Teacher DP, Math Coordinator at School U

This teacher continued to resist reform mathematics, and her work, not surprisingly, showed an emphasis on computation. Where there were questions that required the students to verbalize or to make minor generalizations, the questions were identical to those in the textbook. This meant that the items were routine questions; the students had already worked on them and needed only to remember what they did in the past.

One test included a bonus question, but it placed emphasis on skills of precision in drawing, rather than encouraging mathematical creativity. The question was as follows:

Here are five squares. [There are five squares (3×3 cm.) drawn on the test paper.] Draw one-quarter in each square, in a different way.

Only precise work, using a sharpened pencil and a ruler, will receive points.

Teacher DL, Math Coordinator at School H

This teacher used some problems which required multiple solutions, some which required non-algorithmic thinking, and some which required generalizations and justifications. Nevertheless, her tests remained conventional in that they included regular word problems, and each one tested the exact subject matter that was studied, e.g. division of simple fractions, and percentages.

Teacher DF, Math Coordinator at School A

This teacher seemed to be unsure what exactly it was that she wanted to assess. One test included 42 computation exercises and six word problems, which is too many for any student. Nevertheless, two other tests displayed new influences: some called for multiple solutions and others included questions which tested understanding.

Teacher DZ, Math Coordinator at School G

This teacher was new to the in-service project that year. Her tests showed some openness, but nevertheless, she concentrated on computation skills. Even in quite open questions, such as “Provide two fractions whose sum is larger than three”, she only required the students to write one solution. In addition, there were no words on the tests, and not even any standard word problems.

7.3 The Fourth Year of the Sharon Project

During the subsequent year, 1998-99, the last one of the Sharon project, I held monthly meetings with the school mathematics coordinators. Assessment was not a major topic discussed.

7.4 Follow-up Questionnaire and Interviews

At the end of the fourth year of the project, I administered a questionnaire to all mathematics teachers in the Sharon community. Three of the six math coordinators filled out this questionnaire. Almost two years later, during the spring of 2001, I interviewed four of the six math coordinators. The views and attitudes of each of these four coordinators are presented here, as they were expressed at the end of the project.

7.4.1 Teacher ED, Math Coordinator at School O

This teacher continually gave tasks with worthwhile mathematics, although her tasks were solely numerical ones, that is, not involving any real-life situations. She used a variety of types of assessment tools, including journal writing, individual reports and weekly summaries. She wrote on the questionnaire at the end of the in-service project that the project had helped her understand the importance of the student's thought processes, rather than merely the final solutions, and had introduced her to the various types of mathematical thinking and consequently to the criteria. She did not make use of the Assessment Task Bank.

At the final interview, almost two years later, she was still writing tasks that required verbalization and some generalization and justification. Now a second-grade teacher, she talked about introducing investigation tasks to her young students:

There are things that they didn't understand at all at the beginning, like finding patterns and making generalizations, and they still have difficulty with them, but I work a lot in this direction. ... and they enjoy it very much.

7.4.2 Teacher DL, Math Coordinator at School H

This teacher uses conventional assessment tools—tests, classwork and homework—although she says that the tests are written with the additional criteria in mind. On the questionnaire, at the end of the fourth year of the project, she noted that she now recorded different kinds of evaluations on the school report card, using the criteria for mathematics assessment. At that time she was not using the Assessment Task Bank extensively. She had given her fifth-grade students two investigative tasks from the fourth-grade bank, that is, investigations which were not closely related to the material they were learning at the time.

In the course of the interview, she continued to discuss the importance of the additional criteria for assessment, but also talked about the need to “cover” all the

topics required in the syllabus. This teacher was influenced by the workshops and the ideas of reform mathematics, but feels “pressured” by her fellow teachers, the junior high school teachers and parents to continue to stress basic computation skills.

7.4.3 Teacher DA, Math Coordinator at School K

This teacher started studying for a M.Ed. during the fourth year of the project, and chose to write her thesis on facilitating problem solving with fifth-grade pupils. This work focused her thinking and encouraged her emphasis on investigative tasks in her teaching. By the end of the fourth year she was using the Assessment Task Bank extensively in her classes for assessment purposes.

In the interview, she talked about the influence of the Assessment Task Bank on her teaching:

My extensive use of the Assessment Task Bank started when it was in the school and I could look at it frequently. Often, before I started to teach a new topic, I checked what was in the task bank and what I could use as a summary, or sometimes as a starting point. The Assessment Task Bank became a book of rules for me—like the Bible. It helped me determine what I needed to check, what the main points of the topic were.

In addition, she talked about how problems on a test would have to be changed. The types of questions in the Assessment Task Bank influenced her: instead of drawing geometric shapes for the students, for example, she would “first of all give them squared paper, and ask them to draw polygons—to see if they understand by drawing for themselves.”

7.4.4 Teacher DF, Math Coordinator at School A

This teacher indicated in her interview that she had continued to be influenced by the concepts taught in the project even after it ended. The test she brought to the interview included different types of mathematical processes, and not merely calculations. It not

only permitted the students to use manipulatives, but encouraged their use and required generalizations based on this use. There were also questions requiring verbalizations and justifications. She demonstrated understandings that she had not previously applied in her assessment tasks.

7.5 Conclusions of the Sharon Study

The first question that my research addressed was whether the teachers' use of mathematical investigation tasks for assessment purposes would influence their view of mathematics. The findings collected in the course of the Sharon study clearly indicate that their view and concept of mathematics has indeed been affected. Most teachers, after exposure to mathematical investigative tasks, began to talk about finding and formulating generalizations and justifying them, instead of focusing on computational skills.

The second question posed in my research asked in what ways would the teachers' use of mathematical investigation tasks for assessment purposes influence the way they teach. It turns out that there is more emphasis on verbalization in classrooms. Teachers are now using discussions in lessons (part of the textbook that they had previously ignored) and regularly used challenge problems.

The third research question asked whether the teachers' use of mathematical investigative tasks for assessment purposes would influence the way they assess their students. Here too, the answer is positive. The tests that most teachers in the Sharon community pose now regularly include problems with multiple solutions, some requiring non-algorithmic thinking, and questions which require generalizations and justifications. Verbalization has become a regular feature in the mathematics classrooms and is now an integral part of all tests.

The teachers in this community now routinely use additional criteria for assessing the performance of students, rather than merely evaluating the final answer they reach. In addition, these teachers express a desire that report cards be changed to reflect this usage.

CHAPTER 8: CONCLUSIONS

This study was an attempt to determine the relationship between good assessment practices and good teaching methods. I wanted to investigate whether making good assessment tasks with criteria for grading available to elementary-school mathematics teachers would have a positive effect on their teaching.

Research tells us that standardized tests influence instruction (Section 2.3). My research explored whether a national bank of mathematical investigative tasks for school-based assessment—tasks which contain “good mathematics”—could influence instruction in a positive manner. Was it possible by these means to influence the mathematics that teachers taught, the assignments they gave and the types of tests they wrote? The research presented here addressed these questions.

The first research question that this study addressed was:

Will the teachers' use of mathematical investigation tasks for assessment purposes influence their view of mathematics?

Prior to my research, I knew that the general view prevalent among elementary school teachers was that mathematics at the elementary school level consisted primarily of basic arithmetic calculations, and possibly the study of some geometrical shapes.

In Chapter Two (Section 2.2) I discussed the kind of mathematics we would like to have today, termed “reform mathematics”. The ideas of reform mathematics are based on a broadened vision of mathematics with emphasis on higher-order thinking. My research indicated that the use of mathematical investigation tasks helped the teachers in my study come to the awareness that mathematics, even on the elementary school level, is about more than just arithmetic exercises—that it involves generalizations, justifications and even creativity. This is evident, for example, in the responses of many teachers (e.g., teachers AX, AT and AQ in Section 6.1.1) who pointed out that their students found it difficult to formulate generalizations. Their words reflect the underlying notion that generalizing is an integral part of mathematics.

As expressed by teacher CO:

In mathematics, we want to work with the central ideas on which mathematics is based, and not with technical computations. (III.1)

Or as teacher BV wrote:

Less emphasis on basic mathematical skills and more emphasis on finding rules and generalizations (III.2; p.32)

As teacher ED from the Sharon study explained (regarding the use of investigations with second graders):

There are things that they didn't understand at all at the beginning, like finding patterns and making generalizations ... but I work a lot in this direction. (Section 7.4)

Thus, my research demonstrated that the Israeli teachers involved in this study now view mathematics in a new light, more in the direction of reform mathematics.

Schifter and Fosnot (1993, pp.193-194) discussed what is involved in influencing teachers' views of mathematics, by quoting a teacher who says, "It is really difficult to discuss how my beliefs about the nature of math and the learning of math have changed". Change in beliefs is an elusive thing, often hard to pinpoint. They go on to explain that while change may indeed be initiated during professional courses, "... the bulk of what teachers must learn will necessarily come only in their own classrooms with their own students". What is important, according to Schifter and Fosnot (1993), is "inaugurating a complex process of active reflection on what is involved in serious exploration of mathematical concepts, on how to think about student learning, and on the advantages and limits of different kinds of classroom structures". Furthermore, they suggest that this process occurs if long-term support— both technical and emotional—is provided.

Thus, it appears that the teachers and teacher-leaders who participated in the various workshops which provided the data for my research will continue to change, especially if they receive on-going support. This process is evident, for example, in the words of

teacher DF, the math coordinator at School A in the Sharon community, who remarked that since the end of the project she had continued to be influenced by the things she had learned (see Section 7.4). The test she had brought to the final interview involved various types of mathematical processes, and not merely calculations. It not only permitted the students to use manipulatives, but encouraged their use and required generalizations based on this use. It also included questions that required verbalizations and justifications. The way she expressed herself demonstrated understandings that she had not previously applied to her assessment tasks.

My second research question was:

Will the teachers' use of mathematical investigation tasks for assessment purposes influence the way they teach, and if so, in what ways?

Prior to my research, I knew that teachers invested most of their efforts into teaching routine procedures and that the work sheets prepared by teachers for the most part involved single-answer exercises. My research indicated that the use of mathematical investigation tasks indeed influenced the way teachers teach. The importance of discourse in the study of mathematics has been noted in the literature (Kazemi, 1998, NCTM, 2000). For all the teachers who participated in my study, verbalization—having the students explain “why”—has become an integral part of their teaching. See, for example, comments by coordinators of the Sharon project (Section 7.4).

As discussed in Chapter 2 (Section 2.2), the mathematical tasks that the students are given have a major influence on the kind of mathematics they learn. Nowadays, the Israeli teachers that I worked with look for investigation tasks to use in their classrooms—often referred to as “authentic tasks”, since they deal with real-life situations that involve some mathematics. Unfortunately, sometimes the tasks are not planned properly, and the situation itself becomes much more important than the mathematics involved. In this case the mathematics remains on the level of simple

computations or the students get so involved in projects such as writing letters to carpenters that they “forget” what is really important (Birenbaum, 1999, p.163).

Flewelling and Higginson (2000) are concerned about this problem as well:

Put simply, most of these tasks do not give students the opportunity to learn very much of value. They create an illusion of learning, an illusion of relevancy. At the completion of a set of such tasks, students add little to their pool of conceptual and procedural knowledge. Their ability to select, or adapt, tools and procedures and fit them to a situation, in a purposeful and integrated fashion, is scarcely improved. Their problem solving, reasoning, critical thinking, and communication skill sets are neither enlarged nor refined. (p.32)

Thus, one finding that emerges from my research—one that is indeed supported by the literature (e.g., NCTM, 2000; Hiebert et.al., 1997)—is that the type of tasks chosen by the teachers is crucial. The problem becomes even more acute when teachers—and even teacher leaders—download tasks from the internet or write their own tasks, but lack the ability to distinguish between tasks that contain “worthwhile” mathematics and tasks which may be “authentic”, but whose mathematical content is minimal. The source of this problem may be the teachers’ lack of mathematical knowledge, as described in Chapter 2 (Section 2.4.2). Without a broad picture of what it is that constitutes mathematics, the teachers seem to be unable to make this important distinction.

This year the Israeli Ministry of Education implemented a new mathematics curriculum for elementary schools which places emphasis on number sense and learning through investigations. When I talk to teachers and teacher-leaders who have been exposed to the mathematics Assessment Task Bank and/or who have participated in my professional development courses, they frequently express surprise at the fact that these ideas are considered innovative. I often hear comments such as: “But this is what we have been doing already” or “This is what the assessment task bank taught us”. In other words, the teaching methods that are now being implemented in the new

mathematics curriculum are already being taken for granted by teachers who have been exposed to the Assessment Task Bank. This is a clear indication that the answer to my second research question is positive: that exposure to the task bank indeed had an effect on the way teachers teach.

My third research question was:

Will the teachers' use of mathematical investigation tasks for assessment purposes influence the way they assess their students, and if so, in what ways?

Prior to my research, I knew that teachers were not using mathematical investigation tasks as assessment tools. For most teachers a standard test would consist of many unconnected short-answer questions. Most test questions required only a single answer and the questions generally tested for knowledge, rather than for understanding.

My research demonstrated that teachers attending my professional courses found the mathematical investigation tasks to be useful for assessment purposes; they found that the tasks provided them with additional information about their pupils, which they could not obtain through conventional assessment methods. The additional criteria used for evaluating the pupils' work aided in defining these additional areas. As teacher AW wrote:

The results of the tasks illustrate clearly what students know well and what I need to work on more ... The grading of the questions enabled me to gain a better understanding of where each child's difficulties lie.
(I.1)

As put by teacher AQ:

Because of this course I have begun to learn to assess differently. I've started to relate to the process and not the final answer, to vary the thinking, to emphasize the creativity of the students.

In tests, I've been opening tasks and I try to vary their levels. (III.2)

Another significant finding of my research related to the way these teachers used the tasks for assessment: they were quite willing to use the mathematical investigation

tasks to supplement the conventional tests, but were reluctant to use them as replacements. Teacher AQ wrote:

These tasks are good for assessment on condition that they are used in addition to (conventional) tests and not instead of tests. (I.1)

Teacher BW also felt that there is “a need for alternative assessment in addition to conventional assessment” (III.2). In other words, investigative tasks are helpful, but can not replace the need for conventional tests.

However, despite the promising findings of my research, which suggested a modicum of change in initial attitudes towards assessment, talks that I have held with teacher-leaders now—several years after the national courses on assessment—suggest that they do not readily use the tasks of the Assessment Task Bank for assessment purposes (see, for example, the interviews at the end of Chapter 6, Section 6.4.3).

Despite the possibility that long-term change is still elusive, I did find that exposure to the Assessment Task Bank with its mathematical investigations influenced, to a certain degree, the way the teachers in my study assessed. The tests of the teachers in my study now regularly include elements which were previously absent: questions requiring explanations and questions with more than one possible answer. These changes are reflected in the test/task analysis of Chapter 6 (Section 6.4.4) and in the follow-up interviews of the Sharon project (Chapter 7, Section 7.4). This process has been supported by the changes in the Israeli national achievement test (see Section 8.1).

8.1 Implications for Teaching and Learning

Assessment is an integral part of reform mathematics. Consequently, if the teachers' methods of assessment have not changed, then it is reasonable to assume that their view of mathematics and their teaching methods have not changed either. As quoted in Chapter 2 (Section 2.3), “It is through our assessment that we communicate most

clearly to students which activities and learning outcomes we value” (Clarke, Clarke and Lovitt, 1990, p.118). Thus monitoring and helping teachers change their assessment practices is critical for reforming instruction.

My research found that teachers seem to need some external force to convince them—or remind them—to use the Assessment Task Bank. One such “external force”, for example, was a regional achievement test containing an investigation task which was held in one area of the country last year, creating a demand for obtaining and using the Assessment Task Bank for practice purposes. Similarly, Grant, Peterson and Shojgreen-Downer (1996) found that teachers’ responses to reform policies were affected by their school’s standardized testing policy. Their findings came from case studies of three California teachers in the context of system reform. Aschbacher found that “use of alternative assessments by teachers ... requires a significant paradigm shift” (1993, p.27).

There is much research that suggests that the changes needed for implementing reform mathematics—which is directly related to alternative assessment practices—fail unless they receive support from national assessment systems (see Section 2.3). Indeed, in recent years the Israeli Ministry of Education has implemented changes in the national achievement tests, delineated below, which have contributed to the trend towards reform mathematics and have provided support for the implementation of alternative assessment practices.

In 1989, a countrywide assessment and evaluation of fourth-grade mathematics was undertaken by the Israeli Ministry of Education. The test, which was given to only a sample of schools, consisted of 40 multiple-choice items, mostly requiring only computations. In June 1996, a sample of fourth graders were again tested. This test had a completely different format, although it continued to be confidential and administered by proctors, rather than by the classroom teachers. The most striking change was the fact that the items were open, and not multiple-choice questions. In addition, there was a scaffolded investigation question. In two-thirds of the test the

students were allowed to use calculators. And although many items continued to test knowledge, the fact that the student had to create the answer himself, rather than choose it from a selection, conveyed a different message to the teachers.

The achievement tests administered in March 2001 once again assumed a different format. This time the teachers were in charge of administering the test and correcting it. The use of calculators was permitted throughout the test. And the format again had become more open. There were questions which required explanations: "Explain how you reached the solution" or "Justify your answer". Once again there was a scaffolded investigation question. In addition, there were some questions which required more than one solution.

Although the teachers of my study were increasingly opening up their tests and using questions that required higher-order thinking, the tendency continued to be to use the tests as summative assessment, rather than formative assessment. In other words, many teachers found it difficult to use test results for planning their subsequent lessons. While they are able to analyze their students' work and could report in some detail on each student's performance, they failed to understand how this should affect their teaching. Before they were exposed to the tasks, the teachers were often not even trying to understand the students' thinking. A test is administered in order to provide a grade, and then one goes on to teach the next unit in the syllabus. This is an issue which still requires much work.

Why have these changes occurred? Research tells us that the change process is quite complex and often very threatening, and that change occurs very slowly (Friedlander, Bruckheimer & Albert, 1987; Fullan, 1999).

I think that the situation regarding assessment in Israel helped the change process. First, although as discussed above, national testing policy is crucial in affecting the willingness of teachers to implement reforms. The fact that we generally have no high-stakes testing on the elementary-school level in Israel has enabled the teachers to practice using the assessment tasks without an accompanying fear of failure. In the

professional courses on assessment, we continuously discussed the use of the tasks for “teachers’ purposes”, rather than for student grades. The positive reactions of the pupils to the tasks, even when they did not succeed completely in solving them, helped convince the teachers to try these tasks.

The fact that the results of the pupils’ work remained within the confines of the school and were not reported to the educational authorities was an additional factor which aided in the implementation of change. Teachers discovered that the results were helpful to them, both in evaluating their pupils and in planning their future classes, and that nobody would criticize their teaching on the basis of the results. Thus, they were willing to use the tasks from the Assessment Task Bank.

Another important factor which helped influence the teachers was the professional development courses I provided. In Chapter 2 (Section 2.4.2), I described principles I use for planning such courses. The results of Chapters 6 and 7 demonstrates that such courses may be quite effective. It is worthwhile reiterating three major points regarding such courses:

1. The teachers were provided with good materials and investigated their use. However, they were not required to create the materials themselves.
2. The use of assessment as a focus is one way to encourage teachers to delve into students’ thinking. There are other methods, but assessment is an effective one.
3. The courses were built according to a constructivist model—that is, looking at the teachers’ needs and finding ways of meeting them. None of the courses had a pre-determined syllabus; the content was modified with accordance with the needs that arose. While the overall goals of a professional course may be pre-ordained, the immediate steps *en route* to achieving these goals must be negotiated with the participants.

8.2 The Limitations of the Study

One of the limitations of my study is that it involved work only with Israeli elementary school teachers. The findings, therefore, may be inapplicable to other levels of schooling. This is certainly true of the Israeli school system, where senior high school mathematics is highly affected by school-leaving examinations, leading to reluctance on the part of the teachers to innovate without “orders from above”.

The teachers in my study participated on a voluntary basis, although many were encouraged to do so by their local Ministry of Education supervisors. This led to two requirements: 1) the professional course had to be interesting in itself, in order to keep them coming; and 2) the course had to conform with their attitudes sufficiently so that they could accept its ideas. I am reminded again of the parallel between working with teachers and working with students and of the need to work with students within their “zone of proximal development”, as suggested by Vygotskii, in order for learning to occur. If the ideas I present are too remote from the participating teachers, they will be unable to accept them.

The fact that the teachers I worked with were Israeli and that we were working within the Israeli school system should not limit the transferability of my findings. A research study conducted in the United States, whose purposes were quite similar to mine, is described by Fuchs, Karns, Hamlett & Katzaroff in *Mathematics Performance Assessment in the Classroom: Effects on Teacher Planning and Student Problem Solving* (1999, p.613). These researchers also suggested that “with routine use of classroom assessments that reflect challenging, authentic tasks, teachers’ knowledge about and understanding of the reform curriculum will increase, their instructional plans will begin to reflect a new vision of mathematics education, and student capacity to engage in mathematical problem solving will improve.” Some of their workshops were similar to mine, although they write about only eight teachers and the duration of their study was only one year. Their conclusions were similar to mine:

Participation in classroom-based PA-driven [performance assessment] instruction did appear to shape teachers' understanding of assessment, curriculum, and instruction in substantial and desirable ways. (Fuchs et.al., 1999, p.635)

My findings, like theirs, are limited in that I did not carry out classroom observations and relied solely on teacher questionnaires and interviews. Fuchs et al. (1999, p.635), describing this problem, quoted Firestone, Mayrowetz & Fairman (1998) who "showed that classroom observations of mathematics programs corroborated teacher reports".

The importance of good tasks is discussed above. Stein and her colleagues from the Quasar project researched the use of higher-order thinking tasks in the classroom and found two occasions where the cognitive level of the tasks declined (Stein et al., 1996). Observing actual classroom practice may have improved my understanding of how the teachers utilized the investigation tasks made available to them.

The main limitation of my study, therefore, is that it is based totally on the teachers' own perceptions of their behavior, rather than on their actual behavior. Nevertheless, as explained in Chapter 3 (Section 3.5), the reliability of a naturalistic inquiry like mine depends upon the length of time and the amount of contact between the researcher and the participants. Since my research continued for a period of five years and since I had extensive contact with the participating teachers, I believe that my findings are reliable despite the caveat mentioned above. Throughout the period of the research, I talked extensively with the teachers, not just during the workshops but also informally by telephone. In addition, I met them on other occasions, such as at mathematical leader conferences. I was definitely a participant-observer in their community, a fact that was extremely advantageous in terms of assessing the reliability of their reports.

Another limitation of my study the inability to determine just how much mathematics the teachers actually learned. Research informs us (Section 2.4.2) of the importance of the teachers' confidence in their subject matter for their ability to teach "reform mathematics". While my professional courses included mathematical enrichment for the teachers, I do not know how much their subject matter knowledge increased.

8.3 Areas for Further Research

There are issues that still need to be resolved. The mathematical education community is still asking itself what mathematics elementary-school teachers need to learn and what kind of professional development can help them learn it. This subject was discussed at the Mathematics Teacher Preparation Content Workshop two years ago at the National Academy of Sciences in Washington, D.C. (MSEB, 2001). A conference on this topic is being planned in Israel for February 2002.

An additional issue is how to improve student learning. The disappointing results of the TIMSS tests (IEA's Third International Mathematics and Science Study, 1997) are of concern in Israel (in the United States there is concern too) (Harmon et al., 1997). We have begun trying to learn from the Japanese and implement study lessons as a way to encourage more staff cooperation and in-school professional development (Yoshida, 1999). This too is a subject for future research.

It has recently been decided by the Ministry of Education to develop specialist math teachers for the Israeli elementary schools. Each school is expected to choose three to five teachers who will participate in an intensive in-service program (four hours weekly over the course of three years) and will concurrently teach mathematics in two to three classrooms (each six hours a week). These responsibilities are in addition to the teacher's regular duties as a "home-room teacher", in charge of the education and social development of one particular class, as well as teaching this class one or two minor subjects. (Israel elementary school teachers spend approximately 24 teaching hours a week in the classroom, in addition to preparation and other duties.) In the coming school year, 2001–02, approximately 3,000 teachers from some 650 elementary schools will be involved in such a program. I will be working at the National Resource Center for Elementary School Mathematics, whose staff will be acting as advisors for the teacher-lecturers of this project and running monthly workshops for these teacher educators. One of the challenges I will face in this capacity will be how to utilize my knowledge about professional development. Another

goal of the resource center will be to compare the various programs that will cater to these teacher educators, from the thirty institutions that will provide in-service courses in the framework of this project.

Another issue that arose in my study was the conflict between reform mathematics and traditional report cards (see, for example, the interviews held with the teachers in the Sharon Project, Section 7.1.3). Graue and Smith (1996), discussing changes in assessment that occurred along with changes in instruction, found a number of constraints:

A major bone of contention for these teachers was the struggle they waged against student and parental beliefs about the nature of mathematics and its instruction. ... The idea of the report card lurked in the background as they looked over student learning and growth. (pp.130, 132)

This problem was one that the teachers who participated in my research also experienced. The new teaching and assessment methods to which teachers are being exposed have led to complaints about the need to give traditional numerical grades in mathematics. Many schools have begun developing new types of report cards. The report card developed by one school, see below, for example, includes a table in which teachers rate their pupils. It is obvious that this report card has been influenced by the additional criteria used in the Assessment Task Bank:

Arithmetic

	Very good	Good	Adequate	Has difficulty
Understanding the task				
Understanding the concepts				
Computational skills				
Verbalization				
Creativity and diversity				

Geometry

	Very good	Good	Adequate	Has difficulty
Understanding the task				
Understanding the concepts				
Creativity and diversity				

A further issue for research relates to the criteria for assessment. Today there are many examples of rubrics (for some of these, see Sections 4.3.1 and 4.4.2), but no research has yet been carried out to identify which bear the strongest influence on teaching and learning mathematics.

An additional issue for research is the connection between the change that teachers are undergoing on an individual level and changes that the entire school staff experiences. I worked with individual teachers and teacher-leaders and did not identify the optimal conditions to enable them to influence all the staff members of their schools. What qualities does a teacher-leader need to enable her to mobilize the staff of a school?

A related issue is the role of school principals in changing programs: how can they best introduce reforms, such as alternative assessment for mathematics (Nelson, 1998). This is a major import if the principal is not a mathematics teacher. And yet, in Israel, like in other countries, the principal of an elementary school is the central person who enables reform to progress.

8.4 Discussion

The most striking thing I continually learn, both from my research and from my teaching, is how similar certain processes are. For example, there is a definite parallelism between my naturalistic research methodology and the desired methods for assessing students' achievements. That is, as I wrote in Section 3.5, "adopting the

naturalistic inquiry paradigm meant that my research was based not on statistical analyses of teachers' responses to questions or other 'objective' data, but instead, on a holistic approach to the teachers' reactions to the assessment workshops and the tasks made available to them." Reform mathematics asks teachers to assess their students' achievements in a similar fashion—not using conventional tests which can be analyzed in a statistical manner, but rather using a holistic approach based on varied assessment tools which each contribute to learning about different abilities of the student. (See Hiebert, 1998, for a related idea.)

Another parallel which I find most often and which is described in my thesis, is the use of constructivist approaches to teaching with teachers. We need to practice what we preach; we should be working with teachers the same way we teach them to work with student. This means, for example, building social and socio-mathematical norms, requesting different solutions to problems and creating a supportive environment. Often professors think that adults need to be taught in different ways, generally by more lecturing (because there is lots of material "to cover" and not enough time in the course). This is certainly not so! Maybe the students are older, and maybe the situations are different (e.g., didactical problems sometimes instead of mathematical ones), but the teaching methods need to be the same!

Another idea which I embraced completely in my research is Gattegno's notion of "subordination of teaching to learning" (1988). This should be interpreted, to my mind, within the framework of the naturalistic inquiry paradigm (see Section 3.5), as "getting out of the way so that the students can go about the business of learning". As a teacher, I therefore force myself to "stop talking", remove myself from center stage and allow the students to dominate the lessons. This holds true for all the frameworks in which I teach, whether with young children or with teacher-leaders—they all require the same basic conditions to enable them to "be learners".

Mason, in his plenary talk at PME in Lisbon 1994, discussed how teachers often go into the classroom already armed with the solutions to investigations; this prevents

them from hearing the students, who may come up with different ways of solving the problems. This strikes a resonant chord within me, as I have often contemplated this problem and have come to the conclusion that it is better to **avoid** investigating a problem before a workshop. This forces me to listen carefully to the teachers and to really hear what they have to say, without being hindered by my own preconceptions. In a workshop that I gave last year, I had teachers work in pairs together with pairs of students on an investigation that they had **not** solved beforehand. They found the experience to be very helpful, although they were not sure whether they were ready to face a full class “unprepared”.

Confrey (2000), discussing the need for reforming math education in the United States, claims that “developing effective approaches to teachers’ professional growth is the thorniest and most troublesome challenge in reforming education because classroom teachers are the key to good instruction” (p.99). Professional programs, she claims, do not address this need effectively:

Two-hour or one-day workshops on trendy topics still dominate professional programs, and when given opportunities to engage in intellectually challenging professional development, too many teachers shy away. Furthermore, much of what is provided in teachers’ preparation programs or in-service training is weak, poorly informed by research or content expertise, and only minimally connected to classroom practices. Even when teachers are given strong in-service activities, frequently they are under no obligation to implement the methods in their classes. (Confrey, 2000, p.99)

In my opinion, the national courses on assessment that I conducted and the work that I performed in the Sharon community met with success because these programs avoided the pitfalls enumerated by Confrey. Unlike the professional programs she criticizes, my courses were long-term, were intellectually challenging for the teachers and were based on the findings of research. Most importantly, the courses bore strong connections to the teachers’ work in their schools. Over and over again, the teachers tried out the

assessment tasks in their classrooms, analyzed the results and planned future lessons on the basis of the students' knowledge.

Heinz, Kinzel, Simon and Tzur (2000) include self-reflection as a key component of learning.

Students' activities provide them [teachers] with records of experiences upon which they can reflect, allowing them to abstract regularities in the activity-result relationship. These abstracted regularities (patterns) are the basis for new or modified conceptions. (p.85)

Philipp, Flores, Sowder and Schappelle (1994) discuss the conceptions and practices of exceptional mathematics teachers. Their findings also focus on reflection as a central component in success:

Together with their commitment and willingness to undergo change, this reflective attitude of mind has very likely helped them judge their successes and failures realistically and formulate instructional decisions appropriate for themselves and for their students, without climbing on bandwagons or becoming unduly discouraged along the way. (p.176)

The main condition required for the development of reflection is, in their opinion, time.

My professional courses have all included reflection as an integral component. I have found that teaching how to reflect means providing time for it, as an integral part of all workshops and **not** as a homework activity. In general, teaching on the elementary level in Israel—as elsewhere—involves a major time commitment in terms of hours of preparation, as well as actual hours in the classroom. Thus, I found that if I wanted teachers to practice new skills, I needed to allocate time for this in the workshops themselves. These skills ranged from reflecting to analyzing student work, from reading a research article to “opening up problems”. Until such activities became second nature for the teachers, they needed support in carrying them out, and this—combined with the time element—meant that the activities had to be performed during the workshop, making it possible for them to receive support from me, and more importantly, from their colleagues, the other participating teachers.

Clarke (1997) discussed the changing role of the mathematics teacher associated with the use of innovation materials, describing the learning process that the teachers involved in his project underwent:

In the same way that reform documents describe student learners as active constructors of mathematical meaning in a social context, in this study teachers, too, were viewed as learners who actively create knowledge through experience and through reflection on that experience, in large part through communication with other teachers.
(p.283)

This is in accordance with my own methods. Clarke also discussed the important role of reflection in the process of change, which in his study was facilitated by the presence of the researcher in the teacher's classroom. Like Schifter and Fosnot, he emphasized that "change is a gradual, difficult, and often painful process", noting that one teacher exhibited little professional growth, evidence that "nothing can be guaranteed, even when many factors supporting professional growth are present" (p.303).

Simon (1995) discussed the need for teachers to have a well-defined conception of mathematics teaching: "teachers would need to develop abilities beyond those already currently focused on in mathematics education reform, particularly the ability to generate hypotheses about students' understandings ..." (p.142). In my opinion, the extensive work that my teacher-leaders did with investigative assessment tasks helped them develop these abilities.

Mason (1994) noted that there is "a strong sense in which the most important effect of educational research is on the researcher themselves. What a researcher finds out most about is themselves ..." (p.179). I feel that my research indeed had this effect. Although in many ways the teachers with whom I had worked underwent change—both in terms of their view of mathematics and in terms of their teaching and assessment practices—I was the person who had changed the most during the five years of the research. I proved to myself that I was capable of undertaking a research project, and most importantly, that I was able to write the results. Throughout my

schooling, writing had been difficult for me and when I had completed my first degree I learned that I was dyslectic, that there were objective reasons for my writing difficulties. One of the reasons I had chosen to study mathematics and to become a math teacher was because there was no writing involved. Like many of the teachers with whom I worked on investigations, I have had to learn to include verbalization skills in my teaching—and this has not been easy for me.

Baumann (1996) wrote about his experience in balancing teaching and research:

By struggling with ways to integrate inquiry into their work, teacher researchers come to know themselves better as teachers and persons, learn to understand their students and families in ways heretofore unknown, increase their professional esteem and credibility, share their learning with colleagues locally and beyond, and, most importantly, help their students develop intellectually, socially, and emotionally. (p.35)

For me too, this balancing act was an issue. I like to teach. I enjoy the interaction with my students, young or old. The process of research, which requires stepping back, reflecting and writing has been a continuous “battle” for me, although in the end I came to enjoy it. My natural instincts were to be involved continually in the “doing”, rather than taking the time to analyze in depth why certain things succeeded and how I could share this learning with my colleagues. This crucial element of doing research seems to be important for teachers too.

Meloy (1994) wrote about many of the issues with which I was struggling. I believe that the title of her book, *Writing the Qualitative Dissertation—Understanding by Doing*, is exceptionally pertinent. I too deal with the question of how much students learn by doing, and in my thesis I have attempted to demonstrate that teachers and teacher-leaders also need to learn in this way. Meloy has expanded this feeling to researchers too. The subtitle of the first chapter, “The End is the Beginning”, is particularly apt, for she suggests that the novice qualitative researcher gains a degree of certainty and confidence only at the end of the process. This again has been my

experience: now that I am completing my research, I understand how it should have been done!

Meloy describes how the thesis often gets written in a non-linear fashion: despite the neat, tidy appearance of the final product, the thesis appears to be an artifact of the research process, rather than the result of any *a priori* guidelines followed by the researcher. The processes of pulling together, shifting, organizing and writing our thoughts are challenging, because the “coming together” occurs in non-linear, halting manner and in multiple ways. The task of choosing which strands to pursue, when to pursue them and how to pursue them, as well as how to organize them and write them up is a difficult one, recurring throughout. And in the words of Paul Simon’s song, “the nearer your destination the more you’re slip-sliding away”.

Rhine (1998) wrote about the role of research in professional development, concluding that what was crucial was not “teachers’ understanding of the development of students’ thinking ... but rather, their objective of inquiry into students’ thinking” (p.29). I believe that my teachers, by using investigation tasks for assessment and analyzing the resulting student work, began to look more closely at the way their students think. This process led to meaningful learning, and indeed, Rhine wrote that “when teachers are diligent about assessing students’ meaning making within activity, they can become cognizant of opportunities to transform meaningless activity into meaningful learning.”

Rhine quotes a Chinese adage which I find helpful too: “Give a person a fish, and she eats for today. Teach a person how to fish, and she eats for a lifetime.” I find that Israeli teachers who come to in-service workshops want to go home with work sheets that they can use in their classrooms the next day, and often judge a workshop by the quantity of pages they receive. My policy is to use the fishing adage, with an adaptation. I tell these teachers that I will teach them how to fish, so they can “eat for a lifetime”, but that meanwhile I will give them some fish, so they will not be hungry while they are learning. This means that each workshop must contain two kinds of

activities: one involving long-term learning and the other with immediate application. It has been my experience that the teachers are willing to accept this method.

Driscoll and Bryant (1998) described the stages of concern that adopters of innovations typically experience. They felt that “professional development focused on mathematics assessment is a necessary but not sufficient support for changes in classroom practice” (p.42). This is my sense too: that without the follow-up of our new curriculum, my teachers would have been unwilling to invest in large changes in their classroom practice.

A final question which should be raised relates to Boaler’s findings, as reported in *Experiencing School Mathematics* (1997). She compared two secondary schools in London, describing how teaching skills alone, the method pursued in one school, leads only to the ability to pass tests, whereas the project method employed in the other school develops students who can use mathematics. In her words,

The findings of this study illustrate the inherent complexity of the learning process and, crucially, that it is wrong to believe that assessments merely indicate whether a student has *more* or *less* knowledge. (p.143)

She adds:

All of these findings indicate that the most important aim for teachers should be to engage students and to provide worthwhile activities that they find stimulating. (p.146)

This comment should apply to all our teaching, not just the teaching of mathematics, and to all learners, teachers as well as students.

REFERENCES

- Albert, J. & Friedlander, A. (1995). Collection of mathematical assessment tasks—Grade 4. Tel Aviv: Center for Educational Technology. (in Hebrew).
- Albert, J. & Friedlander, A. (1997). Collection of mathematical assessment tasks—Grades 5 and 6. Tel Aviv: Center for Educational Technology. (in Hebrew).
- Ann Arbor Public Schools. (1993). Alternative assessment—Evaluating student performance in elementary mathematics. Palo Alto, CA: Dale Seymour Publications.
- Arcavi, A. (2000). Problem-driven research in mathematics education. Journal of Mathematical Behavior, 19, 141–173.
- Aschbacher, P. R. (1993). Issues in innovative assessment for classroom practice: Barriers and facilitators. Los Angeles, CA: National Center for Research on Evaluation, Standards, and Student Testing (CRESST).
- Assessment options for Houghton Mifflin mathematics. (1995). Boston, MA: Houghton Mifflin Company.
- ATM Working Group on Assessment. (1989). Exploring assessment—An ATM discussion book. Derby: Association of Teachers of Mathematics.
- Australian Educational Council. (1991). A national statement on mathematics for Australian schools. Victoria, Australia: Australian Educational Council and Curriculum Corporation.
- Avital, S. M. & Shettleworth, S. J. (1968). Objectives for mathematics learning: Some ideas for the teacher. Toronto, Canada: The Ontario Institute for Studies in Education.
- Balanced assessment for the mathematics curriculum—Elementary portfolio support package. (1995). Cambridge, MA: Educational Technology Center, Harvard Graduate School of Education.
- Ball, D. L. (1991). Research on teaching mathematics: Making subject matter knowledge part of the equation. In J. Brophy (Ed.), Advances in Research on Teaching, Vol. 2 (pp. 1–48). Greenwich, CT: JAI Press.

- Ball, D. L. (1997). What do students know? Facing challenges of distance, context, and desire in trying to hear children. In B. J. Biddle *et al.* (Eds.), International handbook of teachers and teaching (pp. 769–818). Dordrecht, The Netherlands: Kluwer Academic Publishers.
- Ball, D. L. & Cohen, D. K. (1996). Reform by the book: What is—or might be—the role of curriculum materials in teacher learning and instructional reform? Educational Researcher, 25 (9), 6–8, 14.
- Ball, D. L., Lubienski, S. T. & Mewborn, D. S. (2001). Research on teaching mathematics: The unsolved problem of teachers' mathematical knowledge. In V. Richardson (Ed.), Handbook of research on teaching (4th ed.) (pp. 433–456). Washington, DC: American Educational Research Association.
- Barber, J., Bergman, L., Goodman, J. M., Hosoume, K., Lipner, L., Sneider, C. & Tucker, L. (1995). Assessments for great explorations in math and science. Berkeley: GEMS—Lawrence Hall of Science.
- Barnes, M., Clarke, D. & Stephens, M. (2000). Assessment: The engine of systemic curricular reform? Journal of Curriculum Studies, 32, 623–650.
- Baron, J. B. (1996). Developing performance-based student assessments: The Connecticut experience. In J. B. Baron & D. P. Wolf (Eds.), Performance-based student assessment: Challenges and possibilities (pp. 166–191). Chicago: University of Chicago Press.
- Baumann, J. F. (1996). Conflict or compatibility in classroom inquiry? One teacher's struggle to balance teaching and research. Educational Researcher, 25 (7), 29–36.
- Baxter, G. P., Shavelson, R. J., Herman, S. J., Brown, K. A. & Valadez, J. R. (1993). Mathematics performance assessment: Technical quality and diverse student impact. Journal for Research in Mathematics Education, 24 (3), 190–216.
- Bazzini, L. (1993). The teaching/learning process and assessment practice: Two intertwined sides of mathematics education. In M. Niss (Ed.), Cases of assessment in mathematics education—An ICMI study (pp. 99–106). Dordrecht, The Netherlands: Kluwer Academic Publishers.

- Ben Simon, A. (1998). National feedback for the education system— Achievement test in mathematics for fourth grade students, June 1996. Jerusalem: Ministry of Education, Office of the Chief Scientist; National Center for Testing and Evaluation. (in Hebrew).
- Birenbaum, M. (Ed.). (1999). Mindful assessment—From theory to practice. Tel Aviv: Tel Aviv University. (in Hebrew).
- Black, P. (1998). Testing: Friend or foe? Theory and practice of assessment and testing. London: Falmer Press.
- Black, P. & Wiliam, D. (1998). Inside the black box: Raising standards through classroom assessment. Phi Delta Kappan, 80, 139–148.
- Blane, D. (1992). Curriculum planning, assessment and student learning in mathematics: A top-down approach. In G. C. Leder (Ed.), Assessment and learning of mathematics (pp. 290–304). Hawthorn, Victoria, Australia: Australian Council for Educational Research Ltd.
- Blume, G. W. & Heid, M. K. (Eds.). (1993). Alternative assessment in mathematics. University Park, PA: Council of Teachers of Mathematics.
- Boaler, J. (1997). Experiencing school mathematics—Teaching styles, sex and setting. Buckingham, England: Open University Press.
- Brandt, R. S. (Ed.). (1992). Performance assessment—Readings from educational leadership. Alexandria, VA: Association for Supervision and Curriculum Development.
- Bright, G. W. & Joyner, J. M. (Eds.). (1998). Classroom assessment in mathematics—Views from a national science foundation working conference. Lanham, MD: University Press of America, Inc.
- Brooks, J. G. & Brooks, M. G. (1999). In search of understanding: The case for constructivist classrooms. Alexandria, VA: Association for Supervision and Curriculum Development.
- Brown, M. (1993). Assessment in mathematics education: Development in philosophy and practice in the United Kingdom. In M. Niss (Ed.), Cases of assessment in mathematics education—An ICMI study (pp. 71–84). Dordrecht, The Netherlands: Kluwer Academic Publishers.

- Brown, M. (1998). The paradigm of modeling by iterative conceptualization in mathematics education research. In A. Sierpiska & J. Kilpatrick (Eds.), Mathematics education as a research domain: A search for identity (pp. 263–276). Dordrecht, The Netherlands: Kluwer Academic Publishers.
- Brown, S. I. & Water, M. I. (1990). The art of problem posing. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Bryant, D. & Driscoll, M. (1998). Exploring classroom assessment in mathematics—A guide for professional development. Reston, VA: National Council of Teachers of Mathematics.
- Burton, L. (1996). Assessment of mathematics—What is the agenda? In M. Birenbaum & F. Docky (Eds.), Alternatives in assessment of achievements, learning processes and prior knowledge (pp. 31–61). Boston: Kluwer Academic Publishers.
- Bush, W. S. (Ed.). (2000). Mathematics assessment—Cases and discussion questions for grades 6–12. Reston, VA: National Council of Teachers of Mathematics.
- Bush, W. S. & Leinwand, S. (Eds.). (2000). Mathematics assessment—A practical handbook for grades 6–8. Reston, VA: National Council of Teachers of Mathematics.
- Cai, J., Lane, S. & Jakabsin, M. S. (1996). The role of open-ended tasks and holistic scoring rubrics: Assessing students' mathematical reasoning and communication. In P. C. Elliott & M. J. Kenney (Eds.), Communication in mathematics and beyond: 1996 yearbook (pp. 137–145). Reston, VA.: National Council of Teachers of Mathematics.
- Carpenter, T. P., Fennema, E., Farnke, M. L., Levi, L. & Empson, S. B. (1999). Children's mathematics: Cognitively guided instruction. Portsmouth, NH: Heinemann.
- Carpenter, T. P., Fennema, E., Peterson, P. L., Chiang, C. P. & Loef, M. (1989). Using knowledge of children's mathematics thinking in classroom teaching: An experimental study. American Educational Research Journal, 26, 499–531.
- Carpenter, T. P. & Lehrer, R. (1999). Teaching and learning mathematics with understanding. In E. Fennema & T. A. Romberg (Eds.), Mathematics classrooms that promote understanding (pp. 19–32). Mahwah, NJ: Lawrence Erlbaum Associates.

- Clarke, D. (1992a). Assessment alternatives in mathematics—Professional development package. Canberra, South New Wales, Australia: Curriculum Development Centre.
- Clarke, D. (1992b). The role of assessment in determining mathematics performance. In Leder, G. C. (Ed.), Assessment and learning of mathematics (pp. 145–168). Hawthorn, Victoria, Australia: Australian Council for Educational Research.
- Clarke, D. (1995). Quality mathematics: How can we tell? Mathematics Teacher, 88 (4), 326–328.
- Clarke, D. (1996). Assessment. In A.J. Bishop, K. Clements & C. Keitel (Eds.), International handbook of mathematics education, Vol. 1 (pp. 327–370). Dordrecht, The Netherlands: Kluwer Academic Publishers.
- Clarke, D. (1997). Constructive assessment in mathematics: Practical steps for classroom teachers. Berkeley, CA: Key Curriculum Press.
- Clarke, D. & Stephens, M. (1996). The ripple effect: The instructional impact of the systemic introduction of performance assessment in mathematics. In M. Birenbaum & F. Dochy (Eds.), Alternatives in assessment of achievements, learning processes and prior knowledge (pp. 63–92). Boston: Kluwer Academic Publishers.
- Clarke, D. J., Clarke, D. M. & Lovitt, C. J. (1990). Changes in mathematics teaching: Call for assessment alternatives. In T. J. Cooney & C. R. Hirsch (Eds.), Teaching and learning mathematics in the 1990s, 1990 yearbook (pp. 118–129). Reston, VA: National Council of Teachers of Mathematics.
- Clarke, D. M. (1997). The changing role of the mathematics teacher. Journal for Research in Mathematics Education, 28 (3), 278–308.
- Clarke, D. M. (1999). Classroom reform five years down the track: The experiences of two teachers. Mathematics Education Research Journal, 11 (1), 4–24.
- Cobb, P., Boufi, A., McClain, K. & Whitenack, J. (1997). Reflective discourse and collective reflection. Journal for Research in Mathematics Education, 28 (3), 258–277.
- Cobb, P., Yackel, E. & Wood, T. (1991). Curriculum and teacher development: Psychological and anthropological perspectives. In E. Fennema, T. P. Carpenter & S. Lamon (Eds.), Integrating research on teaching and learning mathematics (pp. 83–119). Albany, NY: State University of New York Press.

- Conference Board of the Mathematical Sciences. (2001). The mathematical education of teachers. Washington, DC: American Mathematical Society, Mathematical Association of America.
- Confrey, J. (2000). Improving research and systemic reform toward equity and quality. In A. E. Kelly & R. A. Lesh (Eds.), Handbook of research design in mathematics and science education (pp. 87–106). Mahwah, NJ: Lawrence Erlbaum Associates.
- Darling-Hammond, L., Ancess, J. & Falk, B. (1995). Authentic assessment in action—Studies of schools and students at work. New York: Teachers College Press.
- Davis, R. B. (1992). Understanding “understanding”. Journal of Mathematical Behavior 11 (3), 225–241.
- Davis, R. B. & Maher, C. A. (1993). Introduction: Schools, mathematics, and the world of reality. In R. B. Davis & C. A. Maher (Eds.), Schools, mathematics, and the world of reality (pp. ix–xii). Boston: Allyn and Bacon.
- Davis, R. B., Maher, C. A. & Noddings, N. (Eds.) (1990). Constructivist views on the teaching and learning of mathematics. Reston, VA: National Council of Teachers of Mathematics.
- Delaney, K. (2001). Teaching mathematics resourcefully. In P. Gates (Ed.), Issues in mathematics teaching (pp. 123–146). London and New York: Routledge.
- Department for Education. (1995). Mathematics in the national curriculum. London, England: HMSO.
- Department for Education and Employment. (1999). The national numeracy strategy—Framework for teaching mathematics from reception to year 6. Sudbury, England: DfEE Publications.
- Dole, S., Nisbet, S., Warren, E. & Cooper, T. J. (1999). Teacher collaboration in developing rich assessment tasks in mathematics as a professional development activity. Mathematics Teacher Education and Development, 1, 38–49.
- Dossey, J. A. & Swafford, J. O. (1993). Issues in mathematics assessment in the United States. In M. Niss (Ed.), Cases of assessment in mathematics education—An ICMI study (pp. 43–57). Dordrecht, The Netherlands: Kluwer Academic Publishers.

- Driscoll, M. (1999). Crafting a sharper lens: Classroom assessment in mathematics. In M. Z. Solomon (Ed.), The diagnostic teaching: Constructing new approaches to professional development (pp. 78–103). New York and London: Teachers College Press.
- Driscoll, M. & Bryant, D. (1998). Learning About assessment, learning through assessment. Washington, D.C.: National Academy Press.
- Duckworth, E. (1987). Some depths and perplexities of elementary arithmetic. Journal of Mathematical Behavior, 6 (1), 43–94.
- Duit, R., Treagust, D. F. & Mansfield, H. (1996). Investigating student understanding as a prerequisite to improving teaching and learning in science and mathematics. In D. F. Treagust, R. Duit & B. J. Fraser (Eds.), Improving teaching and learning in science and mathematics (pp. 17–31). New York and London: Teachers College Press.
- Even, R. & Tirosh, D. (2002). Teacher knowledge and understanding of students' mathematical learning. In L. English (Ed.), Handbook of international research in mathematics education. (pp.219-240). Mahwah, NJ: Laurence Erlbaum.
- Extended assessment for interactive mathematics. (1995). New York: Glencoe.
- Falk, B. (2000). The heart of the matter—Using standards and assessment to learn. Portsmouth, NH: Heinemann.
- Ferrini-Mundy, J. (1998). What have we learning for the future? In J. Ferrini-Mundy, K. Graham, L. Johnson & G. Mills (Eds.), Making change in mathematics education—Learning from the field (pp. 129–141). Reston, VA: National Council of Teachers of Mathematics.
- Firestone, W. A., Mayrowetz, D. & Fairman, J. (1998). Performance-based assessment and instructional change. Educational Evaluation and Policy Analysis, 20, 95–113.
- Flewelling, G. & Higginson, W. (2000). A handbook of rich learning tasks—Realizing a vision of tomorrow's mathematics classroom. Kingston, Ontario, Canada: Centre for Mathematics, Science and Technology Education, Queen's University.

- Frechtling, J. (2001). What evaluation tells us about professional development programs in mathematics and science. In C. R. Nesbit, J. D. Wallace, D. K. Pugalee, A. Miller & W. J. DiBiase (Eds.), Developing teacher leaders—Professional development in science and mathematics (pp. 17–42). Columbus, OH: ERIC Clearinghouse for Science, Mathematics and Environment Education.
- Freedman, R. L. H. (1994). Open-ended questioning—A handbook for educators. Menlo Park, CA: Addison-Wesley.
- Friedlander, A., Bruckheimer, M. & Albert, J. (1987). Counselling: A model for in-service teacher education. Zentrablatt für Didaktik der mathematik, 4, 143–149.
- Friel, S. N. & Bright, G. W. (1997). Reflecting on our work: NSF reacher enhancement in K-6 mathematics. Lanham, MD: University Press of America, Inc.
- Friel, S. N. & Bright, G. W. (2001). Effective professional development for teacher leaders: Lessons learned from K-6 mathematics enhancement program. In C. R. Nesbit, J. D. Wallace, D. K. Pugalee, A. Miller & W. J. DiBiase (Eds.), Developing teacher leaders—Professional development in science and mathematics (pp. 71–88). Columbus, OH: ERIC Clearinghouse for Science, Mathematics and Environment Education.
- Fuchs, L. S., Fuchs, D., Karns, K., Hamlett, C. L. & Katzaroff, M. (1999). Mathematics performance assessment in the classroom: Effects on teacher planning and student problem solving. American Educational Research Journal, 36 (3), 609–646.
- Fullan, M. (1999). Change forces: The sequel. Philadelphia, PA: Falmer Press, Taylor & Francis Inc.
- Fullan, M. G., & Stiegelbauer, S. (1991). The new meaning of educational change. London: Cassell Educational Ltd.
- Gattegno, C. (1988). Reflections on forty years of work on mathematics teaching. For the Learning of Mathematics, 8 (3), 41–42.
- Ginsburg, H. P., Jacobs, S. F. & Lopez, L. S. (1993a). Assessing mathematical thinking and learning potential. In R. B. Davis & C. A. Maher (Eds.), Schools, mathematics, and the world of reality (pp. 237–262). Boston: Allyn and Bacon.

- Ginsburg, H. P., Jacobs, S. F. & Lopez, L. S. (1993b). Assessing mathematical thinking and learning potential in primary grade children. In M. Niss (Ed.), Investigations into assessment in mathematics education—An ICMI study (pp. 157–167). Dordrecht, The Netherlands: Kluwer Academic Publishers.
- Glaser, B. G. & Strauss, A. L. (1967). The discovery of grounded theory: Strategies for qualitative research. New York: Aldine de Gruyter.
- Goetz, J. P. & LeCompte, M. D. (1984). Ethnography and qualitative design in educational research. Orlando, FL: Academic Press, Inc.
- Goldin, G. A. (1993). Mathematics education—The context of the crisis. In R. B. Davis & C. A. Maher (Eds.), Introduction: Schools, mathematics, and the world of reality (pp. 1–7). Boston: Allyn and Bacon.
- Graeber, A. O. & Tirosh, D. (1988). Multiplication and division involving decimals: Preservice elementary teachers' performance and beliefs. Journal of Mathematical Behavior, 7 (3), 263–280.
- Graham, K. & Ferrucci, B. J. (1998). What mathematics is important? In J. Ferrini-Mundy, K. Graham, L. Johnson & G. Mills (Eds.), Making change in mathematics education—Learning from the field (pp. 19–41). Reston, VA: National Council of Teachers of Mathematics.
- Grant, S. G., Peterson, P. I. & Shojgreen-Downer, A. (1996). Learning to teach mathematics in the context of systemic reform. American Educational Research Journal, 33 (2), 509–541.
- Graue, M. E. & Smith, S. Z. (1996). Shaping assessment through instructional innovation. Journal of Mathematical Behavior 15, (2), 113–136.
- Greenwood, J. J. (1993). On the nature of teaching and assessing “mathematical power” and “mathematical thinking”. Arithmetic Teacher, 41 (3), 144–152.
- Grouws, D. A. & Meier, S. L. (1992). Teaching and assessment relationships in mathematics education. In G. C. Leder (Ed.), Assessment and learning of mathematics (pp. 83–105). Hawthorn, Victoria, Australia: Australian Council for Educational Research Ltd.
- Guba, E.G. & Lincoln, Y. S. (1989). Fourth generation evaluation. London: Sage Publications.

- Guba, E. G. & Lincoln, Y. S. (1994). Competing paradigms in qualitative research. In N. K. Denzin & Y. S. Lincoln (Eds.), Handbook of qualitative research (pp. 105–117). London: Sage Publications.
- Hajar, P. (1997). Awakening to the mathematician within: One teacher's story. In A. L. Goodwin (Ed.), Assessment for equity and inclusion: Embracing all our children (pp. 121–140). New York: Routledge.
- Hand, B. (1996). Diagnosis of teachers' knowledge bases and teaching roles when implementing constructivist teaching/learning approaches. In D. F. Treagust, R. Duit & B. J. Fraser (Eds.), Improving teaching and learning in science and mathematics. New York: Teachers College Press.
- Hargreaves, A., Earl, L. & Schmidt, M. (2002). Perspectives on alternative assessment reform. American Educational Research Journal, 39, 69–95.
- Hargreaves, A. & Fullan, M. (1998). What's worth fighting for out there. New York: Teachers College Press.
- Harmon, M., Smith, T. A., Martin, M. O., Kelly, D. L., Beaton, A. E., Mullis, I. V. S., Gonzalez, E. J. & Orpwood, G. (1997). Performance assessment in IEA's third international mathematics and science study. Chestnut Hill, MA: TIMSS International Study Center, Boston College.
- Heaton, R. M. (2000). Teaching mathematics to the new standards. New York, NY: Teachers College Press.
- Heinz, K., Kinzel, M., Simon, M. A. & Tzur, R. (2000). Moving students through steps of mathematical knowing—An account of the practice of an elementary mathematics teacher in transition. Journal of Mathematical Behavior, 19 (1), 83–107.
- Henningsen, M. & Stein, M. K. (1997). Mathematical tasks and student cognition: Classroom-based factors that support and inhibit high-level mathematical thinking and reasoning. Journal for Research in Mathematics Education, 28 (5), 524–549.
- Herman, J. L., Aschbacher, P. R. & Winters, L. (1992). A practical guide to alternative assessment. Alexandria, VA: Association for Supervision and Curriculum Development.

- Herrington, A., Herrington, J., Sparrow, L. & Oliver, R. (1998). Learning to teach and assess mathematics using multimedia: A teacher development project. Journal of Mathematics Teacher Education, 1, 89–112.
- Herrington, T., Sparrow, L., Herrington, J. & Oliver, R. (1997). Investigating assessment strategies in mathematics classrooms. Perth, Western Australia: Edith Cowan University.
- Hershkowitz, R. (1989). Geometrical concept images of students and teachers. Unpublished doctoral dissertation, The Hebrew University of Jerusalem, Jerusalem. (in Hebrew).
- Heuvel-Panhuizen, M. van den (1994). Improvement of (didactical) assessment by improvement of problems: An attempt with respect to percentage. Educational Studies in Mathematics, 27, 341–372.
- Heuvel-Panhuizen, M. van den (1996). Assessment and realistic mathematics education. Utrecht: Freudenthal Institute.
- Hiebert, J. (1998). Aiming research toward understanding: Lessons we can learn from children. In A. Sierpiska & J. Kilpatrick (Eds.), Mathematics education as a research domain: A search for identity, Vol. 1 (pp. 141–152). Dordrecht, The Netherlands: Kluwer Academic Publishers.
- Hiebert, J., Carpenter, T. P., Fennema, E., Fuson, K., Human, P., Murray, H., Olivier, A. & Wearne, D. (1996). Problem solving as a basis for reform in curriculum and instruction: The case of mathematics. Educational Researcher, 25 (4), 12–21.
- Hiebert, J., Carpenter, T. P., Fennema, E., Fuson, K. C., Wearne, D., Murray, H., Olivier, A. & Human, P. (1997). Making sense: Teaching and learning Mathematics with understanding. Portsmouth, NH: Heinemann.
- Hofstein, A. & Even, R. (2001). Developing chemistry and mathematics teacher leaders in Israel. In C. R. Nesbit, J. D. Wallace, D. K. Pugalee, A. Miller & W. J. DiBiase (Eds.), Developing teacher leaders—Professional development in science and mathematics (pp. 189–208). Columbus, OH: ERIC Clearinghouse for Science, Mathematics and Environmental Education.
- Howson, G. (1993). The relationship between assessment, curriculum and society. In M. Niss (Ed.), Investigations into assessment in mathematics education—An ICMI study (pp. 47–56). Dordrecht, The Netherlands: Kluwer Academic Publishers.

- Joffe, L. S. (1992). The English experience of a national curriculum and assessments. In G. C. Leder (Ed.), Assessment and learning of mathematics (pp. 190–222). Hawthorn, Victoria, Australia: Australian Council for Educational Research Ltd.
- Johnson, B. (1996). Performance assessment handbook. Princeton, NJ: Eye on Education.
- Jones, L. V. (2001). Assessing achievement versus high-stakes testing: A crucial contrast. Educational Assessment, 7, 21–28.
- Kazemi, E. (1998). Discourse that promotes conceptual understanding. Teaching Children Mathematics, 4 (7), 410–414.
- Kieran, C. (1994). Doing and seeing things differently: A 25-year retrospective of mathematics education research on learning. Journal for Research in Mathematics Education, 25 (6), 583–607.
- Kilpatrick, J. (1992). A history of research in mathematics education. In D. A. Grouws (Ed.), Handbook of research on mathematics teaching and learning (pp. 3–38). New York: Macmillan Publishing Co.
- Kilpatrick, J., Swafford, J. & Findell, B. (Eds.). (2001). Adding it up: Helping children learn mathematics. Washington, DC: National Academy Press.
- Kline, M. (1973). Why Johnny can't add—The failure of the new math. New York: St. Martin's Press.
- Kohn, A. (2000). The case against standardized testing—Raising the scores, ruining the schools. Portsmouth, NH: Heinemann.
- Krulik, S. & Rudnick, J. A. (1998). Assessing reasoning and problem solving—A sourcebook for elementary school teachers. Boston: Allyn and Bacon.
- Kuhs, T. M. (1992). Mathematics assessment: Alternative approaches. Reston, VA: National Council of Teachers of Mathematics.
- Kulm, G. (Ed.). (1990). Assessing higher order thinking in mathematics. Washington, DC: American Association for the Advancement of Science.
- Kulm, G. (1994). Mathematics assessment—What works in the classroom. San Francisco, CA: Jossey-Bass Publishers.

- Lambdin, D. V. (1993). The NCTM's 1989 evaluation standards: Recycled ideas whose time has come? In N. L. Webb & A. F. Coxford (Eds.), Assessment in the mathematics classroom, 1993 yearbook (pp. 7–16). Reston, VA: National Council of Teachers of Mathematics.
- Lambdin, D. V. (1998). What roles does assessment play in mathematics education reform? In J. Ferrini-Mundy, K. Graham, L. Johnson & G. Mills (Eds.), Making change in mathematics education – Learning from the field (pp. 87–101). Reston, VA: National Council of Teachers of Mathematics.
- Lambdin, D. V., Kehle, P. E. & Preston, R. V. (Eds.). (1996). Emphasis on assessment—Readings from NCTM's school-based journals. Reston, VA: National Council of Teachers of Mathematics.
- Lambdin Kroll, D., Masingila, J. O. & Tinsley Mau, S. (1992). Grading cooperative problem solving. Mathematics Teacher, 85 (8), 628–632.
- Lampert, M. (1991). Connecting mathematical teaching and learning. In E. Fennema, T. P. Carpenter & S. Lamon (Eds.), Integrating research on teaching and learning mathematics (pp. 121–152). Albany, NY: State University of New York Press.
- Lange, J. de (1992). Critical factors for real changes in mathematics learning. In Leder, G. C. (Ed.), Assessment and learning of mathematics (pp. 305–329). Hawthorn, Victoria, Australia: Australian Council for Educational Research Ltd.
- Lange, J. de, Burrill, G., Romberg, T. & Reeuwijk, M. van (1993). Learning and testing mathematics in context—The case: data visualization. Madison: National Center for Research in Mathematical Sciences Education.
- Leadership Development and Enabling Change Project. (1994). Lessons in mathematics for the classroom—and for inservice sessions with alternative assessment procedures. Cedar Falls, IA: University of Northern Iowa.
- LeCompte, M. D. (2000). Analyzing qualitative data. Theory into Practice, 39 (3), 146–154.
- Leder, G. C. & Forgasz, H. J. (1992). Perspectives on learning, teaching and assessment. In G. C. Leder (Ed.), Assessment and learning of mathematics (pp. 1–23). Hawthorn, Victoria, Australia: Australian Council for Educational Research Ltd.

- Leinwand, S. (2000). Sensible mathematics: A guide for school leaders. Portsmouth, NH: Heinemann.
- Linchevski, L. (1985). The meanings that elementary school teachers give to concepts they use in teaching arithmetic and geometry. Unpublished doctoral dissertation, The Hebrew University of Jerusalem, Jerusalem. (in Hebrew).
- Lincoln, Y. S. & Guba, E. G. (1985). Naturalistic inquiry. London: Sage Publications.
- Linn, R. L. (1991). Dimensions of thinking: Implications for testing. In L. Idol & B. F. Jones (Eds.), Educational values and cognitive instruction: Implications for reform (pp. 179–208). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Linn, R. L. (2000). Assessments and accountability. Educational Researcher, 29 (2), 4–16.
- Linn, R. L. & Gronlund, N. E. (1995). Measurement and assessment in teaching. Upper Saddle River, NJ: Simon & Schuster.
- Linn, R. L. & Herman, J. L. (1997). A policymaker's guide to standards-led assessment. Denver, CO: National Center for Research on Evaluation, Standards and Student Testing (CRESST).
- Loucks-Horskey, S., Hewson, P. W., Love, N. & Stiles, K. E. (1998). Designing professional development for teachers of science and mathematics. Thousand Oaks, CA: Corwin Press, Inc.
- Ma, L. (1999). Knowing and teaching elementary mathematics—Teachers' understanding of fundamental mathematics in China and the United States. Mahwah, NJ: Lawrence Erlbaum Associates.
- Madfes, T. J. & Muench, A. (1999). Learning from assessment—A middle school mathematics professional development resource. San Francisco, CA: WestEd.
- Malone, J. A. & Taylor, P. C. S. (Eds.). (1993). Constructivist interpretations of teaching and learning mathematics. Perth, Western Australia: National Key Centre for School Science and Mathematics.
- Manfre, E., Moser, J. M., Lobato, J. E. & Morrow, L. (1996). Performance assessments for Heath mathematics connections. Lexington, MA: D.C. Heath and Company.

- Marshall, S. P. (1988). Assessing problem solving: A short-term remedy and a long-term solution. In R. I. Charles & E. A. Silver (Eds.), The teaching and assessing of mathematical problem solving, (pp. 159–177). Reston, VA: National Council of Teachers of Mathematics.
- Mason, J. (1994). Researching from the inside in mathematics education—Locating an I-you relationship. Proceedings of the Eighteenth International Conference for the Psychology of Mathematics Education, University of Lisbon, 1, 176–194.
- Mathematical Sciences Education Board. (1989). Everybody counts—A report to the nation on the future of mathematics education. Washington, D.C.: National Academy Press.
- Mathematical Sciences Education Board. (1993a). Measuring up—Prototypes for mathematics assessment. Washington, D.C.: National Academy Press.
- Mathematical Sciences Education Board. (1993b). Measuring what counts—A conceptual guide for mathematics assessment. Washington, D.C.: National Academy Press.
- Mathematical Sciences Education Board. (2001). Knowing and learning mathematics for teaching. Washington, D.C.: National Academy Press.
- Maykut, P. & Morehouse, R. (1994). Beginning qualitative research—A philosophic and practical guide. London: The Falmer Press.
- Medina, N. & Neill, D. M. (1990). Fallout from the testing explosion: How 100 million standardized exams undermine equity and excellence in America's public schools. Cambridge, MA: FairTest.
- Meloy, J. M. (1994). Writing the qualitative dissertation—Understanding by doing. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Mokros, J., Russell, S. J. & Economopoulos, K. (1995). Beyond arithmetic. Changing mathematics in the elementary classroom. Palo Alto, CA: Dale Seymour Publications.
- Moon, J. (1997). Developing judgment—Assessing children's work in mathematics. Portsmouth, NH: Heinemann.

- Moon, J. & Schulman, L. (1995). Finding the connections—Linking assessment, instruction, and curriculum in elementary mathematics. Portsmouth, NH: Heinemann.
- Morgan, C. (2000). Better assessment in mathematics education? A social perspective. In J. Boaler (Ed.), Multiple perspectives on mathematics teaching and learning (pp. 225–242). Westport, CT and London, England: Ablex Publishing.
- Moschkovich, J. N. & Brenner, M. E. (2000). Integrating a naturalistic paradigm into research on mathematics and science cognition and learning. In A. E. Kelly & R. A. Lesh (Eds.), Handbook of research design in mathematics and science education (pp. 457–486). Mahwah, NJ: Lawrence Erlbaum Associates.
- Nathan, L. (2002). The human face of the high-stakes testing story. Phi Delta Kappan, 84, 595–600.
- National Council of Supervisors of Mathematics. (1996). Great tasks and MORE!!—A Source Book of Camera-Ready Resources on Mathematics Assessment. Golden, CO: National Council of Supervisors of Mathematics.
- National Council of Teachers of Mathematics. (1989). Curriculum and evaluation standards for school mathematics. Reston, VA: National Council of Teachers of Mathematics.
- National Council of Teachers of Mathematics. (1991). Professional standards for teaching mathematics. Reston, VA: National Council of Teachers of Mathematics.
- National Council of Teachers of Mathematics. (1995). Assessment standards for school mathematics. Reston, VA: National Council of Teachers of Mathematics.
- National Council of Teachers of Mathematics. (2000). Principles and standards for school mathematics. Reston, VA: National Council of Teachers of Mathematics.
- Nelson, B. S. (1998). Lenses on learning: Administrators' views on reform and the professional development of teachers. Journal of Mathematics Teacher Education, 1 (2), 191–215.

- Nesbit, C. R., DiBiase, W. J., Miller, A. S. & Wallace, J. D. (2001). In their own words: What science and mathematics teacher leaders say are important aspects of professional development. In C. R. Nesbit, J. D. Wallace, D. K. Pugalee, A. Miller & W. J. DiBiase (Eds.), Developing teacher leaders—Professional development in science and mathematics (pp. 227–244). Columbus, OH: ERIC Clearinghouse for Science, Mathematics and Environment Education.
- Nesher, P. (1996). National testing and assessment in Israel. Jerusalem: The Ministry of Education.
- Nevo, D. (1991). Assessment as a tool for improving teaching and learning. In M. Zilberstein (Ed.), Assessment in planning learning and teaching—A teacher's perspective (pp. 5–12). Jerusalem: Ministry of Education and Culture. (in Hebrew).
- Niss, M. (1993a). Assessment in mathematics education and its effects: An introduction. In M. Niss (Ed.), Investigations into assessment in mathematics education—An ICMI study (pp. 1–30). Dordrecht, The Netherlands: Kluwer Academic Publishers.
- Niss, M. (Ed.). (1993b). Cases of assessment in mathematics education—An ICMI study. Dordrecht, The Netherlands: Kluwer Academic Publishers.
- Noddings, N. (1993). Constructivism and caring. In R. B. Davis & C. A. Maher (Eds.), Introduction: Schools, mathematics, and the world of reality (pp. 35–50). Boston: Allyn and Bacon.
- Pandey, T. (1991). A sampler of mathematics assessment. Sacramento, CA: Department of Education.
- Pandey, T. (1992). Test development profile of a state-mandated large-scale assessment instrument in mathematics. In T. A. Romberg (Ed.), Mathematics assessment and evaluation—Imperatives for mathematics educators (pp. 100–127). Albany, NY: State University of New York Press.
- Paris, S. G. (1998). Why learner-centered assessment is better than high-stakes testing. In N. M. Lambert & B. L. McCombs (Eds.), How students learn: Reforming schools through learner-centered education (pp. 189–209). Washington, DC: American Psychological Association.
- Partners in Change Project. (1997). The Partners in Change handbook—A professional development curriculum in mathematics. Boston: Boston University.

- Paulos, J. A. (1988). Innumeracy—Mathematical illiteracy and its consequences. New York: Hill and Wang.
- Philipp, R. A., Flores, A., Sowder, J. T. & Schappelle, B. P. (1994). Conceptions and practices of extraordinary mathematics teachers. Journal of Mathematical Behavior, 13 (2), 155–180.
- Pressley, M. & McCormick, C. (1995). Cognition, teaching, and assessment. New York: HarperCollins College Publishers.
- Resnick, L. B. (1987). Education and learning to think. Washington, D.C.: National Academy Press.
- Resnick, L. B & Resnick, D. P. (1992). Assessing the thinking curriculum: New tools for educational reform. In B. R. Gifford & M. C. O’Conner (Eds.), Changing assessments: Alternative views of aptitude, achievement and instruction (pp. 37–76). Boston, Dordrecht and London: Kluwer Academic Publishers.
- Reys, B. J. & Long, V. M. (1995). Teacher as architect of mathematical tasks. Teaching Children Mathematics, 5, 296–299.
- Rhine, S. (1998). The role of research and teachers’ knowledge base in professional development. Educational Researcher, 27 (5), 27–31.
- Romberg, T. A. (1997). Mathematics in context: Impact on teachers. In E. Fennema & B. S. Nelson (Eds.), Mathematics teachers in transition (pp. 357–380). Mahwah, NJ: Lawrence Erlbaum Associates.
- Romberg, T. A. & Kaput, J. J. (1999). Mathematics worth teaching, mathematics worth understanding. In E. Fennema & T. A. Romberg (Eds.), Mathematics classroom that promote understanding (pp. 3–17). Mahwah, NJ: Lawrence Erlbaum Associates.
- Romberg, T. A., Zarinnia, E. A. & Collis, K. F. (1990). A new world view of assessment in mathematics. In G. Kulm (Ed.), Assessing higher order thinking in mathematics (pp. 21–38). Washington, D.C.: American Association for the Advancement of Science.
- Rothman, R. (1995). Measuring up—Standards, assessment, and school reform. San Francisco, CA: Jossey-Bass Publishers.

- Schifter, D. (1993). Mathematics process as mathematics content: A course for teachers. Journal of Mathematical Behavior, 12 (3), 271–283.
- Schifter, D. & Fosnot, C. T. (1993). Reconstructing mathematics education—Stories of teachers meeting the challenge of reform. New York: Teachers College Press.
- Schoenfeld, A. H. (1994). Reflections on doing and teaching mathematics. In A. H. Schoenfeld (Ed.), Mathematical thinking and problem solving (pp. 53–70). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Schorr, R. Y. (2000). Impact at the student level: a study of the effects of a teacher development intervention on students' mathematical thinking. Journal of Mathematical Behavior, 19 (2), 209–231.
- Schwartz, J., Kenney, J. M., Kelly, K. A., Sienkiewicz, T., Sivan, Y., Steinbok, V. & Yerushalmy, M. (1995). Assessing mathematical understanding and skills effectively—An interim report of the Harvard group, balanced assessment in mathematics project. Cambridge, MA: Harvard Graduate School of Education.
- Shafer, M. C. & Romberg, T. A. (1999). Assessment in classrooms that promote understanding. In E. Fennema & T. A. Romberg (Eds.), Mathematics classrooms that promote understanding (pp. 159–184). Mahwah, NJ: Lawrence Erlbaum Associates.
- Shannon, A. (1999). Keeping score. Washington, DC: National Academy Press.
- Sheingold, K. J., Heller, J. I. & Paulukonis, S. T. (1995). Actively seeking evidence: Teacher change through assessment development. Princeton, NJ: Educational Testing Service.
- Shepard, L. A. (2000a). The role of assessment in a learning culture. Educational Researcher, 29 (7), 4–14.
- Shepard, L. A. (2000b). The role of classroom assessment in teaching and learning. Los Angeles, CA: National Center for Research on Evaluation, Standards, and Student Testing (CRESST).
- Silver, E. A. & Kenney, P. A. (1995). Sources of assessment information for instructional guidance in mathematics. In T. A. Romberg (Ed.), Reform in school mathematics and authentic assessment (pp. 38–86). Albany, NY: State University of New York.

- Silver, E. A. & Kilpatrick, J. (1988). Testing mathematical problem solving. In R. I. Charles & E. A. Silver (Eds.), The teaching and assessing of mathematical problem solving (pp. 178–186). Reston, VA: National Council of Teachers of Mathematics.
- Silver, E. A., Kilpatrick, J. & Schlesinger, B. (1995). Thinking through mathematics—Fostering inquiry and communication in mathematics classrooms. New York: College Entrance Examination Board.
- Silver, E. A. & Lane, S. (1993). Assessment in the context of mathematics instruction reform: The design of assessment in the Quasar project. In M. Niss (Ed.), Cases of assessment in mathematics education—An ICMI study (pp. 59–69). Dordrecht, The Netherlands: Kluwer Academic Publishers.
- Simon, M. A. (1993). Prospective elementary teachers' knowledge of division. Journal for Research in Mathematics Education, 24 (3), 233–254.
- Simon, M. A. (1995). Reconstructing mathematics pedagogy from a constructivist perspective. Journal for Research in Mathematics Education, 26 (2), 114–145.
- Simon, M. A. & Blume, G. W. (1994). Mathematical modeling as a component of understanding ratio-as-measure: A study of prospective elementary teachers. Journal of Mathematical Behavior, 13 (2), 183–197.
- Smith, T. (2000). Bridging the research-practice gap: Developing a pedagogical framework that promotes mathematical thinking and understanding. Mathematics Teacher Education and Development, 2, 4–16.
- Solomon, M. Z. (Ed.). (1999). The diagnostic teacher: Constructing new approaches to professional development. New York and London: Teachers College Press.
- Solomon, M. Z. & Morocco, C. C. (1999). The diagnostic teacher. In M. Z. Solomon (Ed.), The diagnostic teacher: Constructing new approaches to professional development (pp. 231–246). New York and London: Teachers College Press.
- Sowder, J. T., Philipp, R. A., Armstrong, B. E. & Schappelle, B. P. (1998). Middle-grade teachers' mathematical knowledge and its relationship to instruction. Albany, NY: State University of New York Press.
- Stecher, B. M. & Mitchel, K. J. (1995). Vermont teachers' understanding of mathematical problem solving and "good" math problems. Los Angeles, CA: National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

- Steen, L. A. (Ed.). (1990). On the shoulders of giants—New approaches to numeracy. Washington, D.C.: National Academy Press.
- Steen, L. A. (Ed.). (1997). Why numbers count—Quantitative literacy for tomorrow's America. New York: College Entrance Examination Board.
- Stein, M. K., Grover, B. W. & Henningsen, M. (1996). Building student capacity for mathematical thinking and reasoning: An analysis of mathematical tasks used in reform classrooms. American Educational Research Journal, 33 (2), 455–488.
- Stein, M. K., Smith, M. S., Henningsen, M. A. & Silver, E. A. (2000). Implementing standards-based mathematics instruction: A casebook for professional development. Reston, VA: National Council of Teachers of Mathematics. New York and London: Teachers College Press.
- Stenmark, J. K. (1989). Assessment alternatives in mathematics—An overview of assessment techniques that promote learning. Berkeley: EQUALS—Lawrence Hall of Science.
- Stenmark, J. K. (Ed.). (1991). Mathematics assessment—Myths, models, good questions, and practical suggestions. Reston, VA: National Council of Teachers of Mathematics.
- Stephens, M. & Money, R. (1993). New developments in senior secondary assessment in Australia. In M. Niss (Ed.), Cases of assessment in mathematics education—An ICMI study (pp. 155–171). Dordrecht, The Netherlands: Kluwer Academic Publishers.
- Sullivan, P. & Clarke, D. (1991). Catering to all abilities through “good” questions. Arithmetic Teacher, 39 (2), 14–18.
- Sullivan, P. & Lilburn, P. (1997). Open-ended maths activities—Using “good” questions to enhance learning. South Melbourne: Oxford University Press.
- Swan, M. (1993). Improving the design and balance of mathematical assessment. In M. Niss (Ed.), Investigations into assessment in mathematics education—An ICMI study (pp. 195–216). Dordrecht, The Netherlands: Kluwer Academic Publishers.
- Thompson, A. G. & Thompson, P. W. (1996). Talking About rates conceptually: Part II. Mathematical knowledge for teaching. Journal for Research in Mathematics Education, 27 (1), 2–24.

- Tobin, K. (Ed.). (1993). The practice of constructivism in science education. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Walter, M. I. (1993). Generating problems from almost anything. In S. I. Brown & M. I. Walter (Eds.), Problem posing—Reflections and applications (pp. 302–316). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Watson, A. & Mason, J. (1998). Questions and prompts for mathematical thinking. Derby, England: Association of Teachers of Mathematics.
- Webb, N. L. (1992). Assessment of students' knowledge of mathematics: Steps toward A theory. In D. A. Grouws (Ed.), Handbook of research on mathematics teaching and learning—A project of the national council of teachers of mathematics (pp. 661–683). New York: Macmillan Publishing Company.
- Webb, N. L. (1993). Assessment for the mathematics classroom. In N. L. Webb & A. F. Coxford (Eds.), Assessment in the mathematics classroom, 1993 yearbook (pp. 1–6). Reston, VA: National Council of Teachers of Mathematics.
- Webb, N. L. & Briars, D. (1990). Assessment in mathematics classrooms, K-8. In T. J. Cooney & C. R. Hirsch (Eds.), Teaching and learning mathematics in the 1990s, 1990 yearbook (pp. 108–117). Reston, VA: National Council of Teachers of Mathematics.
- Webb, N. L. & Romberg, T. A. (1992). Implications of the NCTM *standards* for mathematics assessment. In T. A. Romberg (Ed.), Mathematics assessment and evaluation—Imperatives for mathematics educators (pp. 37–60). Albany, NY: State University of New York Press.
- Wheeler, D. (1993). Epistemological issues and challenges to assessment: What is mathematical knowledge? In M. Niss (Ed.), Investigations into assessment in mathematics education—An ICMI study (pp. 87–95). Dordrecht, The Netherlands: Kluwer Academic Publishers.
- Wiggins, G. (1998). Educative assessment—Designing assessments to inform and improve student performance. San Francisco, CA: Jossey-Bass Publishers.
- Wilcox, S. K. & Lanier, P. E. (Eds.). (2000). Using assessment to reshape mathematics teaching. Mahwah, NJ: Lawrence Erlbaum Associates.

- Wilson, S. M. & Berne, J. (1999). Teacher learning and the acquisition of professional knowledge: An examination of research on contemporary professional development. Review of Research in Education, 24, 173–209.
- Yackel, E. & Cobb, P. (1996). Sociomathematical norms, argumentation, and autonomy in mathematics. Journal for Research in Mathematics Education, 27 (4), 458–477.
- Yoshida, M. (1999). Lesson study: A case study of a Japanese approach to improving instruction through school-based teacher development. Unpublished doctoral dissertation, University of Chicago, Chicago, IL.

Appendix A: Investigative Tasks

Add-a-Digit

Animals

Bridges

Geoboard Quadrilaterals

Getting Results

Parallelograms

Population

Table Arrangements

Trains

ADD-A-DIGIT

The sign \approx means an approximate answer.

For example: $41 \times 3 \approx 120$

1. Circle the closest possible answer for each exercise:

a. $5 \times 49 \approx \square$

25 200 250

b. $7 \times 23 \approx \square$

25 200 250

2. Choose “appropriate” or “not appropriate” for each exercise:

a. $39 \times 5 \approx 200$ appropriate / not appropriate

b. $3 \times 41 \approx 1200$ appropriate / not appropriate

c. $72 \times 8 \approx 6000$ appropriate / not appropriate

d. $4 \times 198 \approx 800$ appropriate / not appropriate

3. Write a digit in each box, so that the product of each multiplication exercise will be as close as possible to the exact answer:

a. $\square \times 32 \approx 200$

c. $3 \times 4\square \approx 150$

b. $4 \times \square 1 \approx 250$

d. $\square 2 \times 3\square \approx 1000$

4. Write a digit – 1, 2, 3, 4 – in each box, so that the product of the exercise will be as close to 300 as possible.

Use each digit only once.

$$\square\square \times \square\square \approx 300$$

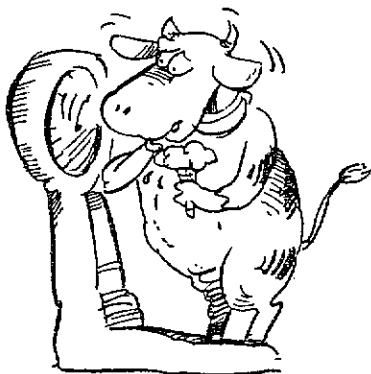
Explain how you decided where to write each digit.

ANIMALS

This table gives information about five types of animals.

The data in the table are approximate and do not include exceptional cases.

The animal	Adult weight	Number of babies at one birth	Length of life	Amount of food for a week
Cat	2 kg	5 kittens	15 years	1 kg
Dog (medium sized)	8 kg	5 puppies	15 years	3 kg
Horse	200 kg	1 foal	30 years	70 kg
Cow	400 kg	1 calf	22 years	80 kg
Sheep	80 kg	1 kid	20 years	30 kg



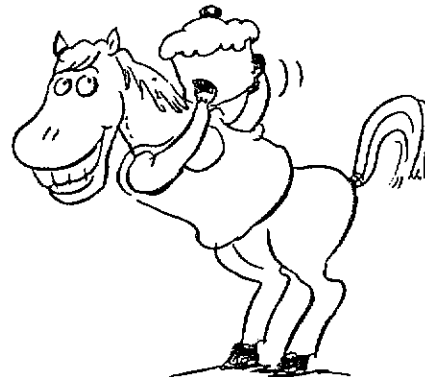
1. Use the data in the table to complete these sentences:

- a) A dog eats _____ kg of food in a week.
- b) An animal who lives 30 years is a _____.
- c) An animal whose weight is less than 5 kg is a _____.
- d) The heaviest animal is a _____ and his weight is _____.
- e) The animals with the most babies at birth are _____.

2. Write two additional sentences using the information from the table about how much food the animals eat, in an interesting way:

- a) _____

- b) _____

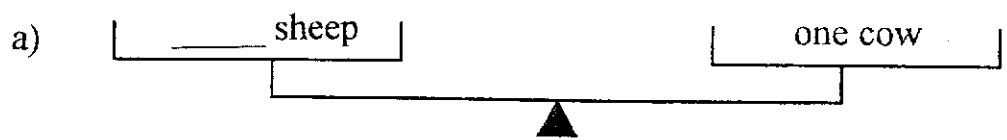


3. Assume that we have very large scales, and we can put animals on them.

Equal weights on both pans of the scales stabilizes them.

Use the animals' weights in the table above to stabilize the scales.

Complete the missing data:



Find additional examples:



4. How many more years does a horse live than a sheep?
5. How many times longer is a horse's life than a cat's?
6. Who eats a larger amount of food in a week:
3 cows or 6 sheep?
20 dogs or 1 horse?
7. What is the approximate amount of food that a horse eats in one month?
8. Uncle Moses has 3 cows and 2 horses.
How much food does he need to feed them for a week?
9. Find three different ways to feed cows, horses and sheep with 200 kg of food each week.

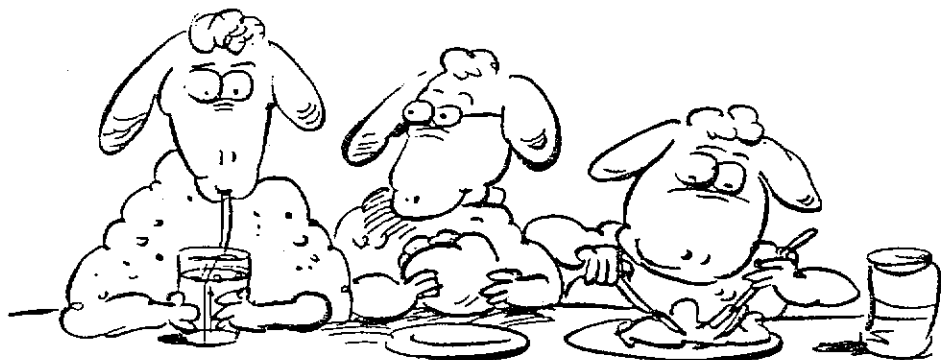
Note: you can use combinations of animals who eat a little less than 200 kg.

Example: 2 cows and 1 sheep

First combination:

Second combination:

Third combination:



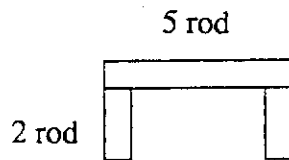


my name: _____

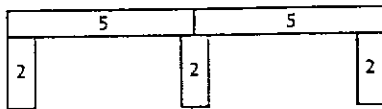
my class: _____

BRIDGES

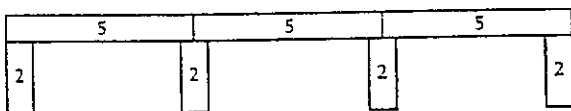
You can see here three rod bridges.



Bridge of one arch (length 5).



Bridge of 2 arches (length 10).



Bridge of 3 arches (length 15).



1. Build a bridge of 4 arches.
 - a. How many rods of 5 did you use? ____
 - b. How many rods of 2 did you use? ____
 - c. What is the length of this bridge? ____

2. Assaf built a bridge of 6 arches.
 - a. How many rods of 5 did he use? ____
 - b. How many rods of 2 did he use? ____
 - c. What is the length of this bridge? ____

3. Shira built a bridge of 10 arches.
 - a. How many rods of 5 did she use? ____
 - b. How many rods of 2 did she use? ____
 - c. What is the length of this bridge? ____

4. Nogah built a bridge of 40 arches.
 - a. How many rods of 5 did she use? ____
 - b. How many rods of 2 did she use? ____
 - c. What is the length of this bridge? ____

5. Ayelet built a bridge that stood on 30 "pillars" (rods of 2).
 - a. How many rods of 5 did she use? ____
 - b. What is the length of this bridge? ____



6. Neomi built a bridge with an **odd** number of arches.

a. Circle the numbers that could possibly be the length of Neomi's bridge.

15 20 25 27 30 35 37 40 45 .55

b. Do all the circled numbers have anything in common?

7. Oded plans to build a bridge.

a. Explain to him how can he decide how many rods of 5 he needs.

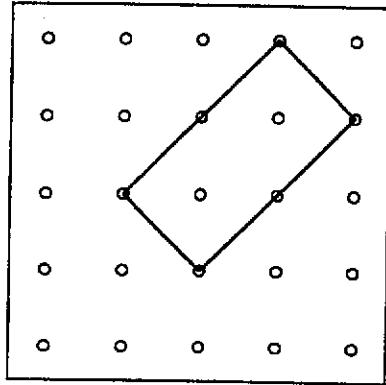
b. Explain to him how can he decide how many rods of 2 he needs.



GEOBOARD QUADRILATERALS

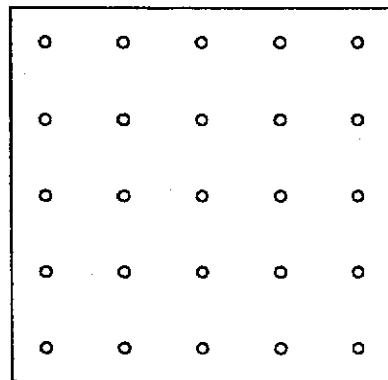
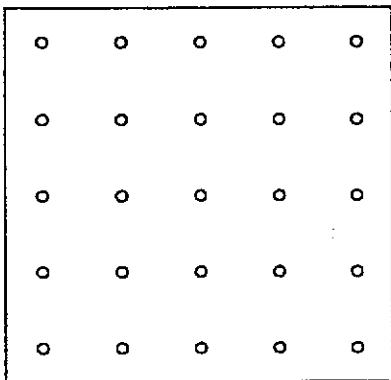
For each question, draw a quadrilateral whose vertices are on the dots and which has the properties described.

Example: A quadrilateral with 4 right angles.



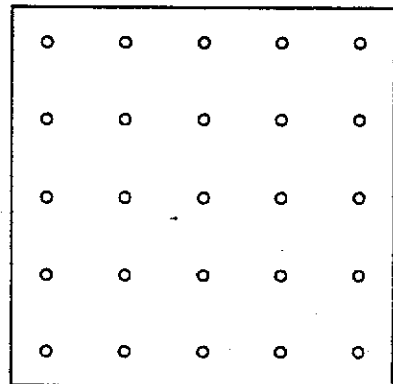
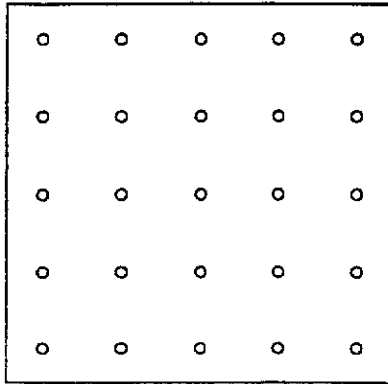
1. A quadrilateral with 2 or more right angles.

(Draw two different examples of such quadrilaterals.)



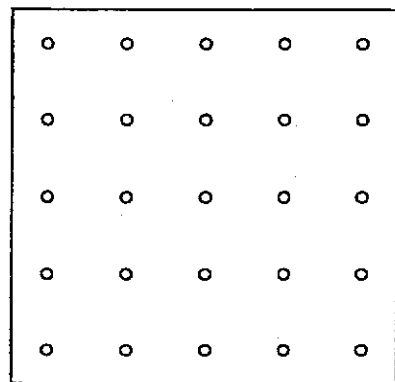
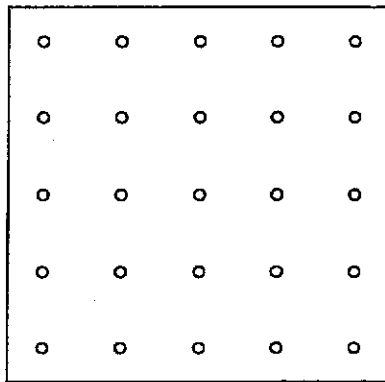
2. A quadrilateral with sides **different in length**.

(Draw two different examples of such quadrilaterals.)



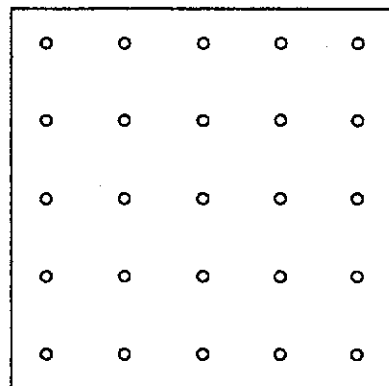
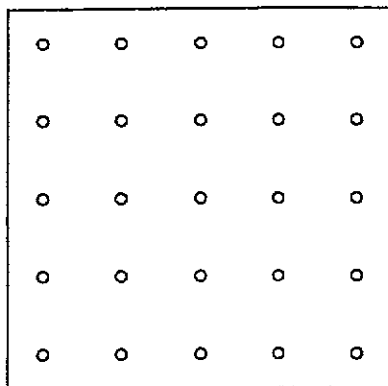
3. A quadrilateral with 2 right angles and 2 equal sides.

(Draw two different examples of such quadrilaterals.)



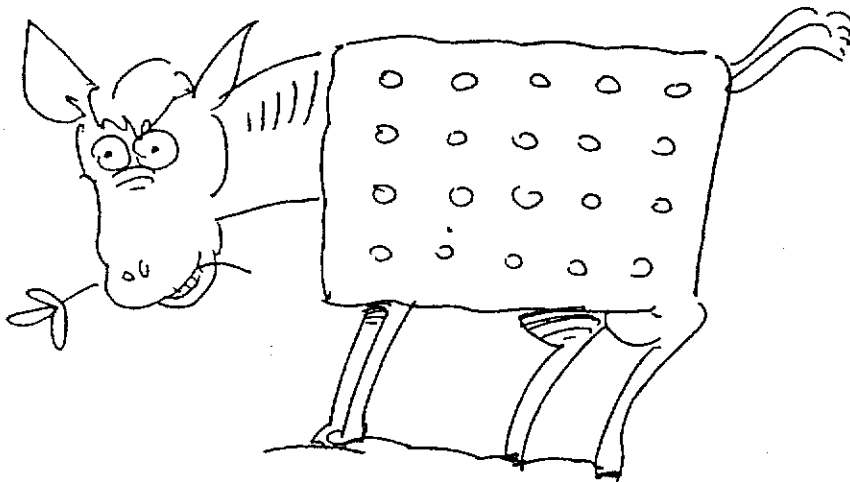
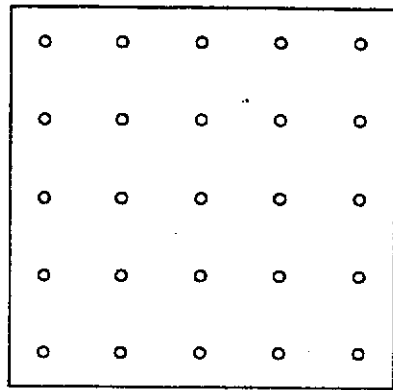
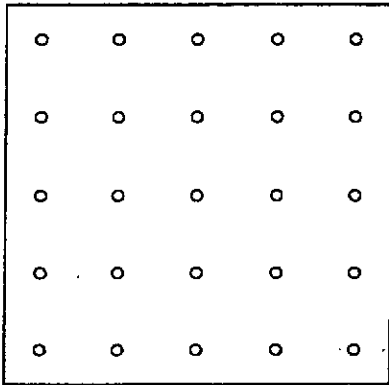
4. A quadrilateral with 2 right angles and all sides different in length.

(Draw two different examples of such quadrilaterals.)



5. A quadrilateral with **only** 2 right angles.

(Draw two different examples of such quadrilaterals.)





my name: _____

my class: _____

Getting Results

1. a. Solve: $(12 - 3) - 2 =$

$$12 - (3 - 2) =$$

b. Why do you think that the two results are different?

2. Solve: $(16 - 6) \times 2 - 1 =$

$$16 - (6 \times 2) - 1 =$$

$$16 - 6 \times (2 - 1) =$$

3. Use parentheses in different ways and solve:

$$17 - 3 \times 4 + 2 =$$

$$17 - 3 \times 4 + 2 =$$

$$17 - 3 \times 4 + 2 =$$



4. Correct the following exercises, by using parentheses, if needed.

$$20 - 5 - 2 = 17$$

$$16 - 6 \times 2 = 4$$

$$7 + 3 \times 4 - 2 = 13$$

$$8 + 16 : 4 \times 2 = 12$$

5. Use **one** pair of parentheses, to get **the largest possible** result:

$$7 \times 4 + 5 - 3 =$$

$$10 - 4 \times 2 + 1 =$$

6. Find a whole number between 1 and 9 and use parentheses, if needed, to get a result as **close as possible to 10**.

$$\square - 4 \times 3 =$$

7. Find a whole number between 1 and 9 and use parentheses, if needed, to get **the largest possible** result.

$$25 - \square \times \triangle =$$

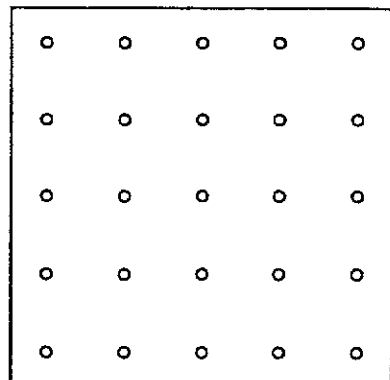
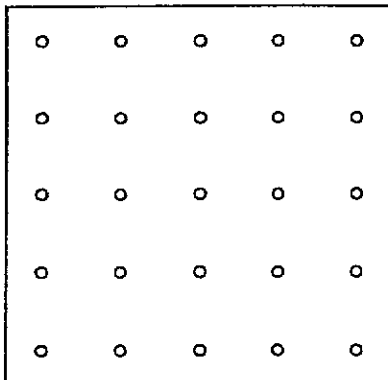
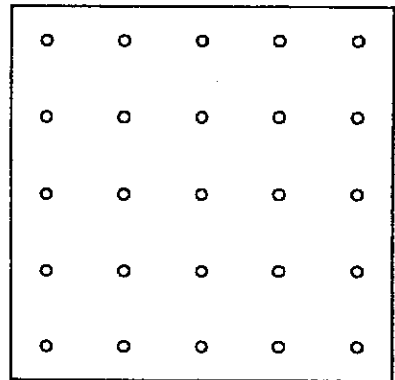
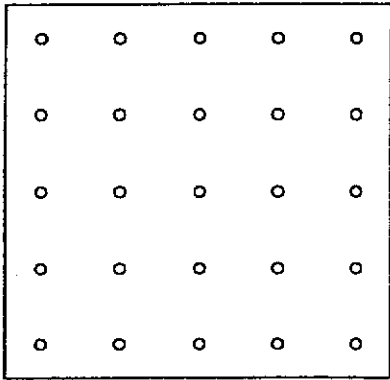
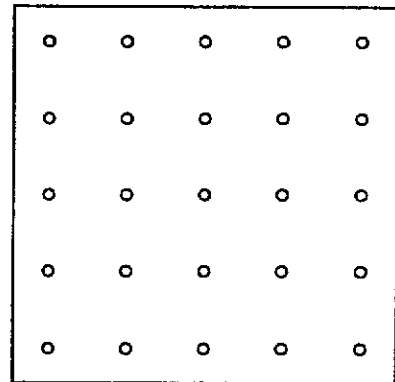
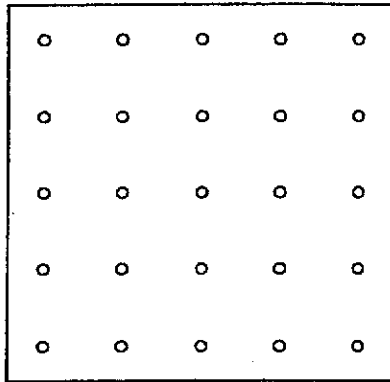


PARALLELOGRAMS

1. In each square, draw a parallelogram with its vertices on the dots.

Draw 6 **different** parallelograms.

(Two parallelograms which coincide with one another after a rotation or a shift are **not** different.)

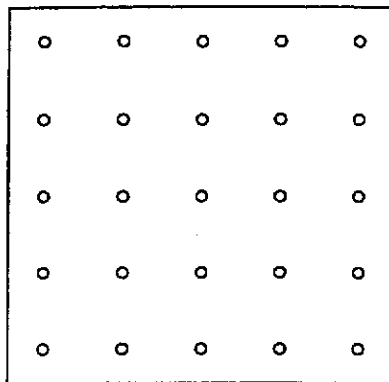


2. Mark in red an acute angle **in each** parallelogram you drew.

(If there are no acute angles, mark x.)

3. Is it possible to draw a parallelogram without even one acute angle?

If so, draw one:

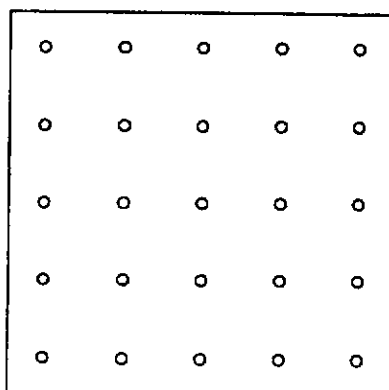


4. Mark in blue one pair of equal sides **in each** parallelogram you drew.

(If there are no equal sides, mark \checkmark .)

5. Is it possible to draw a parallelogram without even one pair of equal sides? _____

If so, draw one:

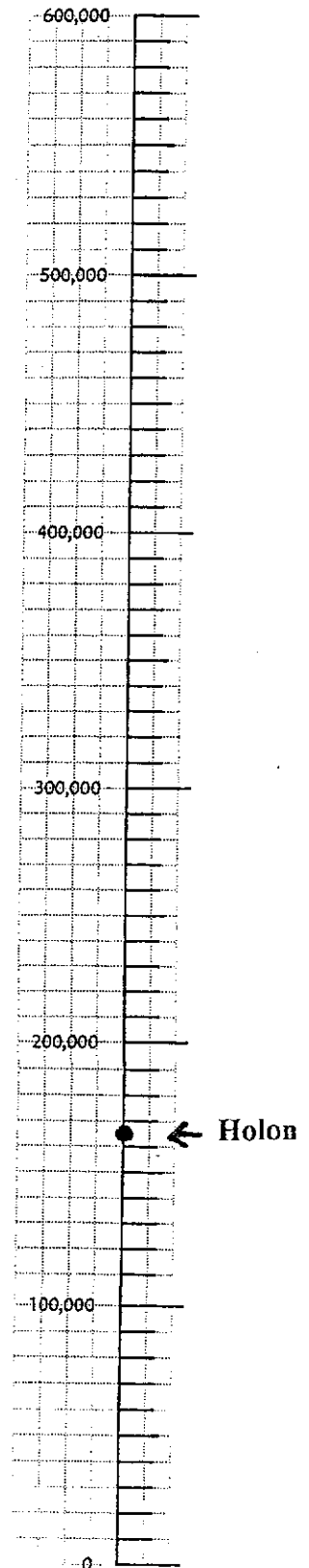


POPULATION

Here is a list of the population of some Israeli cities in 1991:

a. Eilat	29,900
b. Be'er Sheva	128,400
c. Holon	161,800
d. Haifa	251,000
e. Hatzor	7,800
f. Tiberias	37,600
g. Yeruham	6,800
h. Jerusalem	544,200
i. Ma'alot	11,600
j. Akko	46,600
k. Rishon Le'Zion	145,600
l. Tel-Aviv – Jaffa	353,200

1. Mark on the number line an appropriate dot for each city, and write the city's name alongside it. See the example.



2. Divide the list of cities into 3 or 4 groups according to their populations.
Explain how you decided to divide them.

3. What is the approximate population of the three largest cities in Israel together? Circle the closest answer:

12,500

125,000

1,250,000

12,500,000





my name: _____

my class: _____

Table Arrangements

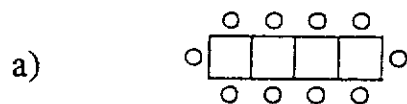
Materials: If you want to, you may use the enclosed 25 cut-out squares.

Suppose that each of the cut-out squares represents a table.

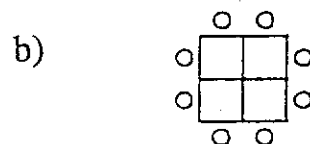
These small square tables can be joined to make larger rectangular tables.

Around such a rectangular arrangement of squares people can sit - where the side of a small square can accommodate one person.

For example, with four square tables we can make the following two rectangular arrangements.

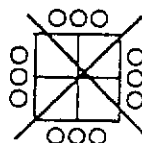


This arrangement can accommodate 10 persons.



This arrangement can accommodate 8 persons.

Notes: A square is also a rectangle (see arrangement b.)
Not more than one person can sit at the side of a small square.



1. At a party, you have 12 square tables.
Use all these tables to make two different rectangular arrangements.
Make a drawing of your arrangements and find the number of persons
that can sit around each of them.

First arrangement:

Number of persons: _____

Second arrangement:

Number of persons: _____



2. At another party, you have 24 square tables.

a) Use all these tables to make two different rectangular arrangements.

Make a drawing of your arrangements and find the number of persons that can sit around each of them.

First arrangement:

Number of persons: _____

Second arrangement:

Number of persons: _____

b) Is there another rectangular arrangement different from the two that you found? _____

If there is, how many small tables are there along the large rectangle and how many across? _____

How many persons can be seated in this arrangement? _____

c) What arrangement would you choose, in order to sit the largest number of people around a rectangular arrangement of 24 small tables?

How many persons can be seated in this arrangement? _____



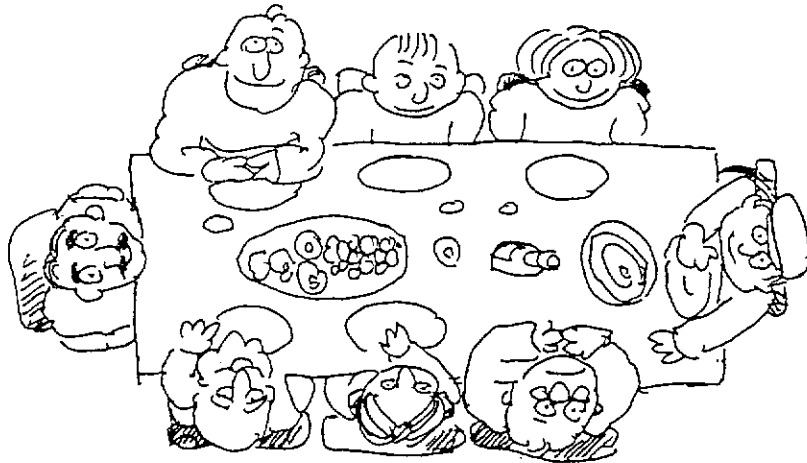
3. Now, suppose that you have 36 square tables.

a) What arrangement would you choose, in order to seat the **largest** number of people around a rectangular arrangement of 36 small tables? _____

How many persons can be seated in this arrangement? _____

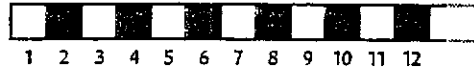
b) What arrangement would you choose, in order to seat the **smallest** number of people around a rectangular arrangement of 36 small tables? _____

How many persons can be seated in this arrangement? _____



TRAINS

1. Here is a “train” of squares. The cars of the train are painted black and white, according to a pattern.



What are the colors of the different cars of the train?

- a) The color of car number 7 is: _____
b) The color of car number 10 is: _____

If we continue the train with the same pattern, what will be the colors of these cars?

- c) The color of car number 18 will be: _____
d) The color of car number 45 will be: _____
e) The color of car number 103 will be: _____
f) Explain how you decided what the color of car number 45 would be (d above):

2. Here is another “train” of squares. The cars of the train are painted black and white according to a different pattern.

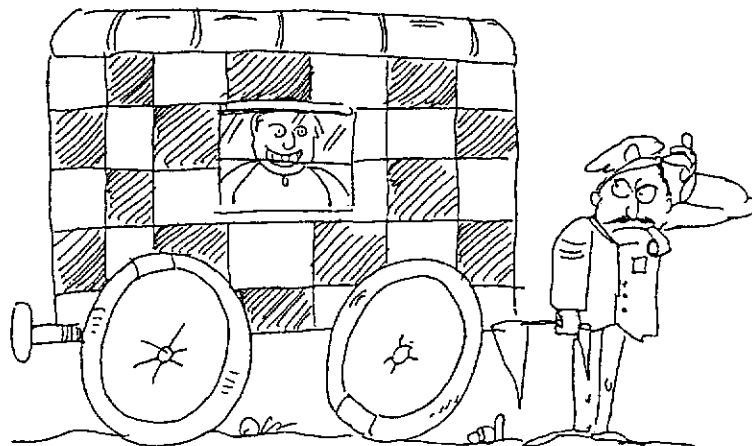


What are the colors of the different cars of the train?

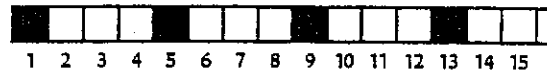
- a) The color of car number 12 is: _____

If we continue the train with the same pattern, what will be the colors of these cars?

- b) The color of car number 16 will be: _____
c) The color of car number 21 will be: _____
d) The color of car number 34 will be: _____
e) The color of car number 102 will be: _____
f) Explain how you decided what the color of car number 34 would be (d above):



3. Here is another “train” of squares. The cars of the train are painted black and white according to a different pattern.

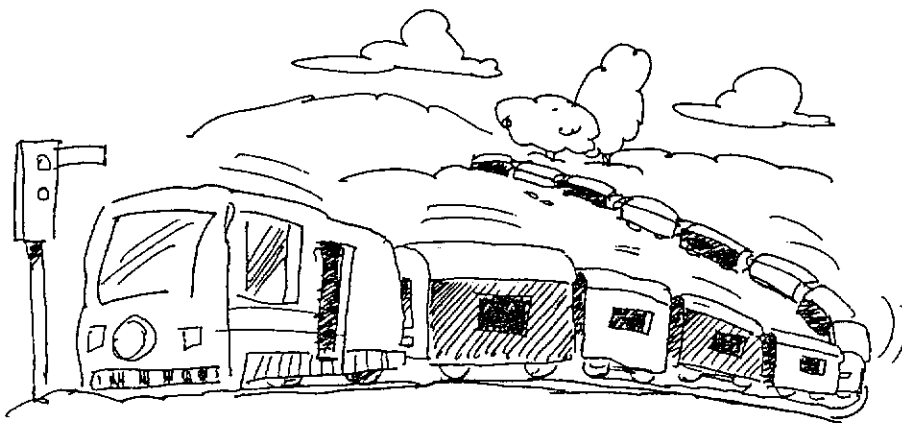


What are the colors of the different cars of the train?

- a) The color of car number 5 is: _____
b) The color of car number 10 will be: _____

If we continue the train with the same pattern, what will be the colors of these cars?

- c) The color of car number 17 will be: _____
d) The color of car number 37 will be: _____
e) The color of car number 79 will be: _____
f) Explain how you decided what the color of car number 37 would be (d above):



Appendix B: Questionnaires

Teacher's Comments – Questionnaire used for pilot of tasks

The First National Course on Assessment

Questionnaire – 2nd meeting

The Second National Course on Assessment

Questionnaire – 6th meeting

The Third National Course on Assessment

Questionnaire – 6th meeting

Questionnaire – 8th meeting

Sharon Project

Background Questionnaire

Follow-up Questionnaire

Teacher's Comments - Questionnaire used for pilot of tasks

1. Describe briefly how the task was used in the classroom (how much time the pupils spent on it, the classroom seating arrangement, etc.).

2. Describe the class atmosphere during the pupils' work on the task.

3. What did you learn from the way the pupils worked, in addition to their written answers?

4. How do you feel about using this task for student assessment?

5. Additional comments and ideas for improving the task.

First National Assessment Course - 2nd meeting (9.3.95)

1. In what type of school did you use the investigative tasks?
2. What were the results of your use of the tasks with the students?
3. Do you think these tasks can be used for assessment?

Second National Course – 6th meeting (21.3.96)

Questionnaire for Math Leaders

Lately we have added criteria for evaluating the tasks, in addition to the answer guide which is attached to each task.

We would like to ask a number of questions which relate to the use of these criteria.

1. How many different tasks have you checked using these criteria? Which tasks?
2. What do you think, in general, about the use of criteria for checking tasks?
3. Did the use of criteria cause:

	very much	some	no difference	definitely not
extending the checking process				
deepening understanding of the pupil's performance				
emphasizing the difference between students				
acquiring more detailed information about the student's level				

4. Which criteria are not clear? That is, which are hard to make decisions about them?
5. Did checking by criteria help you plan your teaching? If yes, give an example.

Third National Assessment Course – 6th meeting (21.11.96)

“Add-a-digit” – an estimation task

1. How was this task different from the tasks that the students usually do?
2. In your experience, does checking tasks with the use of criteria provide you with information that conventional grading doesn't give you? yes / no
If yes, describe this information.
3. Additional comments.

Third National Assessment Course – 8th meeting (27.2.97)

Questionnaire for Math Leaders

Lately we have added criteria for evaluating the tasks, in addition to the answer guide which is attached to each task.

We would like to ask a number of questions which relate to the use of these criteria.

1. How many different tasks have you checked using these criteria? Which tasks?
2. What do you think, in general, about the use of criteria for checking tasks?
3. Did the use of criteria cause:

	very much	some	no difference	definitely not
extending the checking process				
deepening understanding of the pupil's performance				
emphasizing the difference between students				
acquiring more detailed information about the student's level				

4. Which criteria are not clear? That is, which are hard to make decisions about them?
5. Did checking by criteria help you plan your teaching? If yes, give an example.
6. What is different in your class / in your supervision as a result of this course? Explain with examples. (Different mathematical content, different in styles of teaching, different in ways of supervising, etc.)

QUESTIONNAIRE TO TEACHERS IN SHARON

I. Personal information.

A. Teaching Experience:

1. Number of years of teaching experience _____
2. Number of years of teaching mathematics _____
3. This year I teach mathematics in grades _____

B. Education:

1. Graduate of _____ teachers' college
_____ Teaching certificate
_____ B.Ed.
My major was _____
2. Graduate of _____ university
My major was _____

II. Choose one class in which you teach mathematics and answer the following questions:

Which grade is it?

1. What topics did you teach in mathematics in this class last week?
2. Give examples of two tasks which the students were given last week.

III. Ways of assessment:

1. What tools (e.g., tests) do you use to assess your students' mathematics achievements?
2. Why do you check your students' mathematics achievements?

IV. Is it usual in your school for the mathematics staff to work together?

If yes, give details:

1. How often does the staff meet?

- ___ once a week ___ once a month
___ once every two weeks ___ a few times a year

2. Who participates in the meetings?

- ___ all the math teachers of the school
___ all the math teachers of one or two grades

3. What are the subjects usually discussed?

4. What did you do at the last meeting?

V. Requests or expectations from the meetings about assessment.

Follow-up Questionnaire - Sharon Project

Questionnaire (July 1999)

1. Give examples of two tasks that the students performed as class work, from the last week of their studies.
2.
 - a) What tools do you use for student assessment?
 - b) Why do you check student achievement?
3. What was the contribution of our workshops regarding student assessment?
4.
 - a) Did you use a task from the assessment task bank in the last half of the year?
 - b) If so, which task and in which class?

Appendix C: Teachers' Responses to Questionnaires

Extensive Study:

Teacher AF	1
Teacher AG	3
Teacher AH	4
Teacher AL	5
Teacher AV	8
Teacher AY	9
Teacher BA	11
Teacher BC	13
Teacher BD	14
Teacher BF	15
Teacher BG	16
Teacher BJ	18
Teacher BO	20
Teacher BP	23
Teacher BU	26
Teacher BV	30
Teacher BW	33
Teacher CC	36
Teacher CD	39
Teacher CF	42
Teacher CJ	45
Teacher CL	48

Intensive Study:

Teacher DA	51
Teacher DF	54
Teacher DH	56
Teacher DL	58
Teacher DP	61
Teacher EA	63
Teacher EC	65
Teacher ED	67
Teacher EE	70

First National Assessment Course - 2nd meeting (9.3.95)

Teacher AF

1. In what type of school did you use the investigative tasks?

One good school, one integrated, one poor

2. What were the results of your use of the tasks with the students?

All the students had difficulty with the task *Population*. In my opinion the spaces between points on the number line are too little. Most of the pupils had trouble with verbalization – they performed correctly but didn't know how to explain.

I saw weak students who succeeded by using concrete materials – even though their teachers said they were very weak.

3. Do you think these tasks can be used for assessment?

Students need to work on such tasks in class before these tasks are used for assessment.

Second National Course – 6th meeting (21.3.96)

Teacher AF

1. How many different tasks have you checked using these criteria? Which tasks?

2 tasks – *Table 2, Bridges*

2. What do you think, in general, about the use of criteria for checking tasks?

In this way it is possible to check student's thinking – it's possible to check several areas.

3. Did the use of criteria cause:

	very much	some	no difference	definitely not
extending the checking process?		x		
deepening understanding of the pupil's performance?	x			
emphasizing the difference between students?		x		
acquiring more detailed information about the student's level?	x			

4. Which criteria are not clear? That is, which are hard to make decisions about them?

Creativity, verbalization

5. Did checking by criteria help you plan your teaching? If yes, give an example.

Yes – I work more with open problems and put an emphasis on verbalization.

First National Assessment Course - 2nd meeting (9.3.95)

Teacher AG

1. In what type of school did you use the investigative tasks?

Five integrated schools

2. What were the results of your use of the tasks with the students?

The tasks, in general, were hard to do in spite of the fact that on first glance it seemed that the students could succeed. There were students that became discouraged, and even one excellent student quit. There was no correlation between teachers' evaluation of the students and their success on the tasks.

3. Do you think these tasks can be used for assessment?

These tasks can be used as additional assessment tools, and not instead of conventional tests.

I think that teachers need to be given such tasks to use as classwork before they are used for assessment.

The checking of the students' performances takes quite a long time, and is a problem for the teachers.

First National Assessment Course - 2nd meeting (9.3.95)

Teacher AH

1. In what type of school did you use the investigative tasks?

A good school.

2. What were the results of your use of the tasks with the students?

The students, in general, gave good explanations – the teacher has been working for the past few months on oral and written explanations.

3. Do you think these tasks can be used for assessment?

Certainly such tasks can be added to tests, or used alone, on condition that students work in class with similar approaches.

First National Assessment Course - 2nd meeting (9.3.95)

Teacher AL

1. In what type of school did you use the investigative tasks?

An integrated school

2. What were the results of your use of the tasks with the students?

- a) Some of the tasks were easy and clear to the students and some hard.
- b) There were problems with the wording relating to making conclusions.
- c) Not all the students could make conclusions.
- d) Weak pupils couldn't deal with the tasks and stopped in the middle.
- e) In general, the pupils enjoyed the tasks very much, and related to them as interesting work-sheets.
- f) It wasn't clear to the students or to the teachers the connection between the tasks and what was currently being studied.

3. Do you think these tasks can be used for assessment?

- a) It was hard for the teachers to accept these tasks for assessing knowledge and mathematical thinking. They felt that the tasks criticized their teaching – and that if they were for assessment, then the curriculum needs to be changed!
- b) It needs to be made clear that these tasks give new messages, and thus teachers need to do more than teach skills!

Second National Course – 6th meeting (21.3.96)

Teacher AL

1. How many different tasks have you checked using these criteria? Which tasks?

3 tasks – *Table Arrangements, Table 2, Bridges*

2. What do you think, in general, about the use of criteria for checking tasks?
 - a) The criteria help the teacher understand what the child needs to understand and know;
 - b) The emphasis moves from product to process;
 - c) The teacher gets information on the student's way of thinking and performing and how to work with him.

3. Did the use of criteria cause:

	very much	some	no difference	definitely not
extending the checking process?			x	
deepening understanding of the pupil's performance?		x		
emphasizing the difference between students?	x			
acquiring more detailed information about the student's level?	x			

4. Which criteria are not clear? That is, which are hard to make decisions about them?

--

5. Did checking by criteria help you plan your teaching? If yes, give an example.

Yes – for example, in the task *Bridges* there was a problem of collecting all the data in the problem. After the test, I created another task where they needed also to collect data and from the data discover the rule. Like the question: Eti needs to read 90 pages in three days. He decided that on the third day he would read twice what he read on the second day. How many pages did he read on the first day?

First National Assessment Course - 2nd meeting (9.3.95)

Teacher AV

1. In what type of school did you use the investigative tasks?

An integrated school with mostly weak pupils, but some excellent ones.

2. What were the results of your use of the tasks with the students?

After using the tasks with six students, I can comment:

- a) The tasks are scaffolded – easy to hard. The weak students succeeded on the beginning questions, the average students did more, and the strong students finished some tasks with much success.
- b) The students' ability to explain verbally is quite low. They are not able to summarize in writing their explanations, even when their thinking is correct.
- c) The difference between the students arose not so much between the number of tasks they succeed on but rather their level of performance. It enabled getting an interesting picture of the spread of their abilities.
- d) The hardest task was on estimation. Very few children used estimation, rather they calculated everything.

3. Do you think these tasks can be used for assessment?

These tasks can be used for assessment, but first one needs to use similar tasks for teaching in order to familiarize the students with this unusual way of thinking and also to give them the opportunities to explain themselves in writing.

Second National Course – 6th meeting (21.3.96)

Teacher AY

1. How many different tasks have you checked using these criteria? Which tasks?

3 tasks – *Table Arrangements, Trains, Table 2*

2. What do you think, in general, about the use of criteria for checking tasks?

Possible to check if (the) student understands the task, using makes it possible to check different areas of ability.

3. Did the use of criteria cause:

	very much	some	no difference	definitely not
extending the checking process?	x			
deepening understanding of the pupil's performance?		x		
emphasizing the difference between students?		x		
acquiring more detailed information about the student's level?	x			

4. Which criteria are not clear? That is, which are hard to make decisions about them?

Creativity

5. Did checking by criteria help you plan your teaching? If yes, give an example.

a) The students work more with manipulatives, even on tests.

b) I use more open problems.

c) I use less activities which are only exercises—and more do the drill on the computer / and emphasize more activities in which process is important.

d) I require verbalization (and generalization) in the students' work.

Second National Course – 6th meeting (21.3.96)

Teacher BA

1. How many different tasks have you checked using these criteria? Which tasks?

5 tasks – *Table 2, Table 3, Table Arrangements, Trains, Bridges*

2. What do you think, in general, about the use of criteria for checking tasks?

Using criteria helps and focuses, adds a framework

3. Did the use of criteria cause:

	very much	some	no difference	definitely not
extending the checking process?	x			
deepening understanding of the pupil's performance?	x			
emphasizing the difference between students?	x			
acquiring more detailed information about the student's level?		x		

4. Which criteria are not clear? That is, which are hard to make decisions about them?

5. Did checking by criteria help you plan your teaching? If yes, give an example.

The criteria give a clearer picture than just marking right or wrong – after analyzing the student profiles which are generated, one can easily plan for future work – one sees in which areas the students are weak and in which areas average.

In one specific school where I used the task and checked with criteria, I recommended the school's teacher-leader to work with the 4th grade teacher on problem solving strategies.

Second National Course – 6th meeting (21.3.96)

Teacher BC

1. How many different tasks have you checked using these criteria? Which tasks?

4 tasks – *Bridges, Table Arrangements, Table 3, Trains*

2. What do you think, in general, about the use of criteria for checking tasks?

Not so clear, doesn't give the student an immediate assessment.

For me, as an assessor, it helps give focus.

3. Did the use of criteria cause:

	very much	some	no difference	definitely not
extending the checking process?	x			
deepening understanding of the pupil's performance?	x			
emphasizing the difference between students?		x		
acquiring more detailed information about the student's level?			x	

4. Which criteria are not clear? That is, which are hard to make decisions about them?

Understanding the task and understanding the concepts

5. Did checking by criteria help you plan your teaching? If yes, give an example.

It's possible to form groups of students.

Second National Course – 6th meeting (21.3.96)

Teacher BD

1. How many different tasks have you checked using these criteria? Which tasks?

3 tasks – *Bridges, Table 3, Table Arrangements*

2. What do you think, in general, about the use of criteria for checking tasks?

Makes the checking easier – sometimes creates a problem, mostly in checking “thinking” when I want to write a level between high and medium.

3. Did the use of criteria cause:

	very much	some	no difference	definitely not
extending the checking process?				x
deepening understanding of the pupil's performance?	x			
emphasizing the difference between students?	x			
acquiring more detailed information about the student's level?	x			

4. Which criteria are not clear? That is, which are hard to make decisions about them?

thinking

5. Did checking by criteria help you plan your teaching? If yes, give an example.

If I discovered a student who forgot some of the conditions in the problem, then I could teach him how to mark them and then go back and check that he fulfilled all the conditions.

Second National Course – 6th meeting (21.3.96)

Teacher BF

1. How many different tasks have you checked using these criteria? Which tasks?

2 tasks – *Bridges, Table Arrangements*

2. What do you think, in general, about the use of criteria for checking tasks?

Using criteria deepens the knowledge of the student, differentiates between (the) student's lack of knowledge in computation and his way of thinking.

3. Did the use of criteria cause:

	very much	some	no difference	definitely not
extending the checking process?			x	
deepening understanding of the pupil's performance?	x			
emphasizing the difference between students?	x			
acquiring more detailed information about the student's level?	x			

4. Which criteria are not clear? That is, which are hard to make decisions about them?

Not clear the difference between understanding the task and understanding the concepts.

5. Did checking by criteria help you plan your teaching? If yes, give an example.

Yes, to teach the advantages of systematic work, to teach how to carry out a plan, to teach how to make sure you have used all the conditions.

Second National Course – 6th meeting (21.3.96)

Teacher BG

1. How many different tasks have you checked using these criteria? Which tasks?

2 tasks – *Table Arrangements, Bridges*

2. What do you think, in general, about the use of criteria for checking tasks?

I relate very positively to the use of criteria. They enable me to pay attention to elements that previously I didn't grade.

3. Did the use of criteria cause:

	very much	some	no difference	definitely not
extending the checking process?				
deepening understanding of the pupil's performance?	x			
emphasizing the difference between students?	x			
acquiring more detailed information about the student's level?	x			

4. Which criteria are not clear? That is, which are hard to make decisions about them?

I find it hard to decide about the student's ability to abstract – sometimes children draw because they think it's required, even when they could find the answer without a diagram.

5. Did checking by criteria help you plan your teaching? If yes, give an example.

My teaching plan will take into account the students' mistakes and the emphases that I want to make – if the problem is understanding of the task, then I would work on reading comprehension, and if the problem is mistakes in computation, then I would check what was causing the problems.

Second National Course – 6th meeting (21.3.96)

Teacher BJ

1. How many different tasks have you checked using these criteria? Which tasks?

2 tasks – *Table Arrangements, Table 2*

2. What do you think, in general, about the use of criteria for checking tasks?

I think that the use of criteria is important, because then one doesn't just relate to computations – it helps us gain a better understanding of the student's thinking, through his mistakes, why he performed as he did.

3. Did the use of criteria cause:

	very much	some	no difference	definitely not
extending the checking process?		x		
deepening understanding of the pupil's performance?	x			
emphasizing the difference between students?		x		
acquiring more detailed information about the student's level?	x			

4. Which criteria are not clear? That is, which are hard to make decisions about them?

Understanding the task and thinking

5. Did checking by criteria help you plan your teaching? If yes, give an example.

Yes – to relate to each student individually, to build work groups according to the students' levels in relation to the criteria.

For example, if a certain group didn't understand the task, then I can work with them alone.

Third National Assessment Course – 6th meeting (21.11.96)

Teacher BO

1. How was this task different from the tasks that the students usually do?

There is no need for solving computation exercises correctly. Such tasks are usually given to small groups of students (usually four), and this time the student had to work by himself.

2. In your experience, does checking tasks with the use of criteria provide you with information that conventional grading doesn't give you? yes / no

If yes, describe this information.

In my opinion, checking by criteria enables the teacher to look at her students through other lenses.

Criteria checking requires using authentic tasks, which have attributes worthy of looking at.

Checking by criteria enable the student to see exactly what needs to be improved.

3. Additional comments.

We need to encourage the teachers to check with criteria – this will help their teaching, it will improve their testing skills and ultimately, it will improve student achievement.

Third National Assessment Course – 8th meeting (27.2.97)

Teacher BO

1. How many different tasks have you checked using these criteria? Which tasks?

3 tasks – *Getting Results, Bridges, Table Arrangements*

2. What do you think, in general, about the use of criteria for checking tasks?

It's very important to use these criteria. This type of grading gives the teacher quality information about the students. Verbalization makes transparent the learner's thinking and points out mistakes in thinking processes. It causes the student to internalize the material, and to know in which areas he is stronger, in thinking or in computing, etc.

3. Did the use of criteria cause:

	very much	some	no difference	definitely not
extending the checking process?			x	
deepening understanding of the pupil's performance?	x			
emphasizing the difference between students?	x			
acquiring more detailed information about the student's level?	x			

4. Which criteria are not clear? That is, which are hard to make decisions about them?

Thinking, creativity

5. Did checking by criteria help you plan your teaching? If yes, give an example.

In every math topic I try to give expression to as many criteria as I can.

In planning each topic I arrange that each child can have an opportunity to show his creativity, his verbalization, his ability to generalize, so that assessment will reflect the process of teaching.

6. What is different in your class / in your supervision as a result of this course? Explain with examples. (Different mathematical content, different in styles of teaching, different in ways of supervising, etc.)

The criteria for assessment have become a natural part of my work with teachers – the ability to verbalize, both orally and in writing.

The features of the assessment tasks have shaped the activities in the topics which are taught in class.

Third National Assessment Course – 6th meeting (21.11.96)

Teacher BP

1. How was this task different from the tasks that the students usually do?

The task was scaffolded in regard to the levels of thinking required. The task combined different topics: place value, multiplication, and estimation. The student needed to explain his solution. The student needed to choose a strategy for his solution.

2. In your experience, does checking tasks with the use of criteria provide you with information that conventional grading doesn't give you? yes / no

If yes, describe this information.

Conventional grading tells us on which questions the students succeeded/failed, and how many students succeeded, whereas using criteria provides details on what is needed to help each student improve his performance. It gives us which area in the task needs improvement – for example, on questions which needed average mathematical thinking the students succeeded versus those which required higher levels of thinking.

3. Additional comments.

The levels (high, medium, low) for judging each criterion were insufficient for me because they didn't give a clear picture of each part that I checked. Detailed word criteria would help me build more suitable tasks and help the student understand what is required.

Third National Assessment Course – 8th meeting (27.2.97)

Teacher BP

1. How many different tasks have you checked using these criteria? Which tasks?

4 tasks – *Getting Results, Bridges, Trains, Add-a-Digit*

2. What do you think, in general, about the use of criteria for checking tasks?

Use of criteria aids and deepens the work of the student. The problem of the criteria was their generality. For average students I had trouble giving an exact assessment.

3. Did the use of criteria cause:

	very much	some	no difference	definitely not
extending the checking process?				
deepening understanding of the pupil's performance?	x			
emphasizing the difference between students?	x			
acquiring more detailed information about the student's level?	x			

4. Which criteria are not clear? That is, which are hard to make decisions about them?

I found creativity hard – to decide if it's creativity or divergent thinking.

For justification and explanation it was hard assessing the students.

5. Did checking by criteria help you plan your teaching? If yes, give an example.
- The criteria directed me to areas in which the child needed more work – for example, verbalization, to write about the way you worked. The criterion of following instructions trained me to give students the instructions in various ways, when these raised the level of difficulty of the task.
6. What is different in your class / in your supervision as a result of this course? Explain with examples. (Different mathematical content, different in styles of teaching, different in ways of supervising, etc.)
- a) Because of the course and the assessment tasks, I have begun to assess the pupils with different tools and not only with the standard test.
 - b) Because of the course I began using open investigative tasks, which are not directly connected to the topic being studied.
 - c) In building a test, I look at the topic's progression and not just give items which are not related one to another.
 - d) On worksheets and on tests, I try to include various topics.
 - e) I give tasks to heterogeneous groups and not just homogeneous ones.

Third National Assessment Course – 6th meeting (21.11.96)

Teacher BU

1. How was this task different from the tasks that the students usually do?
 - a) Question 1 is familiar but I would prefer to give four possibilities:
 - Completely false
 - Completely true
 - Almost true but false
 - Almost false but true
 - b) About question 2 – usually we ask the children to write correct / not correct and here they we asked to write appropriate / not appropriate. It's too bad that they weren't asked to give an explanation which would have prevented their copying.
 - c) In the answer key there were patterns, but I don't remember such things for estimation. Question 3 was very challenging.
 - d) Question 4 was half-open and very hard. It's too bad that there wasn't a similar question which allowed the weaker students a chance to succeed. The new part of question 4 was "explain ..." record and verbalize the process. The task was scaffolded in regard to the levels of thinking required. The task combined different topics: place value, multiplication, and estimation. The student needed to explain his solution. The student needed to choose a strategy for his solution.

2. In your experience, does checking tasks with the use of criteria provide you with information that conventional grading doesn't give you? yes / no
If yes, describe this information.

Certainly using criteria gives us an additional dimension, especially if we write comments about how the student can improve his performance.

3. Additional comments.

The levels (high, medium, low) for judging each criterion were insufficient for me because they didn't give a clear picture of each part that I checked.

Detailed word criteria would help me build more suitable tasks and help the student understand what is required.

Third National Assessment Course – 8th meeting (27.2.97)

Teacher BU

1. How many different tasks have you checked using these criteria? Which tasks?

3 tasks – *Table 3, Add-a-Digit, Getting Results*

2. What do you think, in general, about the use of criteria for checking tasks?

I think there is a need for qualitative criteria, but there is still confusion about some of them. The clear ones are: understanding the task, understanding the concepts, mathematical thinking and computation.

It's very hard for me to evaluate creativity.

In general, rating criteria on levels doesn't help us know what exactly needs treatment in each criterion.

3. Did the use of criteria cause:

	very much	some	no difference	definitely not
extending the checking process?			x	
deepening understanding of the pupil's performance?	x			
emphasizing the difference between students?	x			
acquiring more detailed information about the student's level?	x			

4. Which criteria are not clear? That is, which are hard to make decisions about them?

Creativity

5. Did checking by criteria help you plan your teaching? If yes, give an example.

Yes! If according to the check it becomes clear that the pupil doesn't know how to verbalize, generalize or justify, then we need to work with him in those areas.

6. What is different in your class / in your supervision as a result of this course? Explain with examples. (Different mathematical content, different in styles of teaching, different in ways of supervising, etc.)

In my class and also in my work with teachers, I emphasize several aspects:

- a) mathematical discussions – pupil with teacher, pupil with pupil, group with group, or the entire class
- b) investigative tasks – at least one each week, accompanied by follow-up work
- c) building tests which are scaffolded (not just on one topic) which have not only one correct answer, which require the need to explain, to justify or to judge
- d) the report card still doesn't include criteria, but I would like to know how to do it so I can recommend it to my school.

Third National Assessment Course – 6th meeting (21.11.96)

Teacher BV

1. How was this task different from the tasks that the students usually do?

The task asks things which we usually only ask of the best students, and here we gave the task to the whole class.

The task requires explaining the way the student worked, and writing the conclusion.

2. In your experience, does checking tasks with the use of criteria provide you with information that conventional grading doesn't give you? yes / no

If yes, describe this information.

Checking with criteria give more detailed information, but there are no "surprises".

3. Additional comments.

Third National Assessment Course – 8th meeting (27.2.97)

Teacher BV

1. How many different tasks have you checked using these criteria? Which tasks?

4 tasks – *Add-a-Digit, Parallelograms, Getting Results, Trains*

2. What do you think, in general, about the use of criteria for checking tasks?

I haven't used criteria grading in classes, but just under experimental conditions – but for teachers who have to give grades on report cards, it seems that there will be problems.

3. Did the use of criteria cause:

	very much	some	no difference	definitely not
extending the checking process?	x			
deepening understanding of the pupil's performance?	x			
emphasizing the difference between students?	x			
acquiring more detailed information about the student's level?	x			

4. Which criteria are not clear? That is, which are hard to make decisions about them?

Creativity, communication

5. Did checking by criteria help you plan your teaching? If yes, give an example.

Certainly yes – for example, if I discovered students who are not managing to reach the right conclusions, then I would give them an easier task and meanwhile try to work with them and little by little bring them to higher levels.

In verbalization, if I discover that students are not verbalizing their solutions we will work together in class on verbalization skills and have many more mathematical discussions, which can be written on transparencies and posters.

6. What is different in your class / in your supervision as a result of this course? Explain with examples. (Different mathematical content, different in styles of teaching, different in ways of supervising, etc.)

- a) In the workshops which I run I implement the ways of working which I learned in this course (copying your way of working for example by opening the workshop with teachers doing an activity, reporting on transparencies, etc.)
- b) In working with teachers, I require more openness in tasks given to students, more “investigating” their mistakes
- c) Change in the mathematical content – less emphasis on basic mathematical skills and more emphasis on finding rules and generalizations (the same topics but tackling at them from a different angle)
- d) Assessment – not yet checking with criteria, but in the process of checking tests with teachers I ask them to relate to specific criteria – concepts, generalizations, etc.

Third National Assessment Course – 6th meeting (21.11.96)

Teacher BW - was not present at this meeting

1. How was this task different from the tasks that the students usually do?
2. In your experience, does checking tasks with the use of criteria provide you with information that conventional grading doesn't give you? yes / no
If yes, describe this information.
3. Additional comments.

Third National Assessment Course – 8th meeting (27.2.97)

Teacher BW

1. How many different tasks have you checked using these criteria? Which tasks?

4 tasks – *Getting Results, Population, Add-a-Digit, Trains*

2. What do you think, in general, about the use of criteria for checking tasks?

The criteria are good, the “ratings” are hard.

The criteria help very much in knowing what to emphasize, and enable making definite conclusions for work with children.

3. Did the use of criteria cause:

	very much	some	no difference	definitely not
extending the checking process?	x			
deepening understanding of the pupil's performance?	x			
emphasizing the difference between students?	x			
acquiring more detailed information about the student's level?	x			

4. Which criteria are not clear? That is, which are hard to make decisions about them?

Creativity and communication. It's very hard to check creative ability.

We need in-between measures for checking creativity.

5. Did checking by criteria help you plan your teaching? If yes, give an example.

Certainly yes. It gives a very clear picture of what I need to emphasize in teaching, what is missing in my math lessons. The thing which stands out is the children's ability to justify.

6. What is different in your class / in your supervision as a result of this course? Explain with examples. (Different mathematical content, different in styles of teaching, different in ways of supervising, etc.)

In working with teachers, I emphasize meta-cognitive processes – the great need for verbalization and mathematical discussions with the students.

I work with them on the need to give investigative tasks of different levels, and **all** students in the class should work on such tasks in heterogeneous groups, so they can help one another. I am still working on myself and the teachers on the need for alternative assessment in addition to conventional assessment. I learned in this course ways of “opening” mathematical pages, and making them better by adding verbalization, reversed thinking, and giving more than one answer to a task.

Third National Assessment Course – 6th meeting (21.11.96)

Teacher CC

1. How was this task different from the tasks that the students usually do?

The need to explain in words their thinking.

2. In your experience, does checking tasks with the use of criteria provide you with information that conventional grading doesn't give you? yes / no

If yes, describe this information.

Conventional grading gives the student a grade and rates him on the class profile. Checking by criteria places the emphasis on what needs to be improved in order to get a better grade, and on the level of knowledge of the student.

3. Additional comments.

- The students enjoyed very much the task
- The students lack strategies for dealing with exercises that are unfamiliar to them
- All the students have problems in written explanations.

Third National Assessment Course – 8th meeting (27.2.97)

Teacher CC

1. How many different tasks have you checked using these criteria? Which tasks?

3 tasks – *Getting Results, Add-a-Digit, Trains*

2. What do you think, in general, about the use of criteria for checking tasks?

The use of different criteria gave us the ability to relate in another way to each pupil. To differentiate between the pupils and to discover new faces of the pupil (that we hadn't seen in frontal teaching). Also, the pupil learned to relate differently to the subject when he received compliments which he was previously not used to getting. Thus we all benefited.

3. Did the use of criteria cause:

	very much	some	no difference	definitely not
extending the checking process?				x
deepening understanding of the pupil's performance?	x			
emphasizing the difference between students?	x			
acquiring more detailed information about the student's level?	x			

4. Which criteria are not clear? That is, which are hard to make decisions about them?
 - a) Explanation of the process – it was hard to evaluate the explanation (even when it was correct). Who should get a higher rating – a student who explains briefly but with mathematical language or one who uses more words and gives examples?

- b) Creativity – should one give a higher rating for originality / unusual (even I didn't think of it) or a similar rating to one who just shows some creativity at all?
5. Did checking by criteria help you plan your teaching? If yes, give an example.
Certainly. Using the criteria taught me the value of mathematical discussions, to work on one's ability to describe one's math doings, to be exact in verbalizing the processes and the conclusions.
6. What is different in your class / in your supervision as a result of this course? Explain with examples. (Different mathematical content, different in styles of teaching, different in ways of supervising, etc.)
- a) mathematical discussions in groups
 - b) focus on verbalization and conceptualization of processes
 - c) focus on process and not just on product
 - d) adding open questions to tests
 - e) adding investigations once a week
 - f) trying to give the students tools so they can solve problems without the teacher
 - g) using investigations with groups.

Third National Assessment Course – 6th meeting (21.11.96)

Teacher CD

1. How was this task different from the tasks that the students usually do?

The task dealt with one topic – estimation.

The task was scaffolded. The students needed to explain in words the solution to an exercise.

2. In your experience, does checking tasks with the use of criteria provide you with information that conventional grading doesn't give you? yes / no

If yes, describe this information.

Conventional grading helped in checking by criteria.

There were a few children whose conventional grades were low in computation, but who made correct conclusions in their explanations.

Checking by criteria helped to understand thinking processes throughout the whole task, and not just on single items.

Checking by criteria gives a clear picture of factors which are important to us in relation to mathematics – for example, understanding concepts, understanding instructions ... in relation to the whole task and not just on individual items.

3. Additional comments.

Third National Assessment Course – 8th meeting (27.2.97)

Teacher CD

1. How many different tasks have you checked using these criteria? Which tasks?

4 tasks – *Add-a-Digit, Table 3, Quadrilaterals, Getting Results*

2. What do you think, in general, about the use of criteria for checking tasks?

The use of criteria add dimensions which weren't visible in conventional grading, and they are important – for example, creativity, the ability to verbalize, the ability to generalize, etc.

3. Did the use of criteria cause:

	very much	some	no difference	definitely not
extending the checking process?				
deepening understanding of the pupil's performance?	x			
emphasizing the difference between students?			x	
acquiring more detailed information about the student's level?	x			

4. Which criteria are not clear? That is, which are hard to make decisions about them?

Justification, generalization, creativity – everything connected to verbal answers.

5. Did checking by criteria help you plan your teaching? If yes, give an example.

Certainly. After using criteria for checking, I have been looking for ways to improve the mathematical language of the students and their ability to express themselves.

Similarly, ways to improve their mathematical thinking – to make it more organized, more systematic – and how to order their prior knowledge.

6. What is different in your class / in your supervision as a result of this course? Explain with examples. (Different mathematical content, different in styles of teaching, different in ways of supervising, etc.)

The course introduced me to the test-bank, and to assessment which was unknown to me.

My work with teachers has consequently changed – new ways of assessment force us to change our ways of teaching.

To think more about authentic tasks in mathematics, to try them, so that the subject will be familiar and not just appear on assessment tasks.

I ask more “why did I get this result” and not just “which” results did I get.

The children learned to investigate, to discover, and also the need to justify what they have found. (Also the teachers have learned these things.)

Third National Assessment Course – 6th meeting (21.11.96)

Teacher CF

1. How was this task different from the tasks that the students usually do?

The task only related to one central topic – estimation in multiplication. The task was scaffolded from easy to hard, with different aspects of the topic. The task required an explanation and justification from the child.

2. In your experience, does checking tasks with the use of criteria provide you with information that conventional grading doesn't give you? yes / no

If yes, describe this information.

Using criteria help one know at which stage of the topic the student is stuck. That is because the task is scaffolded. Using criteria gives the teacher a wider picture of the mathematical ability of the student since the criteria evaluate elements which previously we didn't check. Thus, this evaluation helps the teacher work with the student in his problem areas.

I found that the area most problematic for the students was communication, and thus it requires work.

3. Additional comments.

I found it hard to check for all the criteria – for example, variety and creativity.

The assessment of children who didn't answer one question is very different from those who made mistakes on the same question.

There are children who can explain themselves, and there are those who cannot.

Third National Assessment Course – 8th meeting (27.2.97)

Teacher CF

1. How many different tasks have you checked using these criteria? Which tasks?

5 tasks – Population, Bridges, Getting Results, Table Arrangements, Trains, Parallelograms

2. What do you think, in general, about the use of criteria for checking tasks?

The criteria give the teacher a fuller picture of the understandings and performance of the pupils.

In my class I gave the pupils the criteria I planned to use for evaluating their work and already after their first piece of work they tried to meet the standards.

3. Did the use of criteria cause:

	very much	some	no difference	definitely not
extending the checking process?	x			
deepening understanding of the pupil's performance?	x			
emphasizing the difference between students?	x			
acquiring more detailed information about the student's level?	x			

4. Which criteria are not clear? That is, which are hard to make decisions about them?

The differences in levels in creativity.

More exact assessment for verbalization, generalization, and justification.

Also, for computations sometimes I have difficulties because the explanation behind is missing, and I don't know why the student computed thus – in order for the assessment to be more qualitative.

5. Did checking by criteria help you plan your teaching? If yes, give an example.

After checking the *Population* task, I concluded that I needed sorting activities in mathematics, as well as activities on determining sorting categories. Some pupils performed well on the task but in question 3 which asked for the (approximate) number of people in the three largest cities together, they wrote 125,000. I learned that I need to work on number sense for large numbers, that is, different ways of understanding what a million is.

6. What is different in your class / in your supervision as a result of this course? Explain with examples. (Different mathematical content, different in styles of teaching, different in ways of supervising, etc.)

My ways of teaching have changed in the last few years, as a result of workshops with Teacher PPP and exposure to new approaches to mathematics in varied workshops.

Today my students study differently with the focus on activities on thinking, investigations, open problems.

As a result of this course, I emphasize more verbalization – to describe the process of your work and the solution, to describe your original plan and hypothesis.

My way of assessing has changed to the use of criteria, and also to other assessment tools. The tests for checking knowledge that I “still” use have changed their format, and now have items which the pupil needs to justify, to explain why he thinks something or to reach a generalization.

Third National Assessment Course – 6th meeting (21.11.96)

Teacher CJ

1. How was this task different from the tasks that the students usually do?

What is unusual is the situation where the task gives the student the answer and he needs to create the exercise. This use of “reversed thinking” on a high level requires the student to think and experiment and does not let him use a “formula” he has learned already.

2. In your experience, does checking tasks with the use of criteria provide you with information that conventional grading doesn't give you? yes / no

If yes, describe this information.

I feel checking by the use of criteria is still a problem for me – coming to conclusions, and just understanding the criteria. I think that I have always used criteria for checking student work, without knowing it.

3. Additional comments.

-

Third National Assessment Course – 8th meeting (27.2.97)

Teacher CJ

1. How many different tasks have you checked using these criteria? Which tasks?

3 tasks – *Trains, Getting Results, Add-a-Digit*

2. What do you think, in general, about the use of criteria for checking tasks?

This direction, in general, seems to me to be correct. Turning attention from computation to the way of thinking is critical, and leads to a way of teaching which seems to me to be better. But I ran into many situations which were not one-sided and it was hard for me to decide, in particular with average students.

3. Did the use of criteria cause:

	very much	some	no difference	definitely not
extending the checking process?	x			
deepening understanding of the pupil's performance?	x			
emphasizing the difference between students?	x			
acquiring more detailed information about the student's level?	x			

4. Which criteria are not clear? That is, which are hard to make decisions about them?

Creativity, understanding concepts

5. Did checking by criteria help you plan your teaching? If yes, give an example.

6. What is different in your class / in your supervision as a result of this course? Explain with examples. (Different mathematical content, different in styles of teaching, different in ways of supervising, etc.)

- a) The main difference is more work in groups with emphasis on dialogue between the pupils.
- b) We plan longer tasks which are scaffolded.
- c) We put more emphasis on having the students explain their thinking.
- d) We plan tasks which include different math topics – more complicated tasks.
- e) I have the students write on transparencies; this encourages more precise work, more focused and more aesthetic.

I am generally pressured to accomplish the material in the curriculum, and now I feel less pressured.

I give less time and less importance to regular tests.

I still don't use the tasks as a way of alternative assessment, rather as a way of work in the class. I use the tasks for working with the fourth grade teachers.

Everything that I do in my classes (this year second and sixth grades), I teach the other teachers in my school.

Third National Assessment Course – 6th meeting (21.11.96)

Teacher CL

1. How was this task different from the tasks that the students usually do?

The task was scaffolded from easy to hard, and varied.

A written explanation was required.

2. In your experience, does checking tasks with the use of criteria provide you with information that conventional grading doesn't give you? yes / no

If yes, describe this information.

Using criteria helps us see if the student understood the task, and if he worked on the basis of understanding or just worked by trial and error. It gives us a broader picture than just his computations, and helps us understand if the student really understood or just happened to succeed.

3. Additional comments.

When assessing through the use of criteria I saw that most of the children have trouble explaining.

Third National Assessment Course – 8th meeting (27.2.97)

Teacher CL

1. How many different tasks have you checked using these criteria? Which tasks?

4 tasks – *Add-a-Digit, Getting Results, Table Arrangements, Population*

2. What do you think, in general, about the use of criteria for checking tasks?

I think that using criteria gives a teacher more leeway in evaluating a pupil. The assessment doesn't stop with just giving a right answer, rather relates also to the level of the answer, to the amount of effort in the pupil's thinking. It enables knowing the pupils and their thinking better.

3. Did the use of criteria cause:

	very much	some	no difference	definitely not
extending the checking process?	x			
deepening understanding of the pupil's performance?	x			
emphasizing the difference between students?	x			
acquiring more detailed information about the student's level?	x			

4. Which criteria are not clear? That is, which are hard to make decisions about them?

Communication – each teacher will look differently at a verbal explanation which can be expressed briefly or at length and can be correct or not.

Creativity – the creativity of teachers is different and thus here also there can be great differences between two teachers' assessments.

5. Did checking by criteria help you plan your teaching? If yes, give an example.

Evaluating by criteria enables me to determine which areas need to be emphasized in class and with which students.

For example, if I noticed difficulties with formulating generalizations or justifications, I would need to emphasize tasks which require these.

6. What is different in your class / in your supervision as a result of this course? Explain with examples. (Different mathematical content, different in styles of teaching, different in ways of supervising, etc.)

As a result of this course I've added new ways of assessing, both in my class and in my work with teachers. We've moved to a final task instead of a final test.

Also, my teaching now stresses ways of thinking.

I give tasks which are not directly related to the curriculum. These tasks are analyzed afterwards in class and conclusions are made.

Background Questionnaire to Teachers in Sharon – 1st meeting (27.1.97)

Teacher DA

I. Personal information.

A. Teaching Experience:

1. Number of years of teaching experience 24
2. Number of years of teaching mathematics 22
3. This year I teach mathematics in grades 4, 5

B. Education:

1. Graduate of Kibbutz Seminary teachers' college
 Teaching certificate
 B.Ed.
My major was _____
2. Graduate of _____ university
My major was _____

II. Choose one class in which you teach mathematics and answer the following questions:

Which grade is it? 4th

1. What topics did you teach in mathematics in this class last week?
Solving complex problems in mathematics.
2. Give examples of two tasks which the students were given last week.
 - a) Verbalization of problems without giving their solutions.
 - b) Weekly thinking puzzle in arithmetic

III. Ways of assessment:

1. What tools (e.g., tests) do you use to assess your students' mathematics achievements?
Tasks – like solving a problem.
A test with exercises for checking.
2. Why do you check your students' mathematics achievements?
Where the class is at in the syllabus.
The amount of understanding of each pupil.
Assessing the student's work.

IV. Is it usual in your school for the mathematics staff to work together? yes

If yes, give details:

1. How often does the staff meet?

once a week once a month
 once every two weeks a few times a year

2. Who participates in the meetings?

all the math teachers of the school
 all the math teachers of one or two grades

3. What are the subjects usually discussed?

Planning lessons, sharing materials from the in-service workshops

4. What did you do at the last meeting?

We solved mathematical problems, challenges, that we are going to use
in the class

V. Requests or expectations from the meetings about assessment.

How to assess in other ways, not using a test.

End-of-Year Questionnaire to Teachers in Sharon (July, 1999)

Teacher DA – 6th grade

1. Give examples of two tasks that the students performed as class work, from the last week of their studies.

a) build regular polyhedrons from the set of triangles
build regular polyhedrons from the “black set”
nets for regular polyhedrons

b) tasks from the textbook, investigating regular polyhedrons to find the number of vertices, sides, and faces

2. a) What tools do you use for student assessment?

Mostly tasks from the Assessment Task Bank, or I write tests like the end of the chapter summaries in the textbook and which include all the material from the chapter.

b) Why do you check student achievement?

Today I check student achievement to see the progress of mathematical thinking, and to see how each student from his place has progressed. I check the students’ understandings in the topics we studied and this also is feedback for me.

3. What was the contribution of our workshops regarding student assessment?

The biggest contribution was that I stopped assessing students with grades. Building new types of tests which assess understanding, like understanding concepts, investigations. Dealing with situations which are relevant to the students’ lives, and also building additional criteria for assessment like generalizations, justifications, etc.

Background Questionnaire to Teachers in Sharon – 1st meeting (27.1.97)

Teacher DF

I. Personal information.

A. Teaching Experience:

1. Number of years of teaching experience 23
2. Number of years of teaching mathematics 23
- 3.

This year I teach mathematics in grades 5

B. Education:

- X 1. Graduate of Beit Berl teachers' college
X Teaching certificate
_____ B.Ed.

My major was _____

- X 2. Graduate of Bar Elan university
My major was special education

II. Choose one class in which you teach mathematics and answer the following questions:

Which grade is it? 5th

1. What topics did you teach in mathematics in this class last week?
Order between fractions. Diagonals in polygons.
2. Give examples of two tasks which the students were given last week.
 - a) Order the following fractions by size from the smallest to the largest:
 $1/10$, $1/4$, $1/12$. Use your fraction rulers to help if you need them.
 - b) Build on the geoboard a quadrilateral with perpendicular diagonals.

III. Ways of assessment:

1. What tools (e.g., tests) do you use to assess your students' mathematics achievements?
Tests, checking daily work, computer drill work.
2. Why do you check your students' mathematics achievements?
Mainly in order to plan.

IV. Is it usual in your school for the mathematics staff to work together? yes

If yes, give details:

1. How often does the staff meet?

- once a week ___ once a month
___ once every two weeks ___ a few times a year

2. Who participates in the meetings?

- ___ all the math teachers of the school
 all the math teachers of one or two grades

3. What are the subjects usually discussed?

Planning lessons

4. What did you do at the last meeting?

Talked about what we had taught, planned future lessons

V. Requests or expectations from the meetings about assessment.

Background Questionnaire to Teachers in Sharon – 1st meeting (27.1.97)

Teacher DH

I. Personal information.

A. Teaching Experience:

1. Number of years of teaching experience 27
2. Number of years of teaching mathematics 27
3. This year I teach mathematics in grades 5

B. Education:

- X 1. Graduate of Shine teachers' college
X Teaching certificate
_____ B.Ed.
My major was literature, Bible
- X 2. Graduate of Hebrew university
My major was Bible, literature

II. Choose one class in which you teach mathematics and answer the following questions:

Which grade is it? 5th

1. What topics did you teach in mathematics in this class last week?
Fractions on the number line, addition and subtraction
Challenge problems
Geometry – the sum of the angles in regular polygons
2. Give examples of two tasks which the students were given last week.
 - a) Arithmetic -- place the following fractions on the number line:
 $\frac{1}{2}$, $\frac{1}{3}$, $\frac{5}{5}$, $\frac{13}{12}$, $\frac{0}{4}$ and explain how you decided where to place them
 - b) Geometry – to find the sum of the angles in an octagon

III. Ways of assessment:

1. What tools (e.g., tests) do you use to assess your students' mathematics achievements?

Individual report after a lesson

Checking students' knowledge in solving exercises while sitting in a small group

2. Why do you check your students' mathematics achievements?

Checking knowledge to enable each child to progress from the point where he is at

To check the students' progress in relation to the curriculum

IV. Is it usual in your school for the mathematics staff to work together? yes

If yes, give details:

1. How often does the staff meet?

once a week once a month
 once every two weeks a few times a year

2. Who participates in the meetings?

all the math teachers of the school
 all the math teachers of one or two grades

3. What are the subjects usually discussed?

Different approaches to teach a subject

Comments on a lesson which was given

Planning lessons

4. What did you do at the last meeting?

Planned the new topic in geometry – heights and areas

V. Requests or expectations from the meetings about assessment.

A simple easy fast way of assessing

Background Questionnaire to Teachers in Sharon – 1st meeting (27.1.97)

Teacher DL

I. Personal information.

A. Teaching Experience:

1. Number of years of teaching experience 25
2. Number of years of teaching mathematics 25
3. This year I teach mathematics in grades 5

B. Education:

- X 1. Graduate of Kibbutz Seminary teachers' college
X Teaching certificate
_____ B.Ed.
My major was mathematics
- _____ 2. Graduate of _____ university
My major was _____

II. Choose one class in which you teach mathematics and answer the following questions:

Which grade is it? 5th

1. What topics did you teach in mathematics in this class last week?

Arithmetic – addition and subtraction of fractions whose denominators have no common factors.

Geometry – perimeter and area of rectangles and squares.

2. Give examples of two tasks which the students were given last week.

a) $1\frac{1}{2} + 3\frac{1}{3} =$

b) An investigative problem about perimeters and areas of rectangles and squares.

III. Ways of assessment:

1. What tools (e.g., tests) do you use to assess your students' mathematics achievements?

Tests, quizzes, my observing them in small group work

2. Why do you check your students' mathematics achievements?

I check the students' achievements in order to see if the material was understood – to be certain I can advance to the next stage.

For report card grades.

IV. Is it usual in your school for the mathematics staff to work together? yes

If yes, give details:

1. How often does the staff meet?

once a week once a month
 once every two weeks a few times a year

2. Who participates in the meetings?

all the math teachers of the school
 all the math teachers of one or two grades

3. What are the subjects usually discussed?

Planning lessons, challenge problems, quizzes and tests

4. What did you do at the last meeting?

Planned lessons, planned a test, prepared worksheets, solved a challenge problem

V. Requests or expectations from the meetings about assessment.

End-of-Year Questionnaire to Teachers in Sharon (July, 1999)

Teacher DL – 5th grade

1. Give examples of two tasks that the students performed as class work, from the last week of their studies.

a) $5 + _ < 6$

Each child wrote 5 solutions

By turn, they reported their solutions – are there more?

$_ + _ < 6$

Find 10 solution pairs – what can you say about the numbers?

- b) An exercise from the textbook which uses the calculator

2. a) What tools do you use for student assessment?

Tests – these we write according to the following criteria:

Understanding the task, understanding the concepts,
mathematical thinking, calculations, communication, creativity

Worksheets

Investigation tasks from the Assessment Task Bank

Homework

- b) Why do you check student achievement?

To check if the material learned has been “caught” by the students, and how it was “caught”, and what wasn’t “caught”

I also want to see where each individual student is (even though a teacher knows that without the need of tests)

3. What was the contribution of our workshops regarding student assessment?

The change of the report card at our school.

I now write tests and worksheets which are completely different from the ones I used to write – now I base my tests and worksheets on the criteria: understanding the task, understanding the concepts, mathematical thinking, calculations, communication, creativity

Background Questionnaire to Teachers in Sharon – 1st meeting (27.1.97)

Teacher DP

I. Personal information.

A. Teaching Experience:

1. Number of years of teaching experience 18
2. Number of years of teaching mathematics 17
3. This year I teach mathematics in grades 4, 5

B. Education:

- X 1. Graduate of _____ teachers' college
X Teaching certificate
_____ B.Ed.
My major was geography, history
- _____ 2. Graduate of _____ university
My major was _____

II. Choose one class in which you teach mathematics and answer the following questions:

Which grade is it? 4th

1. What topics did you teach in mathematics in this class last week?

Vertical multiplication.

Conclusion of the subject of exponents.

2. Give examples of two tasks which the students were given last week.

a) Worksheets with vertical multiplication

b) Worksheets

III. Ways of assessment:

1. What tools (e.g., tests) do you use to assess your students' mathematics achievements?

A short test at the end of a topic, another test on several topics in the same chapter, classwork, challenge questions

2. Why do you check your students' mathematics achievements?

In order to know if they understood the material being taught and its objectives.

IV. Is it usual in your school for the mathematics staff to work together? yes

If yes, give details:

1. How often does the staff meet?

once a week once a month
 once every two weeks a few times a year

2. Who participates in the meetings?

all the math teachers of the school
 all the math teachers of one or two grades

3. What are the subjects usually discussed?

Preparing worksheets, tests, planning lessons

4. What did you do at the last meeting?

Planned a final test for quadrilaterals

V. Requests or expectations from the meetings about assessment.

To learn about other ways of assessing, in addition to tests.

Is there a need for comparison tests in a school?

Background Questionnaire to Teachers in Sharon – 1st meeting (27.1.97)

Teacher EA

I. Personal information.

A. Teaching Experience:

1. Number of years of teaching experience 21
2. Number of years of teaching mathematics 20
3. This year I teach mathematics in grades 5,6

B. Education:

- X 1. Graduate of Kibbutz Seminary teachers' college
X Teaching certificate
_____ B.Ed.
My major was Bible, history
- _____ 2. Graduate of _____ university
My major was _____

II. Choose one class in which you teach mathematics and answer the following questions:

Which grade is it? 5th

1. What topics did you teach in mathematics in this class last week?
Common denominators – the denominator of one fraction a factor of the denominator of the second fraction
2. Give examples of two tasks which the students were given last week.
 - a) Predictions about how we can solving the following exercises and why we solve them in that way:
 $\frac{1}{3} + \frac{1}{12} =$ $\frac{3}{8} + \frac{1}{4} =$
 - b) --

III. Ways of assessment:

1. What tools (e.g., tests) do you use to assess your students' mathematics achievements?

Individual work with a student, tests, work with a group of students

2. Why do you check your students' mathematics achievements?

How much the students understand the material, if I taught clearly enough, and if I can continue to the next topic

IV. Is it usual in your school for the mathematics staff to work together? yes

If yes, give details:

1. How often does the staff meet?

once a week once a month
 once every two weeks a few times a year

2. Who participates in the meetings?

all the math teachers of the school
 all the math teachers of one or two grades

3. What are the subjects usually discussed?

Planning lessons

4. What did you do at the last meeting?

Solved challenge problems, planned lessons

V. Requests or expectations from the meetings about assessment.

An additional way to assess students

Background Questionnaire to Teachers in Sharon – 1st meeting (27.1.97)

Teacher EC

I. Personal information.

A. Teaching Experience:

1. Number of years of teaching experience 25
2. Number of years of teaching mathematics 25
3. This year I teach mathematics in grades 4, 5

B. Education:

1. Graduate of _____ teachers' college
 Teaching certificate
_____ B.Ed.
My major was science, mathematics
- _____ 2. Graduate of _____ university
My major was _____

II. Choose one class in which you teach mathematics and answer the following questions:

Which grade is it? 6th

1. What topics did you teach in mathematics in this class last week?
Ratio, word problems on ratio
2. Give examples of two tasks which the students were given last week.
 - a) A worksheet, with word problems
 - b) Ratio, scale factors of personal objects

III. Ways of assessment:

1. What tools (e.g., tests) do you use to assess your students' mathematics achievements?

Short quizzes on the topic being studied

2. Why do you check your students' mathematics achievements?

Checking knowledge, understanding.

Certainly for making a profile one needs to know what to strengthen and to correct.

IV. Is it usual in your school for the mathematics staff to work together? no

If yes, give details:

1. How often does the staff meet?

_____ once a week _____ once a month
_____ once every two weeks _____ a few times a year

2. Who participates in the meetings?

_____ all the math teachers of the school
_____ all the math teachers of one or two grades

3. What are the subjects usually discussed?

4. What did you do at the last meeting?

V. Requests or expectations from the meetings about assessment.

Background Questionnaire to Teachers in Sharon – 1st meeting (27.1.97)

Teacher ED

I. Personal information.

A. Teaching Experience:

1. Number of years of teaching experience 1
2. Number of years of teaching mathematics 1
3. This year I teach mathematics in grades 4, 5

B. Education:

- X 1. Graduate of Levinsky teachers' college
X Teaching certificate
_____ B.Ed.
My major was English
- X 2. Graduate of Tel Aviv university
My major was humanities

II. Choose one class in which you teach mathematics and answer the following questions:

Which grade is it? 4th

1. What topics did you teach in mathematics in this class last week?
Exponents, mathematical thinking (a challenge problem)
2. Give examples of two tasks which the students were given last week.
 - a) Write exercises whose answers are from 0 to 20 using the digits from 1997. Each digit can be used exactly once, all arithmetic operations can be used.
 - b) Create a menu from the Burger Range Menu for a party where you invited 29 children and your budget is 350 shekels.

III. Ways of assessment:

1. What tools (e.g., tests) do you use to assess your students' mathematics achievements?

Individual reports, tests.

2. Why do you check your students' mathematics achievements?

To see if the students understood the material.

If the material was presented in a good, clear way.

IV. Is it usual in your school for the mathematics staff to work together? yes

If yes, give details:

1. How often does the staff meet?

once a week once a month

once every two weeks a few times a year

2. Who participates in the meetings?

all the math teachers of the school

all the math teachers of one or two grades

3. What are the subjects usually discussed?

Planning lessons

4. What did you do at the last meeting?

Planned the next unit for geometry

V. Requests or expectations from the meetings about assessment.

Additional ways for assessing students

End-of-Year Questionnaire to Teachers in Sharon (July, 1999)

Teacher ED – 5th grade

1. Give examples of two tasks that the students performed as class work, from the last week of their studies.

- a) Here is a right triangle – how can we find its area?
- b) Find all possible answers to make this correct:
 $0.39 < 0.3__$

2. a) What tools do you use for student assessment?

Mathematical journals, weekly summaries, quizzes, individual reports

- b) Why do you check student achievement?

To know where each student has difficulty

To know if the students understand the material being learned

To see who can make generalizations

To see at what level the class is

3. What was the contribution of our workshops regarding student assessment?

To understand the importance of the student's thought processes and not merely the final solutions.

To know that there are a number of criteria which are important in mathematical thinking

To differentiate between understanding concepts, generalizations, multiple solutions, and creativity

Background Questionnaire to Teachers in Sharon – 1st meeting (27.1.97)

Teacher EE

I. Personal information.

A. Teaching Experience:

1. Number of years of teaching experience 18
2. Number of years of teaching mathematics 16
3. This year I teach mathematics in grades 5

B. Education:

- X 1. Graduate of Levinsky teachers' college
X Teaching certificate
_____ B.Ed.
My major was Bible, history
- _____ 2. Graduate of _____ university
My major was _____

II. Choose one class in which you teach mathematics and answer the following questions:

Which grade is it? 5th

1. What topics did you teach in mathematics in this class last week?
The place of fractions on the number line.
2. Give examples of two tasks which the students were given last week.
 - a) Place the fraction $\frac{2}{5}$ on the number line
 - b) Place the fraction $\frac{9}{4}$ on the number line

III. Ways of assessment:

1. What tools (e.g., tests) do you use to assess your students' mathematics achievements?
Classwork and checking each student, tests
2. Why do you check your students' mathematics achievements?
In order to check if they understood the topics studied in class

IV. Is it usual in your school for the mathematics staff to work together? yes

If yes, give details:

1. How often does the staff meet?

once a week once a month
 once every two weeks a few times a year

2. Who participates in the meetings?

all the math teachers of the school
 all the math teachers of one or two grades

3. What are the subjects usually discussed?

Planning lessons, ways for assessment

4. What did you do at the last meeting?

Planned lessons

V. Requests or expectations from the meetings about assessment.

To get tools for assessment students

Appendix D: Paper for SEMT '01
International Symposium
Elementary Mathematics Teaching
Prague, August, 2001

Note: For copyright reasons the following paper in Appendix D has not been reproduced in full.

Albert, Jeanne *Working with elementary teachers on challenge problems in mathematics*

(Co-ordinator, ADT Project (Retrospective), Curtin University of Technology, 7.5.03)