

# National Center for Human Resources Development (NCHRD)

Comparative Math and Science Achievement Across the UNRWA Fields of Gaza, Lebanon, Syria, Jordan and the West Bank

Ву

Kapur Ahlawat (NCHRD) Moh'd Shahin (UNRWA)
Reem Damen(UNRWA) Abdallah Qatouni(UNRWA)
And Hisham Al-Dhaleh (NCHRD)

Publications Series No.





**NCHRD** 

Comparative Math and Science Achievement Across the UNRWA Fields of Gaza, Lebanon, Syria, Jordan and the West Bank

#### Authors

Kapur Ahlawat (NCHRD), Moh'd Shahin (UNRWA),

Reem Damen (UNRWA), Abdallah Qatouni (UNRWA),

and Hisham Al-Dhajeh (NCHRD)

The study was supported by a grant from the Government of Germany

#### **ACKNOWLEDGEMENTS**

We gratefully acknowledge the German Government's financial contribution for its keen interest in conducting this study.

We also extend our thanks to the staff of the National Center for Human Resource and Development (NCHRD) in Jordan for their efforts in helping (UNRWA) personnel to use International Assessment of Education progress second study (IAEP II) test and in conducting data analysis.

Special reference is made to the **General Education Division** staff members for their efficient contribution in the preparation and production of this report.

N.Bouchouchi
Director of Education
UNRWA

# **Table of Contents**

Table of Contents	I
List of Tables	Ш
List of Figures	VШ
Executive Summary	IX
Part One (The Context)	
Chapter I	·
Introduction	2
Introduction	$\frac{1}{2}$
- Questions of the Study	3
Chapter II	•
Methodology - Sampling Frame - Instruments - The Data	4 4 6 11
Part Two (International Comparisons)	
Chapter III	
International Comparisons in Mathematics Achievement	13 13 23
Chapter IV	
International Comparisons in Science Achievement	30 30 40

### Chapter V

Areas of Weakness and Strength	44
Areas of Weakness and Strenth in Mathematics	44
Areas of Wekness and Strength in Science	52
	•
Part Three	
(Contextual Factors of Academic Achievement)	
Chapter VI	
Home Background, Classroom Activities Student Attitudes and	
Achievement	61
The family Background, Classroom Activities, and Attitudes of the	
Math Sample Students	61
Home Background, Classroom Activities, Student Attitudes and	
Achievement in Science	73
Chapter VII	-
Major Dagommandations	06
Major Recommendations	86
References	88
	00

# List of Tables

Table (1):	Sampling Distribution of Schools and Students Across Student Gender and Test Subject	4
Table (2):	Sampling Distribution of Schools and Students Across Student Gender and Test Subject	5
Table (3):	Sampling Distribution of Schools and Students Across Student Gender and Test Subject	6
Table (4):	Table of Specifications for the 8th Grade Mathematics Test	. 7
Table (5):	Distribution of Items by Content Area (Math Test)	8
Table (6):	Distribution of Items by Cognitive Skill (Math Test)	8
Table (7):	Table of Specifications for the 8th Grade (Science Test)	8
Table (8):	Distribution of Items by Content Area (Science Test)	9
Table (9):	Distribution of Items by Cognitive Skill (Science Test)	9
Table (10):	Contents of the Family Background Questionnaire (Part V of the Math and Science Tests)	10
Table (11):	Average Percent Correct, Standard Error, and Statistical Significance Between Pairs of Gaza, Lebanon, and Syria as One Set and Between Jordan and West Bank as a Separate Set on the Total Math Test Score and its Five Content Areas	14
Table (12):	Average Percent Correct, Standard Error, and Statistical Significance Between Pairs of Gaza, Lebanon and Syria as One Set and Between Jordan and West Bank as a Separate Set on Math Test Skills	18
	Average Percent Correct, Standard Error, and Statistical Significance Between Pairs of Gaza, Lebanon and Syria as One Set and Between Jordan and West Bank as a Separate Set on Math Test Subscales	19
	Average Percent Correct (M), Standard Deviation (SD) Standard Error (SE) and Statistical Significance Between Male and Female Student Achievement in Math Test in Each Field.	24

Table (15):	Average Percent Correct, Standard Error and Statistical Significance Between Gaza, Lebanon, Syria, As One Set and Between Jordan and West Bank As a Separate Set on the Science Test and its Four Content Areas	32
Table (16):	Average Percent Correct, Standard Error and Statistical Significance Between Gaza, Lebanon, Syria, As One Set and Between Jordan and West Bank As a Separate Set on the Science Test and its Skill Subscales	35
Table (17):	Average Percent Correct, Standard Error and Statistical Significance Between Gaza, Lebanon, Syria, As One Set and Between Jordan and West Bank As a Separate Set on the Science Test and its Subscales	37
Table (18):	Average Percent Correct (M), Standard Deviation (SD) and Statistical significance Between Male and Female Student Achievement on Science Test in Each Field	40
Table (19)	Average Percent Correct (M), Standard Deviation (SD) and Standard Error (SE) of Mean and Statistical Significance of the Difference Between Male and Female Student Achievement in Total Sample of Gaza, Lebanon & Syria.	43
Table (20):	Number of Test Items Under Success Percentages	44
	Percentage of 8th Grade Students Passing Each Item in Part I of the Math Test in Each Field and the Entire Sample	46
	Percentage of 8th Grade Students Passing Each Item in Part II of the Math Test in Each Field and the Entire Sample	48
. ]	Percentage of 8th Grade Students Passing Each Item in Part III of the Math Test in Each Field and the Entire Sample	49
Гable (24): 1 1	Percentage of 8th Grade Students Passing Each Item in Part IV of the Math Test in Each Field and the Entire	51

Table (25):	Number of Test Items With Less than 30% Correct Answer in Each Field Classified by Content and Skill Category Measured by the Science Test	
Table (26):	Number of Test Items in the Range of 30%-50% Correct Answers in Each Field Classified by Content and Skill Category Measured by the Science Test	53
Table (27):	Percentage of 8th Grade Students Passing Each Item in Part I of the Science Test in Each Field and the Entire Sample	55
Table (28):	Percentage of 8th Grade Students Passing Each Item in Part II of the Science Test in Each Field and the Entire Sample	56
Table (29):	Percentage of 8th Grade Students Passing Each Item in Part III of the Science Test in Each Field and the Entire Sample	57
Table (30):	Percentage of 8th Grade Students Passing Each Item in Part IV of the Science Test in Each Field and the Entire Sample	58
Table (31):	Linear Correlation Coefficients of Home Background, Classroom Activities and Student Attitudes with Math Achievement Components in Gaza, Lebanon and Syria	63
	Linear Correlation Coefficients of Home Background, Classroom Activities and Student Attitudes with Science Achievement Components in Gaza, Lebanon and Syria	75
Table (M1):	Q1. What language is usually spoken in your home?	61
Table (M2):	Q2. How many brothers and sisters do you have?	62
Table (M3):	Q3. How many books are there at your home?	65

Çât L

	Q5. How often do you read your own for fun outside of school?	65
Table (M5)	Q6. How much time do you usually spend each day on homework for all school subjects?	66
Table (M6):	Q 4. and Q 5. Percentage of Response in Each Category	67
Table (M7):	Percentage of Yes/No Responses to Q7, Q8, Q9, Q23, and Q24	68
Table (M8):	Q 16. With which of the following statements about Mathematics do you agree?	.* 69
Table (M9):	Percentage of Responses in Each Category of the Items 10 - 15	70
Table (M10)	): Percentage of Responses in Each Category of Items 17 through 21	71
Table (S1):	Frequency, and Percentage of Responses in Each Language Category	73
Table (S2):	Frequency, and Percentage of Responses in Each Response Category	74
, .	Frequency, and Percentage of Responses in Each Response Category Q3: How many books are there in your home? (Do not count newspapers, magazines, or comics)	77
Table (S4):	Percentage of Responses in Each Category of the Items 10 - 15	78
Γable (S5):		78
Γable (S6):	Response Percentage in Each Category for Each Science Learning Activity	80
Table (S7):	Percentage of Students in Each Category of Watching TV at Home on a School Day and Spending Time Doing Homework Each Week?	80

Category. Q6: How much time do you usually spend each day on homework for all school subjects?			
Table (S9):	Frequency, and Percentage of Responses in Each Category. Q5: How often do you read on your own for fun outside of school?	84	
Table (S10):	Frequency, and Percentage of Responses in Each Category. Q13 With which of the following statements about science do you agree?	85	

# List of Figures

Figure (1):	Mean Percent Correct Scores of the Five Arabic and the International IAEP II Samples on the Math Test Content Components	17
Figure(2):	Mean Percent Correct Scores of the Five Arabic and the International IAEP II Samples on the Three Skill-Components of the Math Test	
Figure (3):	Mean Percent Correct Scores of the Five Arabic and the IAEP II Samples on Fifteen Subscales Defined by the Interaction of Five Content Areas and Three Cognitive Skills	20
Figure (4):	Average Percent Correct Scores of the 8th Grade Students of the Different Samples on the Science Test and its Content Subscales	34
Figure (5):	Average Percent Correct Scores of the 8th Grade Students of the Different Samples on the Cognitive Skill Subscales of the Science Test	35
Figure (6):	Average Percent Correct Scores of the 8th Grade Students of Different Samples on Content by Skill Subscales of the Science Test	38

#### Executive Summary

#### Introduction

Standardized tests of achievement are designed to assess overall achievement in one or more subject areas or skills at various points during the child's school experience. The term standardized means that these tests are designed to be administered, scored and interpreted under standard conditions. Because of this feature, they are useful for making comparisons between pupils and between groups even when the tests have been administered at different times and in different places. The rationale for this study emanated from the need to obtain nationally/internationally comparable data by standardized measure on the achievement of UNRWA students in Gaza, Lebanon and Syria. A study recently completed by the National Center of Educational Research and Development (NCERD) comparing the Math and Science achievement of 8th grade students in Jordan and the West Bank further spurred the interest of the UNRWA schools in Gaza, authorities to assess the performance of UNRWA Lebanon and Syria, as well. Availability of testing technology developed by the International Assessment of Educational Progress second study (IAEP II) adapted for and tested on Jordanian and West Bank student populations and willing cooperation of experienced regional researchers provided added impetus to replicate the study in the remaining three fields of the UNRWA schools.

#### **Objectives**

The overall aim of this study was to measure the achievement of 8th grade students in *UNRWA* schools in Mathematics and Science using *IAEP II* standardized instruments and to collect information about classroom activities, student attitudes toward Mathematics and Science and other home and family background factors that could possibly attribute variation in student achievement.

More specifically the objectives of the study were:

- a. To measure the levels of achievement of 8th grade students in UNRWA schools in Mathematics and Science using internationally standardized tests.
- b. To explore similarities and differences in the achievement of 8th grade students in Math and Science across the various fields of *UNRWA* operations.

- c. To identify home and community backgrounds, educational practices and students' school experiences and attitudes, and their relationship with student achievement in Math and Science.
- d. To assist the Palestinian Authority to set broad guidelines for preparing Math and Science curricula for Palestinian students.
- e. To detect the areas of strength and weakness in the achievement of 8th grade students in Math and Science as measured by *IAEP II* Tests in Gaza, Lebanon and Syria.

#### Instruments

The instruments of the study were the same Arabised achievement tests that were administered to the IAEP II study sample in Jordan in 1991. The same instruments were later used in 1992 by the National Center for Educational Research and Development (NCERD) in Jordan to conduct a comparative study of student achievement in Math and Science in Jordan and West Bank. The instruments included 76-item Math test and a 72-item Science test, each including respectively a set of 24 and 22 items tapping student home background and school/classroom practices and out of school activities.

The Math test included 5 topics namely: Number and Operations (NUM), Measurement (MEA), Geometry (GEO), Data Analysis and Probability (DAT), and Algebra (ALG). Each topic included questions to assess student performance on each of the three cognitive skills Conceptual Understanding (CU), Procedural Knowledge (PK) and Problems Solving (PS). Thus, in addition to the total Math test score, there are 23 other subscale sores, 5 for the 5 topics, 3 for the cognitive skills and 15 for the 15 combinations of topics and skills.

The Science test was designed to measure students' Knowledge of facts, concepts and principles (KNO), Application of knowledge to solve simple problems (APP) and Integration of knowledge and principles to solve more complex problems (INT) in 4 major topics of Science namely: Life Sciences (LIF), Physical Sciences (PHY), Earth and Space Sciences (ESP), and Nature of Science (NAT). The Science test contained 18 distinct components covering the knowledge, application and integration skills in the four above mentioned disciplines of Science.

Both tests were administered in Gaza, Lebanon and Syria by a team of *UNRWA* personnel who were specifically trained and rehearsed on *IAEP II* standardized testing procedures at the National Center for Educational Research and Development (*NCERD*) in Jordan.

#### Findings

The findings of the study are described under three major classes, the international comparisons, item analysis for each field and the relationship of home background, school/classroom practices and students' attitudes with their achievement scores on each of the Math and Science tests.

#### International Comparisons in Math Achievement

On the Mathematics Test, comparisons were made between Gaza, Lebanon, Syria, and International averages on total test scores and on 23 components of the Math Test. While with reference to *IAEP II* norms of achievement in Mathematics students in each of the three fields scored very low, Gaza students' achievement was consistently lower than that of their counterparts in Lebanon and Syria at the level of statistical significance  $\alpha = 0.05$ .

In Syria, eighth grade students scored statistically significantly higher than their counterparts in both Gaza and Lebanon on 10 out of 24 components including the total Math test score. Furthermore, students in Lebanon scored statistically significantly higher than their counterparts in Gaza on 7 out of 24 components including the total test score.

When the mean scores of the five fields were compared (taking into consideration the achievement of *UNRWA* students in both West Bank and Jordan from study conducted by *NCERD* in 1992), with other countries (placed on the International scale), only SAR field students total score could be placed before the total test score of Jordan's national sample (representing 13-year- old population as used in *IAEP II* study in 1991). The mean score of the entire *UNRWA* sample when treated as one country places it as third from the bottom among the 20 *IAEP II* study countries, above Mozambique and Brazil.

#### Gender-Wise Comparisons in Math Achievement

In Gaza field, on one component only, namely Geometry Procedural Knowledge, female students scored statistically significantly higher than the male students. On the remaining 23 components, differences were not statistically significant.

In Lebanon, on 17 out of 24 components, male students scored significantly higher than the females.

In Syria, male and female differences were significant on 16 out of the 24 test components in favour of the male students. Female students did not outscore the males on any of the 24 variables.

#### Science Achievement

UNRWA eighth grade students in Syria scored statistically significantly higher than their counterparts in Lebanon on 16 out of 18 components of the Science test including the total test score. They also scored significantly higher than their cohorts in Gaza on all the 18 test components.

UNRWA eighth grade students in Lebanon scored statistically significantly higher than their counterparts in Gaza on the total test score and on 7 of the test components.

The average percent correct score (Mean) for each of the five fields (Jordan and West Bank means as reported by Ahlawat et al., 1992), on total Science Test score were 50.7, 55.3, 61.7, 54.3 and 48.5 for Gaza, Lebanon, Syria, Jordan and West Bank respectively. Upon comparing the total Science Test mean score of each field with the international scale of the 20 IAEP II participating countries, the total science mean score of 8th grade students in Syria can be placed on top of Jordan.

The mean score of the entire *UNRWA* sample when treated as a whole entity, came up to 54.4 which places *UNRWA* third from the bottom (i.e., just above Sao Paulo Brazil and Fortaleza Brazil).

#### Gender - Wise Comparisons in Science Achievement

In Gaza field, on 5 Science Test components out of the 18, the mean score of male students was statistically significantly higher than that of the female students at the  $\alpha=0.05$  level. Moreover, in terms of absolute differences between the means of the two groups, on 7 out of the 18 variables, male students' scores were slightly higher than those of the females. On two subscales the mean scores of male and female students were the same, namely Life Sciences Knowledge (LIF KNO) and Integration (INT), while on two other subscales Nature of Science topic (NAT) and Nature of Science Integration (NAT INT) mean scores of males and females were nearly the same. On the remaining two subscales namely Life Application (LIF APP) and Life Integration (LIF INT) female students' mean scores were slightly higher but only on one of them (LIF INT) the difference was statistically significant at the 0.05 Alfa level.

In Lebanon Field, on 9 out of the 18 variables including the total test score, differences were statistically significant in favor of male students at the  $\alpha = 0.05$  level. As was the case in Gaza field, performance of female students in Life Sciences Integration was statistically significantly higher than that of their male cohorts at the 0.05 Alpha level.

In Syria, on 9 out of 18 variables including the total Science Test score, differences were statistically significant in favor of male students at  $\alpha = 0.05$ . Female students in Syria did not perform significantly better than the male students on any of the 18 variables.

#### Areas of Relative Weakness and Strength

#### **Mathematics**

In terms of different content areas of Mathematics curriculum, in Syria and Lebanon the topic of Geometry ranks first followed by Numbers, Statistics, Algebra, and Measurement which is the weakest of all.

In Gaza, the topic of Numbers comes first followed by Geometry, Statistics, Algebra and finally Measurement. Apparently, the topic of Measurement is the weakest of all in all the three fields.

In terms of cognitive skills in Mathematics, in both Syria and Lebanon, Conceptual Understanding ranks first followed by Procedural Knowledge, and Problem Solving comes last.

Upon comparing the 15 combinations of the 5 Math topics and 3 cognitive skills measured in each of Gaza, Lebanon and Syria, it was quite obvious that students in Gaza were the weakest of all in Measurement Conceptual Understanding (MEA CU), Measurement Procedural Knowledge (MEA PK), Measurement Problem Solving (MEA PS), Geometry Problem Solving (GEO PS), and Data Analysis Procedural Knowledge (DAT PK).

In Lebanon performance of students was the weakest in Measurement Procedural Knowledge (MEA PK), Measurement Problem Solving (MEA PS) and Data Analysis Procedural Knowledge (DAT PK).

As for Syria performance of students was the weakest in Measurement Procedural Knowledge (MEA PK) and Measurement Problem Solving (MEA PS).

#### Science

In Syria, among the four content areas, the topic of Earth and Space Sciences ranks the best followed by Life Sciences, Physical Sciences and Nature of Science, the latter two disciplines being the weakest.

engelegge (1 des 1971) i de 1987 (1 de 1988) i de 1988 (1 de 1988) i de 1988 (1 de 1988) i de 1988 (1 de 1988) Handard (1 de 1988) i de 1988 (1 de 1988) i Handard (1 de 1988) i de 1988 (1 de 1988) i

In Lebanon, Life Sciences ranks first followed by Earth and Space Sciences, Nature of Science and Physical Sciences which turns out to be the weakest area.

In Gaza, Earth and Space Sciences and Life Sciences stand at an equal footing and rank first, next comes Physical Sciences and finally Nature of Science which turns out to be the weakest area.

In terms of cognitive skills in Science, Knowledge takes the first rank in all the three fields (Gaza, Lebanon, and Syria), then comes Application followed by Integration which takes the last place.

Upon comparing the performance of 8th grade students in Gaza, Lebanon and Syria on the 10 combinations of the four topics and the three skills measured (Earth and Space Sciences and Nature of Science were measured on two skills each), students in Gaza were quite weak in Life Sciences Integration (LIF INT) Physical Sciences Application and Integration (PHY APP & PHY INT) followed by Life Sciences Application (LIF APP), Physical Sciences Knowledge (PHY KNO) and Nature of Science Application and Integration (NAT APP & NAT INT).

In Lebanon weak performance was evident in Life Science Integration (LIF INT) and Physical Sciences Knowledge and Integration Skills (PHY KNO & PHY INT) and Nature of Science Integration (NAT INT).

In Syria weak performance was obvious in Physical Sciences Knowledge (PHY KNO) as well as in Integration Skill in both Life Sciences and Nature of Science (LIF INT & NAT INT).

### Home Background and Student Achievement

The two proxy measures of socioeconomic status, (1) the number of brothers and sisters and (2) the number of books in the home were not correlated statistically significantly with achievement in Science, yet the number of books in the home was statistically significantly correlated with Math achievement scores of the students.

Watching TV at home on school days: Amount of watching TV at home on school days was found to have a significant positive linear relationship with students achievement in both Math and Science.

Reading for fun: Outside school, had positive correlation with achievement in Math and Science. Statistical significance, however, was obtained only in case of Math achievement.

Parental aspirations: (My parents want me to do well in Mathematics and parental interest in science (My parents are interested in science), had significant positive relationship with students' performance in Mathematics but no relationship with their achievement in Science.

Parental interest (Does any one at home ever talk to you about what you are learning in (Math/Science) class had no significant relationship with achievement in either Math or Science.

Weekly time spent on home work in the subject had a positive relationship with achievement in Math (at least in two components) and a negative but nonsignificant relationship with achievement in Science.

Daily time spent on homework in all subjects had significant positive relationship with achievement in both subjects, Mathematics and Science.

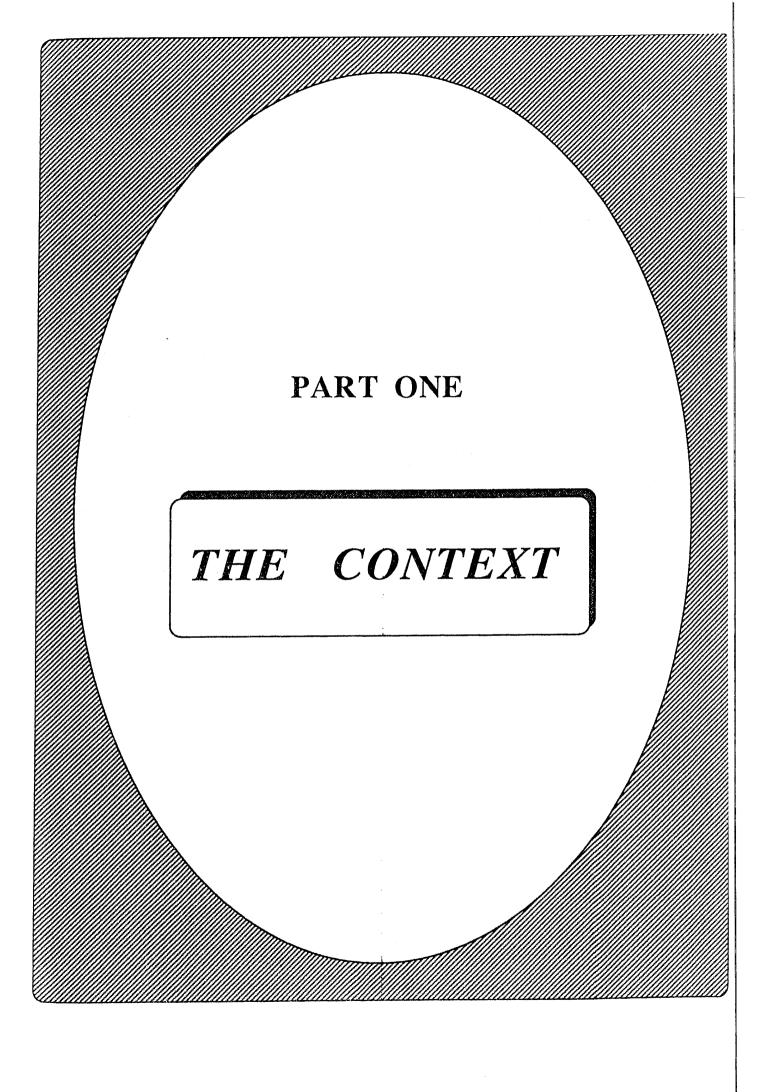
Amount of help in the homework at home had a negative relationship with achievement in both Math and Science. The correlation, however, was statistically nonsignificant, except with three of the Science Test subscales.

Beliefs and Attitudes of students toward Science and Math had a significant positive correlation with achievement in both Math and Science.

Some classroom activities like listening to the teacher give Math lesson and doing Math exercises by oneself had no significant relationship with achievement in Math. But some other activities like problem solving in small groups and working with Math and Geometrical objects had statistically significant negative correlation with several subscale scores of the Math Test. Science classroom activities like doing science experiments, watching teacher do science experiments and watching science programs on TV or Video in school had no significant relationship with achievement in science which is contrary to expectations.

The amount tests and quizzes had negative but nonsignificant correlation with achievement in both Math and Science.

and the second of the second o 





#### Chapter 1

#### **INTRODUCTION**

#### Background

International comparisons of educational systems and pupils' achievement in different subjects started in 1960s, two studies were conducted by the International Association for the Evaluation of Educational Achievement (IEA), one was carried out in 1964 and the other in early 1980. The International Assessment of Educational Progress (IAEP) also conducted two International studies, one in 1988 where only six countries participated in the study and the second (IAEP II) was carried out in 1991, wherein 20 countries namely; Korea, Taiwan, Switzerland, Hungary, Soviet Union, Slovenia, Italy, Israel, Canada, France, England, Scotland, Spain, United States, China, Ireland, Portugal, Jordan, Sao Paulo Brazil and Fortaleza Brazil participated. Jordan was the only Arab country which participated in the IAEP II study in two subjects (Math and Science) for age 13 (grade 8) students where UNRWA was represented by 17 schools in the study sample. Results of the study were disappointing to educators in Jordan. The National Center for Educational Research and Development (NCERD) in Jordan repeated the study in Jordan and West Bank in April 1992. The sample contained 16 and 12 schools from UNRWA schools of Jordan and West Bank, respectively. Results of the study indicated that performance on both Math and Science tests in Jordan and West Bank was extremely low. In May 1993, UNRWA in coordination with NCERD conducted the same study in UNRWA Schools in Syria, Lebanon and Gaza Fields.

Each of the countries which participated in *IAEP II* study did so for its own reasons. Some wanted to compare their results with those of neighbouring countries or competitors. Others wanted to learn about the educational systems and practices of countries whose students seem to achieve high in Mathematics and Science. *UNRWA* embarked on this study to evaluate the effectiveness of its educational system using internal and external information gathered from international comparisons of educational systems and students' levels of achievement in the various fields.

#### Rationale

UNRWA provides education for more than 390,000 Palestinian pupils and devotes a significant portion of its financial resources (more than 60% of its annual budget) for this purpose. The main information available to UNRWA to judge how effectively these resources are utilized is teachermade tests (which may lack validity and reliability) and in some Fields the national tests. Results of UNRWA schools in Lebanon Field in the Brevet National Examination were shocking to educationists with 16% pass in

1991, however, results of Brevet Trial Test of 1993 reached 44% level of success. In SAR Field, the pass percentage of *UNRWA* schools in the national examination at the end of the Preparatory Cycle was higher than the pass percentage of the Public Schools.

In Gaza and West Bank Fields, there was a prolonged closure of schools during the last four years. Eventually, academic performance of students and their attitudes towards education have declined to a great extent. The rationale for this study, in the light of the foregoing factors, stemmed from UNRWA's need to conduct a scientific evaluation of its educational system with special emphasis on pupils' achievement. The UNRWA students' performance in Math and Science was assessed in comparison with that of students from the countries which participated in the IAEP II study. Among other things the study aimed at providing good comparative data on educational achievement in UNRWA's fields of operation. Furthermore, within the context of the current peace process, Palestinians will need valid and reliable information in order to set broad guidelines for preparing Math and Science curricula for the Palestinian students.

#### Questions of the Study

In view of the foregoing perspective, this study aimed at investigating the following questions:

- 1) What are the levels of achievement of 8th grade students in *UNRWA* schools in Math and Science as measured by the *IAEP II* standardized tests?
- What similarities and differences are there in the achievement of 8th grade students in Math and Science across the five fields of *UNRWA* operations and in other countries?
- 3) Are there differences in achievement in Math and Science at the 8th grade level between male and female students?
- What do 8th grade students think about teaching Math and Science and what types of home backgrounds, educational practices, students' school experiences are associated with achievement in Math and Science?
- What are the relative areas of weakness and strength in the achievement of 8th grade students in Math and Science as measured by IAEP II tests?
- What are the indicators which may be derived from the results of the study which may assist in proposing a unified curricula in Math and Science for the Palestinian students?

#### Chapter II

#### **METHODOLOGY**

#### Sampling Design

#### Sampling Frame

As the main objective of this study revolved around comparing the levels of achievement of 8th grade students in *UNRWA* schools in Math and Science in each of Gaza, Lebanon and SAR fields, the following variables were used to define population strata and sampling frame.

- Field: Gaza, Lebanon, and Syria.
- Gender of the school: Male and Female.

#### Gaza Sample

The number of schools, randomly selected from each stratum, and the number of students who participated in both the Math and Science *IAEP* tests are given in Table 1.

Table (1)
Sampling Distribution of Schools and Students Across
Student Gender and Test Subject

Subject	t Schools		Subject Schools Students		
	Male	Female	Male	Female	
Math	4	4	88	87	
Science	4	4	87	88	

Students in each class were randomly divided into two halves. One half was randomly assigned to Science Test and the other to Math Test. In all, 175 students took the Math Test and 175 students took the Science Test.

#### Lebanon Sample

The sampling frame for Lebanon was determined following the same criteria as set for Gaza to define elements of the school population of interest. However, the sampling frame in Lebanon covered 5 educational areas namely, North, Central Lebanon, Baqa'a, Sidon and Tyre. 8 schools were randomly selected on the basis of school gender, 4 Male and 4 Female schools. Table 2 describes the distribution of schools and students across gender and Test subject.

Table (2)
Sampling Distribution of Schools and Students Across
Student Gender and Test Subject

Subject	ubject Schools Students		ents	
	Male	Female	Male	Female
Math	4	4	63	68
Science	4	4.	62	66

In all 131 students were randomly assigned to the Math Test and 128 students were randomly assigned to the Science Test.

#### Syria Sample

The sampling for Syrian schools was determined following the same criteria as set earlier for Gaza and Lebanon. However, it covered 3 areas namely North Area, Damascus Area and South Area. 12 schools were randomly; selected on the basis of school gender, 6 male and 6 female schools.

Table 3 describes the distribution of schools and students across gender and test subject.

Table (3)
Sampling Distribution of Schools and Students Across
Student Gender and Test Subject

Subject	Schools				lents
	Male	Female	Male	Female	
Math	6	6	111	114	
Science	6	6	114	112	

In all, 255 eighth grade students were randomly assigned to the Math Test and 226 eighth grade students were randomly assigned to the Science Test.

#### **Instruments**

The instruments and procedures used in this study were mostly the same as used in the second International Assessment of Educational Progress (IAEP II) study in which 20 countries participated from all over the globe. The development and standardization procedures of those instruments are systematically documented in the Technical Report of the IAEP II study (Adam chu, et al., 1992). In general, the data collection instruments included achievement tests in science and mathematics, a family background questionnaire including classroom experiences and student attitude toward the subject appended to the achievement tests, and a school questionnaire about general school conditions and instructional practices commonly used in schools. The composition and format of both achievement tests is described in the following paragraphs.

#### Math Achievement Test

The IAEP II Mathematics Test was assembled with common consensus of expert teams of the participating countries such that it focused on the common elements of their curricula with emphasis to a varying degree. Three broad areas of cognitive processes and five areas of mathematical contents (topics) consensually agreed upon to be essential elements of mathematics curriculum by all participating countries constituted the test components as detailed in Table 4.

Table (4)

Table of Specifications for the 8th Grade Mathematics Test

		Skills			,
Content	Conceptual Understanding (CU)	Procedural Knowledge (PK)	Problem Solving (PS)	Tot	al %
Numbers and Operations (NUM)	12	9	7	28	36.3 (35)
Measurement (MEA)	3	3	7	13	17.1 (15)
Geometry (GEO)	5	3	3	11	14.5 (20)
Data Analysis Statistics and Probability (DAT)	3	5	1	9	11.8 (15)
Algebra (ALG)	2	7	6	15	19.7 (15)
Total %	25 32.9	27 35.5	24 31.5	76 100	100

Note: In the last column numbers given in parentheses indicate relative emphasis given to different contents in the Math curriculum at that time in Jordan.

Table 4 gives the distribution of the math test items over content and skills and, finally, among fifteen cells, each defined jointly by the interaction of each content and skill category. The bottom row of Table 4 shows that 76 test items are almost evenly distributed among the three skill categories. The imbalance of item distribution among the five content areas, however, is apparent from the last column of the table.

The weight given to different areas of the subject matter varies from 12% (Data Analysis, Statistics and Probability) to 37% (Numbers and Operations). This disproprtionality reflects the differential emphasis that commonly exists in all curricula. Since curricular emphasis, generally, varies, more or less from country to country or from one school system to another, IAEP II test developers had obtained the relative curricular emphasis from all the participating countries. Incidentally, the relative emphasis of final version of the 8th grade (Age -13) Math Test as shown in Table 4 closely approximated the relative emphasis in Jordanian 8th Grade Math curriculum. Jordanian relative emphasis is given in parentheses in the last column of Table 4.

The number of questions included in the Math achievement test to measure students' performance in each content area and that number as a percentage of the whole test (76 questions) describe the structural base for the assessment of Math achievement. The distribution of items over the five content areas is followed by the distribution of items over the three skill levels, respectively, in Table 5 and Table 6.

Table (5)

Distribution of Items by Content Area (Math Test)

Content	Items	<b>%</b>
Numbers and Operations	28	37
Measurement	13	17
Geometry	11	15
Data Analysis, Statistics and Probability	9	12
Algebra	15	20
Total	76	100

Table (6)

Distribution of Items by Cognitive Skill (Math Test)

Content		Items		%		
Conceptual Understanding			25		33	
Procedural Knowledge		Ĭ.	27		35	• . :
Problem Solving	• •		24		32 :	
Total			76		100	· ·

#### Science Achievement Test

The distribution of Science test items by skill and topic is given in Table 7.

Table (7)

Table of Specifications for the 8th Grade Science Test

·		Skills			
Content	Knowledge (KNO)	Application (APP)	Integration (INT)	Tot	al %
Life Sciences (LIF)	16	6	3	25	35
Physical Sciences (PHY)	. 4	14	8	26	36
Earth and Space Sciences (ESP)	-1	8	-	9	13
Nature of Science (NAT)	-	4	8	12	17
Total %	21 29.2	32 44.4	19 26.4	72 100	100

The last row of Table 7 shows that relative emphasis given to different skills ranges from 26% (the least which is assigned to integration of knowledge for solving rather complex problems to 44% (the most which is given to application of knowledge to solve simple problems). The last column of Table 7 indicates varying degree of emphasis attached for

different content areas. Earth and Space Sciences with 12.5% of the items in the test got the least importance while Physical Sciences with 36% items got the highest importance. The four content areas seem to fall in two groups. Life Sciences and Physical Sciences form one group which is given substantially higher weight (35% and 46% respectively) than the group formed by Earth and Space Sciences, and the Nature of Science, the two areas which are given 12.5% and 16.7% weight, respectively.

The content areas of 8th grade Science include several interrelated disciplines of Science. All disciplines, however, do not get equal importance in terms of curricular and instructional emphasis. The following Table 8 describes the relative emphasis given to each content area in the Science Achievement Test.

Table (8)

Distribution of Items by Content Area (Science Test)

Content	Items	%
Life Sciences	25	34.7
Physical Sciences	26	36.1
Earth and Space Sciences	9	12.5
Nature of Science	12	16.7
Total	72	100

The objectives of teaching Science could be described in terms of cognitive processes required to deal with science content at different levels of complexity. The Science Achievement Test includes questions that assess students' performance in Science at three levels of cognitive skills. The following Table 9 shows relative distribution of different cognitive skills in the assessment of science achievement.

Table (9)
Distribution of Items by Cognitive Skill (Science Test)

Content	Items	%
Knows Facts, Concepts and Principles	21	29.2
Uses Knowledge to Solve Simple	- 32	-44.4
Problems Integrates Knowledge to Solve More Complex Problems	19	26.4
Total -	72	100

#### General Structure of Tests

Both Science and Math achievement tests were composed of five parts. The first four parts contained sets of achievement items while the fifth part consisted of 22 or 24 questions about students' family background, home environment, classroom activities and attitudes and beliefs as they were expected to be associated with student achievement in Math and Science. The Science Test contained a set of 22 such questions, whereas, the Math Test contained 24 questions. The two extra questions in Math activities were concerned with use of calculators and computers in solving math problems.

All the 72 items of the Science Test were Multiple-Choice, but in the Math Test 17 out of 76 items were Constructed-Response, i.e., students supplied the answer, while the remaining 59 items were Multiple-Choice type. Everyone of the four achievement parts in both tests was equally and individually timed. The contents included in the student questionnaire constituting Part V of the Math and Science tests are indicated in Table 10.

Table (10)

Contents of the Family Background Questionnaire
(Part V of the Math and Science Tests)

Content	Number of Items		
	Science	Math	
Language spoken at home	1	1	
No. of brothers and sisters	1	1	
No. of books	1	1	
Watching TV on school days	1	1	
Reading for fun	1	1	
Homework	. 2	2	
Beliefs and attitude toward the subject	7	7	
Subject related activities at school	4	4	
Tests and quizzes	. 1	1	
Family interest and support for the subject	. 3	3	
Calculator and Computer	_	2	
Total	22	24	

#### The Data

#### Data Collection

IAEP II study had developed standardized instruments as well as standardized test administration and data collection procedures. A team of UNRWA staff members from Gaza, Lebanon, Syria and UHB(A), was trained and rehearsed at the National Center for Educational Research and Development (NCERD) in the IAEP standardized procedures of test administration and data collection in May 3-7, 1993. In order to maintain the consistency of test administration and data collection methods, every precaution was taken to follow the standard test administration procedures as prescribed by the IAEP II study, step by step, with full fidelity. Pretest experience and adequate practice training was given to all participating classes so that lack of familiarity with multiple choice items and timed tests may not become sources of error in test scores.

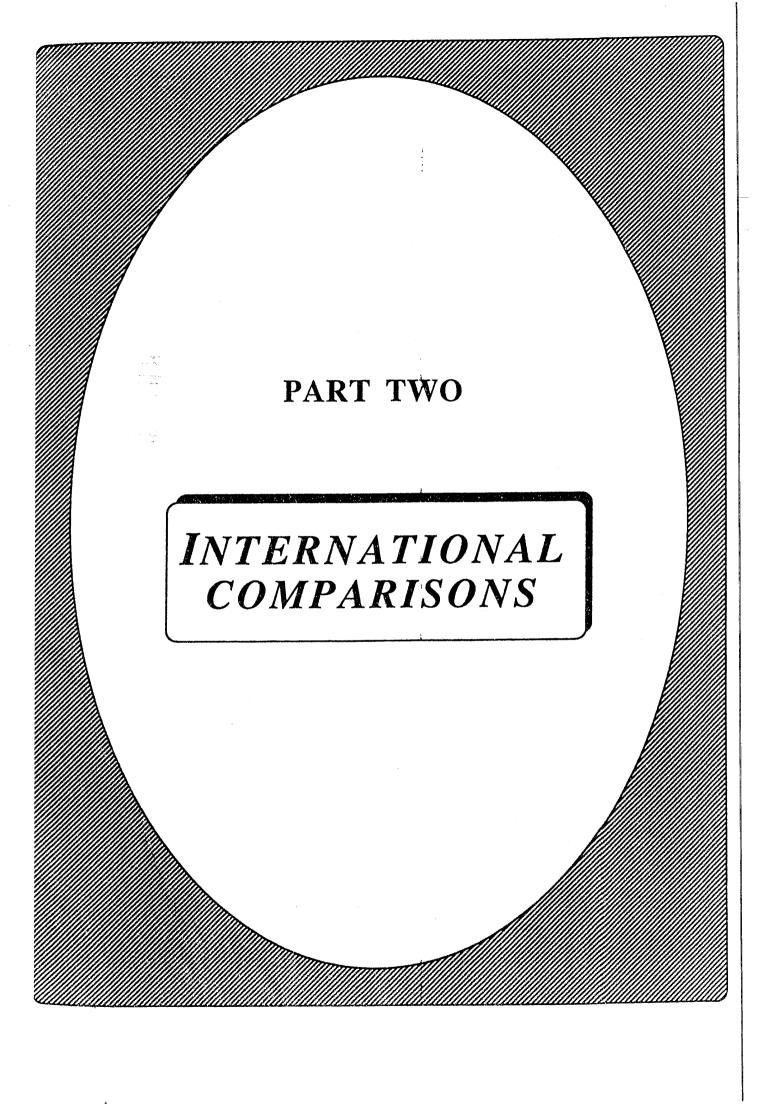
In all three fields (Gaza, Lebanon, Syria), data were collected during May 1993, scoring and data entry were conducted through *IAEP II* procedures at *UNRWA* Headquarters - Amman (UHB - A) in June 1993.

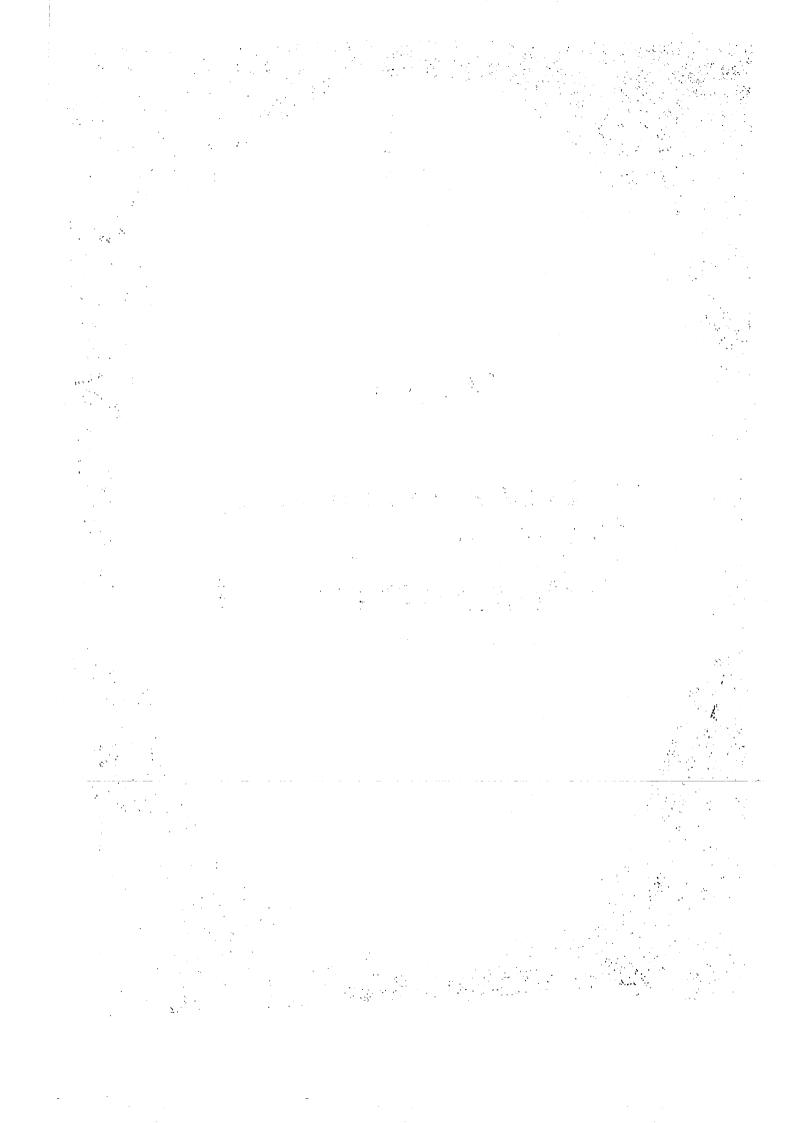
The Math Test was administered to a sample of 175 eighth grade students in Gaza, 131 eighth grade students in Lebanon and 225 eighth grade students in Syria following the standardized testing procedures of the IAEP II study by a team of UNRWA staff members from Gaza, Lebanon, Syria and UHB (A). All the team members had been well trained and rehearsed in IAEP II testing methods. The Science Test was administered to a sample of 175 eighth grade students in Gaza, 128 in Lebanon and 226 in Syria.

## Statistical Analyses of Data

Since Mathematics and Science tests were administered to different students, Science and Mathematics data were treated independently. In conformity with the *IAEP II* setting, randomly assigned half of the students in each intact class were administered the Science test and the other half were administered the Math Test. Both data sets were separately analyzed and subjected to a variety of statistical analyses.

The study questions required the statistical analyses and parameter estimation of dependent variables, i.e., the achievement tests and student questionnaire components or subscales to be treated as individual variables defined by specific sets of items. The test items were treated, generally, at three levels: individual item level, subscale score level, and total test score level.





### Chapter III

# INTERNATIONAL COMPARISONS IN MATHEMATICS ACHIEVEMENT

## Scale of comparison

In this part we will present the performance levels of eighth grade students on the *IAEP II* Mathematics Achievement Test in Gaza, Lebanon, and Syria (Data for Jordan and West Bank was added from an earlier report by Ahlawat et al., 1992) along with the international norms on each component of the Mathematics test in a comparable framework.

IAEP II Mathematics test consisted of 76 items of which 59 were of the Multiple-Choice type. They were designed to measure students' Conceptual Understanding, Procedural Knowledge, and Problem Solving skills in five major topics: Numbers, Measurement, Geometry, Statistical Data, and Algebra. Including total scores on all the 76 items, the mathematics test contained 23 distinct components.

Identifying the area of weakness and strength in student achievement in Mathematics and Science with a view to provide formative feedback to improve curricula, textbooks and instructional practices and to suggest broad guidelines to develop modern curricula and textbooks for the evolving new Palestinian system of education were among the primary objectives of the study. In order to look into these matters one needs to compare the performance of the same group of students on subtests of varying length measuring different contents and skills in order to be able to make some judgement about relative weakness and strength of the students in different areas of curriculum. Although to achieve perfect equivalence of measures of different concepts has been fraught with formidable complexities, to obtain a common unit of measurement is the absolute requirement.

In this respect the item correct scores on the achievement tests and on all the subtests within each test were transformed into percentage correct scores. So it must be clear from the outset that all the statistical computations were based upon the percentage correct scores on the tests and subtests and consequently all the parameters estimated, and comparative inferences made in this report refer to percentage correct score as the basic scale of comparison.

Earlier in 1992, the same *IAEP II* Math Test was administered to a sample of 319 eighth grade students in Jordan and 183 in the West Bank. The null hypothesis of no significant differences in the achievement of students in Gaza, Lebanon, and Syria was tested by using One-way analysis of variance at  $\alpha = 0.05$  level of significance for each component of the Math test. The results of the analyses are presented in Tables 11, 12, and 13.

The average percent correct score (mean) and its standard error were adopted as the population parameters of performance level on each test component. Table 11 includes the average percent correct score, standard error, and statistical significance on total math test and the five content areas in math namely Numbers and Operations (NUM), Measurement (MEA), Geometry (GEO), Algebra (ALG), and Data Analysis Statistics and Probability (DAT).

#### Table (11)

Average Percent Correct, Standard Error, and Statistical Significance Between Pairs of Gaza, Lebanon, and Syria as One Set and Between Jordan and West Bank as a Separate Set on the Total Math Test Score and its Five Content Areas.

				verage Per	cent Cor	rect	
Variable	Statistic	Gaza N=175	Lebanon N=131	Syria N=225	Jordan N=319	W.Bank N=183	INT.
Total Math	Sig Mean S.Error	36.35 1.13	40.10 1.25	G <l<s *<br="">43.53 1.01</l<s>	38.30 0.9	*** 29.00 00.80	58.00
Num (2)	Sig Mean S.Error	34.61 1.07	37.6 1.20	G,L <s* 39.47 1.01</s* 	42.1 00.88	*** 32.60 00.89	63.20
MEA (13	Sig Mean S.Error	25.71 1.05	26.25 1.26	G <l,s* 29.20 1.16</l,s* 	29.00 00.86	*** 21.40 00.84	48.00
GEO (11	Sig Mean S.Error	38.18 1.44	49.45 1.26	G <l,s* 52.53 1.54</l,s* 	40.80 1.31	*** 29.90 1.39	63.70
ALG (15	Sig Mean S.Error	32.99 1.49	38.07 1.63	G <l,s* 39.88 1.42</l,s* 	35.20 1.15	*** 24.40 1.10	55.50
DAT (09	Sig Mean ) S.Error	35.75 1.45	41.45 1.96	G<\$ <l* 47.80 1.31</l* 	42.30 00.99	*** 35.80 1.38	71.30

#### - Notes:

- 1. The number of test items comprising a subscale is given between brackets.
- 2. Mean is taken over the percent correct score on each subscale.
- 3. Statistical significance among Gaza, Lebanon, and Syria was tested via Newman-Keul Multiple comparisons at the 0.05 level, while the significance of t-test between Jordan and West Bank is taken from Ahlawat et al. (1992).

and the first of the second of

The state of the state of the state of

- \* Means statistically significant at  $\alpha < 0.05$ .

  \*\* Means statistically significant at  $\alpha < 0.01$ .

  \*\*\* Means statistically significant at  $\alpha < 0.001$ .

Table 11 also includes the International Norm (Average percent correct score) on each of the above mentioned components. The international norm was computed by the IAEP II study on the basis of aggregated national sample of 20 countries which participated in the study conducted by the Educational Testing Services (ETS) in collaboration with the national team of each country.

The name of Math Variable measured and the number of test items measuring it are given in the first column entitled "Variable".

The level of statistical significance for each of the One-Way Analysis of Variance between the three fields (Gaza, Lebanon, and Syria) was tested at  $\alpha = 0.05$  level.

The bar over G and L in the Symbol G,L< S is used to indicate that there is no significant difference between Gaza and Lebanon but Syria is significantly higher than each of the other two fields.

The statistics about Jordan, West Bank, and IAEP II samples are taken from an earlier study conducted by Dr. Ahlawat et al., 1992.

The overall Math achievement is represented by the average percent correct total Math score which is given in the first row of table 11. The mean percent correct score of eighth grade students in Syria is 43.53 which is higher by 7,3,5, and 14 percentage points than the means of Gaza, Lebanon, Jordan and West Bank, respectively.

In comparison to the International Mean, the mean of students in Syria is almost 15 points below the norm, while Lebanon's is 18 points, Gaza's 22 points, West Bank's 29 points, and Jordan's 20 points below the international mean.

The highest achieving country among the 20 IAEP II participants was China with its average 80 percent correct, and the lowest achieving was Mozambique with 28 percent correct average.

The mean score of the entire *UNRWA* sample when treated as one group was 38 which could be placed as the third from bottom before Mozambique and Brazil.

Table IT shows that relative performance on the Measurement component was the worst among all the topics, across all the UNRWA fields as well as in the *IAEP II* international study. This can be attributed to the fact that the prescribed Mathematics curriculum does not emphasize this topic enough. The Measurement component was identified an area of weakness in Jordan and West Bank also. (Ahlawat et al., 1992).

A graphic display of the average percent correct scores of the five UNRWA fields along with that of the *IAEP II* study population on the above mentioned components is given in Figure 1.

Table 12 includes the average percent correct scores, Standard Errors, and significance on the three main skills measured namely; Conceptual Understanding (CU); Procedural Knowledge (PK); and Problem Solving (PS). Each topic included questions to assess the student performance in each of the three hierarchical cognitive skills: Knowledge, Application and Integration. In addition, the international norms on the three skills are also represented in table 12.

It is quite evident that the Procedural Knowledge skill is an area of major weakness. The highest average percent correct score (40.38) was obtained by Syria, and it is about 20 points less than the international norm (60.30). In regard with international comparison, student achievement on all the three cognitive skills is much below the international level, in all the five UNRWA fields included in this study.

Figure (1)

Mean Percent Correct Scores of the Five Arabic and the International IAEP II
Samples on the Math Test Content Components

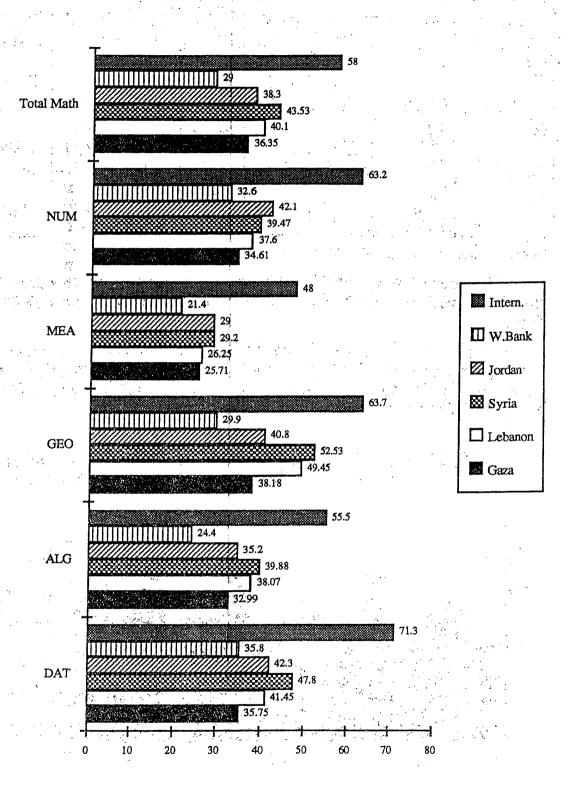


Table (12)

Average Percent Correct, Standard Error, and Statistical Significance
Between Pairs of Gaza, Lebanon and Syria as One Set and Between Jordan and
West Bank as a Separate Set on Math Test Skills

	<del></del>	1	]	Av	erage Per	cent Cori	rect	
Var	riable	Statistic	Gaza	Lebanon	Syria	Jordan	W.Bank	INT.
			N=175	N=131	N=225	N=319	N=183	AVG.
CU	(25)	Sig Mean S.Error	41.23 1.18	44.92 1.35	G <l<s *<br="">51.06 0.99</l<s>	43.90 0.85	*** 36.30 0.95	62.50
PK	(27)	Sig Mean S.Error	33.35 1.45	37.86 1.64	G <l,s* 40.38 1.26</l,s* 	35.70 1.11	*** 23.70 1.07	60.30
PS	(24)	Sig Mean S.Error	34.64 1.07	37.60 1.20	G <s* 39.24 1.01</s* 	35.40 0.88	*** 27.50 0.89	57.00

Note: (Jordan and West Bank are added from earlier study).

A graphical representation of the average percent correct scores on the three main skills along with the *IAEP II* study population mean is given in figure 2.

Table 13 includes the 15 combinations resulting from the 5 topics by the three main skills. Table 13 also includes the average percent correct scores in addition to the standard error, statistical significance, and the international norm for each of the 15 combinations.

Figure (2)

Mean Percent Correct Scores of the Five Arabic and the International IAEP II Samples on the Three Skill-Components of the Math Test

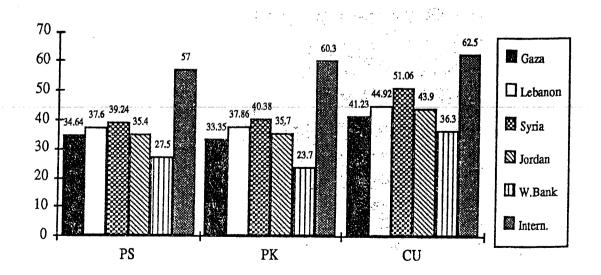


Table (13)
Average Percent Correct, Standard Error, and Statistical Significance
Between Pairs of Gaza, Lebanon, and Syria as One Set and Between Jordan
and West Bank as A Separate Set on Math Test Subscales

Variable   Statistic   Gaza   Lebañon   Syria   Jordan   W.Bank   IN   N=175   N=131   N=225   N=319   N=183   AV	
NUM-CU         Sig Mean (12)         42.67 I.36         43.00 I.46         GL-S* 47.78 I.11         41.80 34.40 34.40 34.40 60.11         60.11           NUM-PK         Sig Mean SError         39.24 I.34 I.14 I.40 I.24 I.33         42.91 38.70 24.40 I.33         24.40 I.33 I.33         61.6           NUM-PS         Sig Mean SError         46.69 I.84 I.84 I.31 I.25 I.50 I.50 I.50         ****         ****           MEA-CU         Sig Mean SError         28.19 30.79 I.34 I.31 I.25 I.50 I.50 I.50 I.50 I.50 I.50 I.50 II.64 II.84 II.93 II.25 II.50 III.50 II.50 II.50 II.50 II.50 II.50 III.50 II.50 III.50 II.50 III.50 II.50 II.50 II.50 III.50 I	G.
Mean   42.67   43.00   47.78   41.80   34.40   60.00	<u>~·</u>
Mean   42.67   43.00   47.78   41.80   34.40   60.00	
(12)   S.Error   1.36   1.46   1.11   0.96   1.14     NUM-PK   Sig   Mean   39.24   41.14   42.91   38.70   24.40   61.6	70
NUM-PK         Sig Mean S.Error         39.24 1.69         41.14 1.84         42.91 1.40         38.70 1.24         24.40 24.40         61.0 61.0 61.0           NUM-PS         Sig Mean S.Error         46.69 1.64         47.66 1.84         51.87 1.31         46.90 1.25         40.40 1.50         68.8 68.8 68.8           (7)         S.Error         1.64         1.84         1.31         1.25         1.50           MEA-CU         Sig Mean         28.19 28.19         30.79 30.79         41.93 41.93         39.60 32.40         32.40 32.40         58.6           (7)         S.Error         1.86         2.33         1.79         1.34         1.70           MEA-PK         Sig Mean         23.81         23.92 25.62         25.44         28.40         17.30         45.4           (9)         S.Error         1.89         2.21         1.84         1.52         1.16           MEA-PS         Sig Mean         25.62         25.30         25.52         24.60         18.30         44.6           (7)         S.Error         1.29         1.42         1.32         1.03         1.10           GEO-CU         Sig Mean         46.67         59.08         64.53         50.90         38.30         65.7<	
Mean   S.Error   1.69   1.84   1.40   1.24   1.33   61.0	*******
NUM-PS         Sig Mean (7)         46.69 S.Error         47.66 1.64         51.87 1.84         46.90 1.25         40.40 40.40 40.40         68.8 68.8           MEA-CU         Sig Mean (7)         30.79 S.Error         41.93 1.79         39.60 1.34         32.40 1.70         58.6           MEA-PK         Sig Mean S.Error         23.81 1.89         23.92 2.21         25.44 1.84         28.40 17.30         17.30 45.4           MEA-PS         Sig Mean Sig Mean         25.62 25.62         25.30 25.52         24.60 24.60         18.30 1.10         44.6           GEO-CU         Sig Mean         46.67 46.67         59.08 64.53         64.53 50.90         50.90 38.30 38.30         65.7 65.7           GEO-PK         Sig Mean         40.19         47.33         51.41         39.90         30.8         62.4	60
Mean (7)         Mean (7)         46.69 (1.64)         47.66 (1.84)         51.87 (1.25)         46.90 (1.25)         40.40 (1.50)         68.8 (1.50)           MEA-CU         Sig Mean (7)         28.19 (1.84)         30.79 (1.93)         39.60 (1.25)         32.40 (1.70)         58.6 (1.70)           MEA-PK         Sig Mean (9)         23.81 (1.79)         23.92 (1.84)         25.44 (1.73)         45.4 (1.73)         45.4 (1.73)         45.4 (1.73)         45.4 (1.73)         45.4 (1.73)         46.6 (1.73) <td></td>	
Mean (7)         Mean (7)         46.69 (1.64)         47.66 (1.84)         51.87 (1.25)         46.90 (1.25)         40.40 (1.50)         68.8 (1.50)           MEA-CU         Sig Mean (7)         28.19 (1.84)         30.79 (1.93)         39.60 (1.25)         32.40 (1.70)         58.6 (1.70)           MEA-PK         Sig Mean (9)         23.81 (1.79)         23.92 (1.84)         25.44 (1.73)         45.4 (1.73)         45.4 (1.73)         45.4 (1.73)         45.4 (1.73)         45.4 (1.73)         46.6 (1.73) <td></td>	
(7)         S.Error         1.64         1.84         1.31         1.25         1.50           MEA-CU         Sig         30.79         41.93         39.60         32.40         58.60           (7)         S.Error         1.86         2.33         1.79         1.34         1.70           MEA-PK         Sig         30.79         41.93         39.60         32.40         58.60           MEA-PK         Sig         30.79         41.93         39.60         32.40         58.60           Mean         23.81         23.92         25.44         28.40         17.30         45.40           MEA-PK         Sig         30.79         1.84         1.52         1.16         11.60           MEA-PS         Sig         30.80         25.52         24.60         18.30         44.60           Mean         25.62         25.30         25.52         24.60         18.30         44.60           GEO-CU         Sig         30.80         64.53         50.90         38.30         65.70           GEO-PK         Sig         30.80         64.53         1.63         1.77           GEO-PK         Sig         30.80         64.53 <td< td=""><td>80</td></td<>	80
MEA-CU         Sig Mean (7)         28.19 S.Error         30.79 1.86         41.93 2.33         39.60 1.79         32.40 1.70         58.60 32.40         58.60 1.70           MEA-PK         Sig Mean (9)         23.81 S.Error         23.92 1.89         25.44 22.1         28.40 17.30         17.30 45.4         45.4           MEA-PS         Sig Mean Sig Mean (7)         25.62 1.29         25.30 1.42         25.52 1.32         24.60 18.30         18.30 44.6         44.6           GEO-CU         Sig Mean Sig Mean Mean 46.67         59.08 64.53 2.38         64.53 1.82         50.90 38.30         38.30 65.7         65.7           GEO-PK         Sig Mean Mean Mean 40.19         47.33         51.41 39.90         30.8         62.4	00
(7)         S.Error         1.86         2.33         1.79         1.34         1.70           MEA-PK         Sig         Mean         23.81         23.92         25.44         28.40         17.30         45.4           (9)         S.Error         1.89         2.21         1.84         1.52         1.16           MEA-PS         Sig         Mean         25.62         25.30         25.52         24.60         18.30         44.6           (7)         S.Error         1.29         1.42         1.32         1.03         1.10           GEO-CU         Sig         G         G         46.67         59.08         64.53         50.90         38.30         65.7           GEO-PK         Sig         G         S         39.90         30.8         62.4	********
MEA-PK         Sig Mean (9)         23.81 (1.89)         23.92 (2.11)         25.44 (2.40)         17.30 (45.41)         45.4 (2.41)         45.4 (2.41)         45.4 (2.41)         45.4 (2.41)         45.4 (2.41)         45.4 (2.41)         45.4 (2.41)         45.4 (2.41)         45.4 (2.41)         45.4 (2.41)         45.4 (2.41)         45.4 (2.41)         45.4 (2.41)         45.4 (2.41)         45.4 (2.41)         45.4 (2.41)         45.4 (2.41)         46.67	60
Mean (9)         23.81 (1.89)         23.92 (2.21)         25.44 (1.52)         28.40 (1.16)         17.30 (1.16)         45.4 (1.52)           MEA-PS         Sig (7)         Mean (1.29)         25.62 (2.30)         25.52 (2.4.60)         18.30 (1.10)         44.6 (1.32)         1.03 (1.10)         44.6 (1.32)         1.03 (1.10)         44.6 (1.32)         44.6	
(9)         S.Error         1.89         2.21         1.84         1.52         1.16           MEA-PS         Sig         Mean         25.62         25.30         25.52         24.60         18.30         44.6           (7)         S.Error         1.29         1.42         1.32         1.03         1.10           GEO-CU         Sig         G <l,s*< td="">         ***         ****           Mean         46.67         59.08         64.53         50.90         38.30         65.7           GEO-PK         Sig         G<s*< td="">         ****         ****           Mean         40.19         47.33         51.41         39.90         30.8         62.4</s*<></l,s*<>	<b>4</b> 0
MEA-PS         Sig Mean S.Error         25.62 1.29         25.30 1.42         25.52 1.32         24.60 1.03         18.30 1.10         44.60 1.10           GEO-CU         Sig Mean         G <l,s* 46.67         G<l,s* 64.53         *** 50.90         38.30 38.30         65.70           GEO-PK         Sig Mean         G<s* 40.19         G<s* 47.33         *** 51.41         39.90         30.8         62.4</s* </s* </l,s* </l,s* 	+0
Mean (7)         25.62 (1.29)         25.30 (1.32)         25.52 (1.32)         24.60 (1.32)         18.30 (1.10)         44.60 (1.32)           GEO-CU         Sig (1.29)         Mean (1.29)         GCL,S* (1.32)         Sig (1.32)         ***         38.30 (1.32)         65.70 (1.32)           (5)         S.Error (1.98)         2.38 (1.82)         1.63 (1.77)         1.77 (1.77)         38.30 (1.77)         65.70 (1	
GEO-CU Sig Mean 46.67 59.08 64.53 50.90 38.30 65.7 (5) S.Error 1.98 2.38 1.82 1.63 1.77 GEO-PK Sig Mean 40.19 47.33 51.41 39.90 30.8 62.4	50
Mean   46.67   59.08   64.53   50.90   38.30   65.7     (5)   S.Error   1.98   2.38   1.82   1.63   1.77     GEO-PK   Sig	
(5) S.Error 1.98 2.38 1.82 1.63 1.77 GEO-PK Sig G< S* *** Mean 40.19 47.33 51.41 39.90 30.8 62.4	70
GEO-PK Sig G< S* *** Mean 40.19 47.33 51.41 39.90 30.8 62.4	/0
Mean 40.19 47.33 51.41 39.90 30.8 62.4	
	40 <sup> </sup>
(3) S.Error 2.34 2.60 2.18 1.77 1.86	
GEO -PS Sig G <l,s* ***<="" td=""><td></td></l,s*>	
Mean 23.81 35.62 33.63 25.00 14.9 61.3 (3) S.Error 1.98 2.83 2.14 1.69 1.8	30
(3) S.Error 1.98 2.83 2.14 1.69 1.8 ALG -CU Sig	
Mean 33.14 34.73 34.44 30.6 26.20 46.9	90
(3) S.Error 2.67 2.81 2.25 1.84 2.29	
ALG- PK Sig G <l,s* ***<="" td=""><td>***************************************</td></l,s*>	***************************************
Mean 32.68 41.33 44.06 37.40 24.10 60.6	50
(5) S.Error 2.13 2.51 2.02 1.58 1.60 ALG -PS Sig 3 3 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4	
ALG -PS Sig G <s* ***="" 24.00="" 32.48="" 34.10="" 35.37="" 36.81="" 51.2<="" mean="" td=""><td>,<sub>0</sub></td></s*>	, <sub>0</sub>
(1) S.Error 1.52 1.62 1.33 1.19 1.28	
DAT -CU Sig G,L <s*< td=""><td></td></s*<>	
Mean 46.76 49.86 61.93 54.10 51.40 78.3	30
(2) S.Error 2.08 2.43 1.68 1.42 2.00	
DAT -PK Sig G< S* *** Mean 24.34 29.92 33.24 29.80 21.40 66.1	
Mean   24.34   29.92   33.24   29.80   21.40   66.1   (7)   S.Error   1.76   2.42   1.72   1.27   1.66	.U
DAT -PS Sig   1.76   2.42   1.72   1.27   1.66	
Mean 60.00 72.52 78.22 69.60 60.70 83.4	10
(6) S.Error 3.71 3.92 2.76 2.58 3.62	

Note: (Jordan and West Bank are added from earlier study).

Table 13 clearly shows that in comparison with the International means the performance of 8th grade students in Gaza was very weak in Measurement Procedural Knowledge (MEA PK), Data Analysis Procedural Knowledge (DAT PK) and Measurement Problem Solving (MEA PS) with average correct scores of 23.81, 24.34 and 25.62 respectively. Performance of 8th grade students in Lebanon was equally weak as has been the case in Gaza in Measurement Procedural Knowledge (MEA PK), Measurement Problem solving (MEA PS) and Data Analysis Procedural Knowledge (DAT PK) with average percent correct scores of 23.92, 25.30, and 29.92 respectively. The same holds true for the performance of 8th grade students in Syria as they obtained 25.44, 25.52, and 33.24 average percent correct scores on (MEA PK), (MEA PS), and (DAT PK) respectively.

It is worth noting that among the 20 participating countries in *IAEP II* Study the lowest average percent correct scores were on (MEA PS) and (MEA PK) with 44.60 and 45.40 mean scores respectively.

On the other hand, Table 13 also shows that average percent correct scores on Data Analysis Problem Solving (DAT PS) were the best of all scores in all the fields, as compared to the other average percent correct scores of all the other combinations. In fact, the average percent correct score of 78.22 obtained by 8th grade students in Syria on Data Analysis Problem Solving Component (DAT PS) is quite close to the international mean of 83.40 on the same component.

A graphic display of the average percent correct scores of these combinations along with the *IAEP II* study population parameters is given in figure 3.

Figure (3)

Mean Percent Correct Scores of the Five Arabic and the IAEP II Samples on Fifteen Subscales Defined by the Interaction of Five Content Areas and Three Cognitive Skills

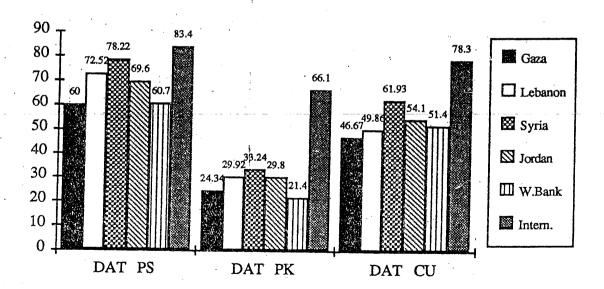


Figure (3) Continued

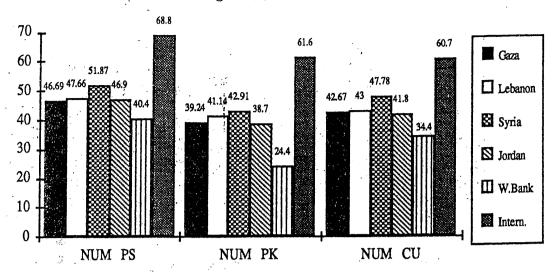
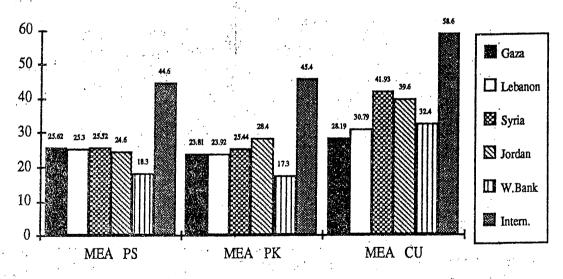


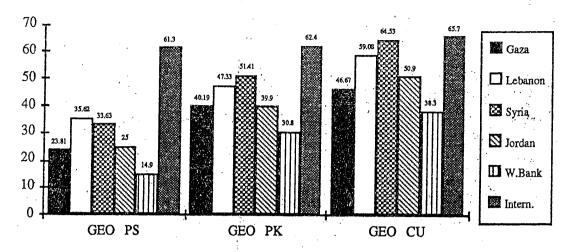
Figure 3 above shows that UNRWA 8th grade level of performance on Numbers topic at the three cognitive skills namely Conceptual Understanding (CU), Procedural Knowledge (PK) and Problem Solving (PS) is almost the same with the exception of Syrian students' performance which is significantly better ( $\alpha \le 0.05$ ) than those of the students of the other two fields on Conceptual Understanding and Problem Solving in the Number area. A plausible explanation to this narrowing of difference in the level of performance could be that the curricula in the host countries are similar as to the extent of emphasis given to the topic of Numbers.

Figure (3) Continued



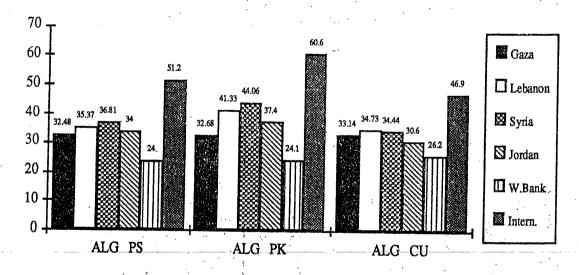
It is quite obvious from the part of figure (3) above that performance of 8th grade *UNRWA* students in all the fields is drastically low on Measurement Procedural Knowledge (MEA PK) and Measurement Problem Solving (MEA PS).

Figure (3) Continued



The part of figure 3 above also shows that *UNRWA* 8th grade students' performance on Geometry Problem Solving (GEO PS) as related to their performance on the other two cognitive skills in Geometry is the weakest. An incredibly low performance of the West Bank *UNRWA* students is obvious in Geometry Problem Solving (GEO PS) with an average percent correct score of 14.9 as compared to the international norm of 62.4%.

Figure (3) Continued



In the area of Algebra, as this part of figure 3 clearly shows, the performance across the three cognitive skills is almost uniform in Gaza, as well as, in the West Bank. In Syria, on the other hand, the average percent correct score on Algebra Procedural Knowledge (ALG PK) is about 10 percentage points higher than that on (ALG CU) and about 7 points higher than that on (ALG PS). As regards international comparisons, the mean percent correct score of Lebanon and Syria on (ALG CU) is 11 points below the IAEP II norm, while that on (ALG PK) is nearly 19 points below the

international norm; on (ALG PS) it is almost 15 points below the international norm.

# Gender and Achievement in Mathematics

## I Comparative Math Achievement of Male and Female Students in Gaza, Lebanon, Syria, Jordan and West Bank

So far the differences studied in the previous section were based upon groups of *UNRWA* students without consideration of their gender. Findings from numerous empirical studies tell us, for certain, that gender differences in students achievement do exist. The following section will investigate the gender differences in Math and Science achievement, in each *UNRWA* field separately.

Table 14 presents the average percent correct score (M), Standard Deviation (SD) and Standard Error (SE) of means and the significance of difference between male and female students' achievement in Mathematics in each field on the total Math test score and its subscales.

## Gender Differences In Gaza

As shown in Table 14, on one Math test component namely Geometry Procedural Knowledge out of the 24, the mean score of the female students is statistically significantly higher than that of the males at  $\alpha = 0.05$ . In terms of absolute differences between the means of the two groups, it is worth mentioning that the means of female students in Gaza tend to be higher than those of the male students in most of the 24 components of the test. This could be attributed to the fact that the girls tend to stay at home and abstain from participating in the Intifada which gives them better opportunity for self learning.

## Gender Differences In Lebanon

Means and Standard Deviations of male and female students' scores on each Math test component including the total Math test score are shown in Table 14. On 17 out of 24 variables male students' scores are significantly higher than those of the female students, some differences are very wide as in the case of Geometry where the difference is more than 22 points. Even on the remaining 7 variables male students outscored the females but the differences were not significant at the 0.05 Alpha level of significance.

**Table (14)** 

Average Percent Correct (M), Standard Deviation (SD) Standard Error (SE) and Statistical Significance Between Male and Female Student Achievement in Math Test in Each Field

	MEA CU		CIMIN	אלן זיא			NOM PK				NUM CU			Math	Cetal		Variable	-
SE M	Sig	SD	Z &	SE	SD	Z	sig	SE	SD	Z	sig	SE	SD	X	sig		le	
25.76 26.13 02.78	02.44	22.62	44.32	02.42	22.72	37.12		02.02	18.95	41.67/		01.63	15.27	34.88		N=88.	Male	. <del>.</del>
30.65 22.86 02.45	02.20	20.52	49.10	02.20	21.88	41.40		01.83	17.10	43.68		01.55	14.45	37.84	***	N=87	Female	Gaza
28.19 24.61 01.86	01.64	21.67	46 69	01.69	22.34	39.24		01.36	18.03	41.67		01.13	14.90	36.35		N=175	Total	.i
37.57 28.34 03.58	02.90		51 03 —	02.69	21.35	46.91	*	-	15.61	45.37		17.74	13.79	45.59	*	N=63	Male	<b>[</b> =4
24.51 23.48 02.85	02.27	18.72	44 54	02.33	19.18	35.78		02.12	17.47	40.81		01.55	12.79	35.02		N=68	Female	Lebanon
30.79 26.67 02.33	01.84	21.09	77 66	01.84	20.93	41 14				43.00		01.25	14.25	40.10		N=131	Total	<b></b>
48.05 28.67 02.72	01.96	20.61	*		21 29	44 85	_	<del></del>		51.65	*	01.45	15.25	47.03	*		Male	
35.97 23.53 02.20	01.61	17.22	3	01.94	20.75	21 23		01.60	17.12	44.01		01.33	14 19	40 13		N=114	Female	Syria
41.93 26.82 01.79	01.31	19.66		01.40	21 06	3 2 1			16.63	47.78		01.01	15.88	43 53		N=225	Total	
38.04 31.35 29.30 23.66 01.81 01.44	01.37	51.42 22.09	*	01.37	32.74	2	01.00	-	17 17	46 78	*	20.00	15.85	43 60	*		Male	Ent
	01.14	46.89 18.73		01.27	39.81	2	01.00	50.10	17 10	23 26.	00.00	78 00	14.03	30 00		N=269	Female Total	Entire Sample
34.65 26.78 01.16	00.90	49.13   20.79		00.93	41.26		2./2	02.71	17.76	3	00.00	27.17	15 10	3	1001	N-531	Total	nple
40.40 24.84 01.96	01.88	47.40 23.87		22.01 01 73	38.40		7.C.TO	01.34	1734	3	01.24	13.09	39.40	<b>.</b>	TOT=NI	N-121	Mala	
38.80 23.15 01.84	01.65	46.40		22.35			01.34	16.83	39./0	30 70	01.18	14.85	37.20	}	OCT =N	NI 150	Eamala	Jordan
39.60 - 01.34	01.25	46.90	7.1.0	21 24	38.70		00.96	3 '	41.80	2	00.86	<u>}</u>	38.30	)	N=319	TETOT	7	
34.90 - 02.55	02.37	39.60	75.20	27 -	22.20		01.90	; ; ;	35.00		01.35	}	28.90		N=83			*
30.30 - 02.28	01.92	41.00	01./8	7 '	26.20		01.73	1	33.70		01.07	'	29.20		N=100	Female	i] ,	West Bank
32.40 - 01.70	01.50	40.40	01.33	3 '	24.40		01.14	ı	34.40		00.80	'	29.00		N=183		_	ak

Table (14) Continued

		Gaza		Н	Lebanon	, D	. tt.	Syria		Ent	Entire Sample	aple		Jordan		₩.	West Bank	F
Variable	Male N=88	Female	Total	Malc N=63	Female Total	Total	Male N-111	Female	Total	Male	Female				Total	Male		Ton
MEA PK sig	_	+		*	į	1		- LE T-1.	11-660	107=107	N=209 N=531		N=161	N=158	N=319	N=83	N=198	N=183
M	23.86	5 23.76	23.81	34.39	14.22	23.92	26.73	23.39	25,44	27.61	21 19	2436	28 85	28 30 	<b>3</b> 6 A	1 5 3 0		1
SD	27.19	22.68		26.75		25.24	28.36	36 48	77 46	77 78	33 05				6.40	ויייים	19.00	17.30
SE			01.89			02.21	02.69	02.49	01 84	01 78	01 46	01 13	00.00	3 2	? ?	3 '		<u>.</u> ,
MEA PS   sig					:	انـ	*			1	3::0	3	04.40	71.70	70.10	1.40	02.13	01.10
X	25.33	3 25.62	25.47	28.12	22.69	25.30	28.83	22.31	25.52	27 48	23 47	25 45	36.40	33 80 -		1000		
SD		9   16.18		16.62	15.46	16.20	20.48	18.72	19.83	18.81	17 14	30 or	10.20		24.00	10.00	00.81	18.30
SE	01.93	3 01.73		02.09	01.87	01.42	01.94	01.75	01.32	01.16	01.05	00.78	01 53		21	21		2 ,
GEO CU   sig				*						*		0				20.10	07.70	OT.TO
×	44.55	5 46.67	45.60	69.52	49.52	59.08	65.05	2	64.53	59.24	54 72	56 95	40 30	ላን ላን		36 70		ງ. ວ
SD	26.91	25.64	26.23	23.52		27.27	27.56	27.03	27.24	28.38		28 11	20 50 28 60		20.50	07.70	<u>.</u>	30.30
SE	02.87			02.96		02.38	02.62	02.53	01.82	01 75	01 60	01 22	02.20	2 6	21 '			? .
GEO PK   sig	*			*			*			*		2	*	04.40	01.00	65.70	14.70	O1.//
×			40.19	57.14	38.24	47.33	57.96	45.03	51.41	49.49	43.99	47 70	42 70	37 10	30 00	1 2 3 1 1 1		1
SD	29.47	7 31.15	31.01	28.98	27.78	29.80	31.05	33.18	32.72	32.06	31.31	31 77	37.75			.,		74.41
-	03.14	1 03.34	02.34	03.65	03.37	02.60	02.95	03.11	02.18	01.98	01.91	01.38	02.58	03 41	27	3	3 '	2 '
GEO PS   sig		* -:-		#						*						70.70	72.37	07.70
M		7 25.67	23.81	47.09	25.00	35.62	39.94	27.49	33.63	35.62	26 27	30 89	78 80	21 10	3	) - -		3
SD	23.10	29.07	26.23	33.14	27.84	32.34	32,36	30.80	32.12	31.34	29.43					Per -70	29.70	50.00
SE	02.46	03.12	01.98	94.17	03.38	02.83	03.07	02.88	02.14	01.94	01 79	01 33			21 60	3 ,	3	? '
											32.1	O1.00	17.70	14.40	V2.09	102.94		

Table (14) Continued

	· · · · · · · · · · · · · · · · · · ·	Gaza		<b>1</b>	Lebanon	<b>.</b> 5		Syria		Ent	Entire Sample	nple		Jordan		<b>*</b>	West Bank	nk
Variable	Male	Remale	Total	Male	Empole	77		E .	7		1							
	Z 11 88				S C		N-111	N-114	TO THE	Maic	remaie 1 otal	10121	Male	Male Female Total	Total	Male	Male Female	Total
ALG CU sig				*						707-11	14=207	TCC=NI	191=N	N=158	N=319	N=83	N=100	N=183
×	33.52	32.76	33.14	40.48	29.41	34.73	38 74	30 26	34 44	37 40	30 06	3		3				
SD	23.10	33.96	35.38	32.20		32 16	37 81	24 20	23.79		3 .00	27.07	30.70	30.40 30.60		23.30	28.50	26.20
SE	02.46		02.67	2.8		02.81	03.11	03.21	02.25	3 ;	2 2	01.67	31.16	34.71		•	·	ı
ALG PK sig								~		*	04.01	/±.10	04.40	07.70	01.84	03.25	03.20	02.29
×	34.09	23.68	33.39	46.26	36.77	41.33	45.56	42.61	<b>4</b>	41 88	37 03	30 07	30 10	3	3	)   	<u> </u>	
SD	27.94	28.65	28.22	25.62		28.69	31.51	29.20	30.33	29.42	20.63	30 46		20,45	ا 40.40	25.60	22.90	14.10
$\vdash$	02.98	03.07	02.13		<u> </u>	02.51	02.99	02.73	02.02	01.83	01 81		03.16	22.67	2	}		1
ALG PS sig							*							02.30	9C.10	02.31	02.22	01.60
X	28.98	36.02	32.48	37.57	33.33	35.37	39.64	34.06	36.81	35 56	34 51	35 03						
SD	18.83	20.79	20.08	18.20	18.66	18.49	21.09	18.60	20.02		1031	10.73			34.00	26.70	21.80	24.00
-	02.01	02.23	01.52	02.29	02.26	01.62	02.02	01.74	01.33	01.25	01 18	20.86	01.40			}	}   .1 	
DAT CU sig										*		00.00	70.10	1.7	67.10	/8.TO	01.74	01.28
×	44.32	49.04			47.06	49.87	64.26	59.65	61.93	54 38 86 43	\$ ₽ -	53 05		73 00				
SD	26.13	28.68	-	30.90		27.85	26.86	23.24	25.14	28.89	36.00	77 45	30.21	31.00	24.10	00.00	52.00	51.40
+	02.78	03.07	02.08	03.89		02.43	02.55	02.18	01.68	01.79	01.59	01 10		01.74				}
DAT PK   sig				#			*			*		31.1.	04.40	01./4 01.42	+	03.10	02.61	02.00
X	24.09	24.60	24.34	39.68	20.88	29.92	38.74	27.90	33.24	2 2 3 3	35.05	20 40						
SD	23.52	23.17	23.28	30.58				22.36	25 70	28.03		27.75			29.80	21.40	21.40   21.40	21.40
SE	02.51	01.48	01.76				02.65	3	01 73	01 73					•	,	,	•
					ŀ	1			3	01.70	⊢	71.10	01.93	2.2	01.27	02.78	02.78   01.97   01.66	<u>?</u>

Table (14) Continued

			MUM				73			***************************************	PK				CU				DAT PS		Variable	
SE	SD	Z	Sig	SE	SD	X	sig	SE	SD	X	sig	SE	SD	X	sig	SE	SD	X	sig		rp	·
01.94	18.18	40.87	. •	01.54	14.45	32.76		02.08	19.54	32.03		01.72	16.13	40.00		05.27	49.45	59.09		N=88	Male	
01.68	15.65	44.29	-	01.47	13.76	36.54		03.02	18.83	34.69		01.61	15.04	42.48		05.26	49.08	60.92		N=87	Female	Gaza
01.29	17.01	42.57		01.07	14.20	34.64		01.45	19.19	33.35		01.18	15.60	41.23		03.71	49.13	60.00		N=176	Total	
01.82	14.46	27.28	*	01.82	14.46	41.73	*	02.37	18.84	45.15	*	01.80	14.29	49.78	*	.04.99	39.58	80.95	*	N=63	Male	F
01.28	14.10	40.13		01.46	12.01	33.76		01.97	16.24	31.15		01.85	15.30	40.41		05.84	48.14	64.71		N=68	Female Total	Lebanon
01.28	14.64	43.57		01.20	13.78	37.60	·	01.64	18.83	37,86		01.35	15.50	44.92		.03.92	44.81	72.52		N=131	Total	
01.38	14.57	50.87	*	01.52	16.04	43.51	*	01.92	20.27	43.34	*	01.29	13.64	54.38	*	.03.51	37.03	83.78	*	N=111	Male	
01.39	14.82	43.70		01.21	12.92	35.09		01.79	19.08	37.49		01.43	15.28	47.82		04.19	44.69	72.81		N=114	Female	Syria
01.01	15.10	47.24		01.01	15.11	39.47	*	01.26	19.83	40.38		00.99	14.83	51.06		02.76	41.37	78.22		N=225	Total	
01.01	16.38	46.65	*	00.98	15.84	39.47		01.26	20.43	40.00	*	00.98	15.91	48.44	*	02.69	43.44	74.81	*	N=262	Male	Entir
00.91	14.96	42.99		00.79	12.97	35.22		01.12	18.43	34.99		00.94	15.49	44.22		02.87	47.14	66.91		N=269	Female	ire Sample
	15.77	44.79		00.63	14.60	37.32		01.13	26.08	24.36		00.94	15.82	46.31		01.97	45.51	70.81		N=531	Total	ıple
01.33	16.94	43.00		01.26	15.99	37.40		01.60	20.29	36.40		01.24	15.75	44.70		03.50	44.38	73.30		N=161	Male	
01.24	15.54	41.10		01.21	15.22	33.30			19.37	35.00		01.17		43.10		03.50 03.79 02.58	47.58	65.80		N=158	Female	Jordan
01.24 00.91	1	42.10 32.10		90.88	•	33.30 35.40		01.11	1	35.00 35.70		00.85		43.10   43.90			,	69.60		N=319	Total	
01.66		32.10		01.32	<b>i</b>	27.70		01.66	1	23.30		01.48	ا <u>ع</u> ــــــــــــــــــــــــــــــــــــ	36.10		05.50	,	54.20		N=83	Male	**
01.66 01.25	i	33.10		01.21		27.40		01.40	1	24.00		01.23	,	36.40		04.76	)	66.00		N=100	Female	West Bank
01.01		32.60		00.89		27.50		01.07		23.70		00.95		36.30		03.62		60.70		N=183	Total	ak

Table (14) Continued

			Gaza			Lebanon			Syria		Entii	ire Sample	ple		Jordan		¥	West Bank	F
		4	_					_											
Variable		Male	Female	Total	Male	Female	Total	Male	Female	Total	Male	Female Total	Total	Male	Female Total	Total	Male	Female Total	Total
		N=88	N=87	N=176	N=63	N=68	N=131	N=111	N=114	N=225	N=262	N=269 N=531		N=161	N=158	N=319	N=83	N=100	N=183
MEA	gis				*		,	*			*		·		,				
	Z	25.09	26.35	25.71	31.75	21.15	49.48	32.78	25.71	29.20	29.95	24.76	27.32	30.10	27.80	29.00 21.70	21.70	21.10	21.40
	SD	14.82	12.94	13.89	14.46	12.49	14.43	18.40	15.55	17.34	16.67	14.11	15.63	15.90	14.86	1	•		'
	SE	01.58	01.39	01.05	01.82	01.51	01.26	01.75	01.46	01.16	01.03		00.68	01.25	01.18	01.18   00.86	01.38	01.04	00.84
GEO	sig				*			*			*			. Ť.					
	X	35.33	41.06	38.18	60.03	39.71	26.25	56.26	48.88	52.53	50.14	44.04	47.05	41.90	39.80   40.80	40.80	28.70	30.80	29.90
	SD	20.28	22.30	21.44	21.32	20.53	14.43	23.34	22.28	23.06	24.25	22.18	23.40	23.93	22.86	i	ľ	,	
	SE	02.16	02.39	01.62	02.69	02.49	01.26	02.21	02.09	01.54	01.50	01.35	01.02	01.89	01.82 01.81		02.04	01.90	01.39
ALG	sig										*				i C				
i i	X	31.97	34.02	32.99	42.01	34.41	38.07	42.28	37.54	39.88	38.75	35.61	37.16	36.80	35.61   37.16   36.80   33.50   35.20   25.80	35.20	25.80	23.20 24.40	24.40
	SD	19.16	20.31	19.71	16.27	19.97	18.61	21.81	20.50	21.24	20.22	20.30	20.30 79.71	79.71	21.37	•	1		'
	SE	02.04	02.18	01.49	02.05	02.42	01.63	02.07	01.92	01.42	01.25	01.24	00.88	01.05	01.70	01.70 01.15	01.54	01.55	01.10
DAT	sig				#			*			*						-		
	X	37.72	36.78	35.75	48.68	34.48	41.31	52.25	43.47	47.80	45.50	39.03	42.23	43.30	43.30 41.30 42.30	42.30	34.90	36.60	35.80
,	SD	, 18.21	20.14	19.17	25.43	16.75	22.45	20.24	17.96	19.58	22.31	18.75	20.82	19.29	15.72	•	ı	,	,
	SE	01.94	01.94   02.16	01.45	03.20	02.03	01.96	01.92	01.68	01.31	01.32	01.14	.00.90	01.00	01.25	00.90 01.00 01.25 00.99 02.31 01.65 01.38	02.31	01.65	01.38

## Gender Differences In Syria

Table 14 clearly shows that, on 16 out of 24 variables of the test, differences are statistically significant in favor of the male students. The differences, though statistically significant, are marginal. The female students did not outscore the male students in any of the 24 variables.

## Gender Differences In Jordan and West Bank

In Jordan, sex differences were significant only in two test components namely Algebra Problem Solving (ALG PS) and Geometry Procedural Knowledge (GEO PK) in favor of male students.

In West Bank, sex differences in Math achievement did not emerge. A plausible explanation seems to be found in the fact, that UNRWA Math performance, irrespective of the sex of the students was across-the-board so poor that the floor effect probably swamped all the potential differences that could have emerged due to sex of the students (Ahlawat et al., 1992).

## II Comparative Math Achievement of Male and Female Students in The Entire Sample (Gaza, Lebanon and Syria)

Table 14 also presents average percent correct score (M), Standard Deviation (SD) and Standard Error of Means between male and female student achievement in the entire sample covering Gaza, Lebanon, and Syria. The asterisk at the upper right corner under the column headed "Entire Sample" indicates significant differences between the means of male and female students on the total test score and its 23 subscales.

Table 14 clearly shows that the male students' scores were significantly higher on 19 out of the 24 test components including the total test score. Even on the remaining 5 non-significant differences male students outscored the females.

## Chapter IV

# INTRENATIONAL COMPARISONS IN SCIENCE ACHIEVEMENT

## Scale of comparison

In this part we will present the performance levels of eighth grade students on *IAEP II* Science achievement test in Gaza, Lebanon, Syria, (Jordan and West Bank from earlier study), along with the International norms on each component of the Science Test in a comparable framework.

IAEP II science test consisted of 72 multiple choice items designed to measure students' Knowledge of facts, concepts and principles (KNOW), Application of knowledge to solve simple problems (APP), and Integration of knowledge and principles to solve more complex problems (INT), in four major topics or disciplines of science namely: Life Sciences (LIF), Physical Sciences (PHY), Earth and Space Sciences (ESP) and Nature of Science (NAT). Including total test score on all the 72 items, the science test contained 18 distinct components.

The item correct scores on the whole Science Test and on its various components were transformed into percent correct scores. All the comparisons, intergroup or intragroup, and all the estimates of different parameters have been based on the percent correct scores. An explanation for doing this have been already provided in Chapter III of this monograph.

Earlier in 1992, the same *IAEP II* Science Test was administered to a sample of 317 eighth grade students in Jordan and 180 eighth grade students in the West Bank.

The null hypothesis of no difference in the achievement of students in Gaza, Lebanon and Syria was tested using One-Way Analysis of Variance at the  $\alpha = 0.05$  level for each component of the Science Test. The results of such analysis are presented in Tables 15, 16, 17.

Table 15 includes average percent correct scores, Standard Error and statistical significance on Total Science Test and the four main content areas in Science namely Life Sciences, Physical Sciences, Earth and Space Sciences and the Nature of Science, as well as, the International norm

(average percent correct score) for each of the Science Test components. The International norm was computed by the *IAEP II* study on the basis of aggregated national samples of 20 countries which participated in 1991 study conducted by Educational Testing Service (ETS) in collaboration with the national teams of the participating countries.

The name of the science variable measured and the number of test items measuring it is given in the first column entitled "variable" in table 15. The level of statistical significance for each of the One-way Analysis of Variance between the three fields (Gaza, Lebanon and Syria) was tested at the  $\alpha = 0.05$  level. The bar over G and L in the symbol  $\overline{G,L} < S$  is used to indicate that there is no significant difference between Gaza and Lebanon but Syria is significantly higher than each of the other two fields. The statistics about Jordan, West Bank, and *IAEP II* samples are taken from an earlier study conducted by (Ahlawat et al., 1992).

The overall science achievement is represented by the Total Science Test score which is given in the first row of Table 15. The mean percent correct score of 8th grade students in Syria is 61.7 which is higher by 6.4, 7.4, 11.0 and 13.2 percentage points than the means of Lebanon, Jordan, Gaza and West Bank respectively. In comparison to the International mean, the mean of students in Syria is almost 5 points below the norm, while Lebanon's is almost 12 points below the norm, Jordan's is 13 points below the norm, Gaza's is almost 16 points below the norm, and West Bank's is almost 19 points below the norm. The highest achieving country among the 20 IAEP II participants was Korea with its average of 78 percent correct, the lowest among them was Fortaleza Brazil with 46% correct average.

The mean score of the entire *UNRWA* sample when treated as one entity came up to 54.4 which places *UNRWA* third from the bottom (i.e., just above Sao Paulo Brazil and Fortaleza Brazil).

Average Percent Correct, Standard Error and Statistical Significance Between Gaza, Lebanon, Syria, As One Set and Between Jordan and West Bank As a Separate Set on the Science Test and its Four Content Areas

			Ave	rage Per	cent Co	rrect	
Variable	Statistic	Gaza N=175	Lebanon N=128	Syria N=226	Jordan N=317	W.Bank N=180	INT. AVG.
Total Science (72	Sig Mean S.Error	50.7 1.0	55.3 1.0	G <l<s *<br="">61.7 0.9</l<s>	54.8 1.13	*** 48.5 00.80	67.1
LIF (25	Sig Mean S.Error	53.6 1.1	60.3 1.2	G <l<s* 64.6 0.9</l<s* 	57.3 0.9	*** 50.5 1.2	66.7
PHY (26	Sig Mean S.Error	48.6 1.1	51.7 1.2	G,L <s* 58.8 1.0</s* 	51.4 0.9	** 47.7 1.2	65.2
ESP (09	Sig Mean ) S.Error	53.8 1.6	55.6 1.5	G,L <s* 66.0 1.4</s* 	57.9 1.2	*** 49.0 1.5	67.1
NAT (12	Sig Mean ) S.Error	46.7 1.5	52.7 1.5	G <l<s* 58.6 1.3</l<s* 	54.5 1.1	*** 46.0 1.6	72.5

#### Notes:

- 1. The number of test items comprising a subscale is given in the brackets.
- 2. Mean is taken over the percent correct score on each subscale.
- Statistical significance of differences among Gaza, Lebanon and Syria was tested Via Newman Keul Multiple Comparisons at the .05 level, while the significance of t-test between Jordan and West Bank is taken from (Ahlawat et al. 1992).

When (in each field of *UNRWA* operations) students' performance on different subscales of the Science Test is compared with respect to judging their relative weakness and strength in different areas of Science, Table 15 clearly shows that the average percent correct scores on the Nature of Science component are the lowest, in all the fields. Interfield comparisons on this subscale, however, place Syria with its 58.6 average percent correct score, at the top position. Despite Syria's first place among the *UNRWA* fields its mean score of 58.6 is 14 percentage points lower than international (*IAEP II*) norm of 72.5.

One plausible explanation of the weakness in this area is that the science curricula to which the 8th grade students were exposed did not put enough emphasis on this content. The same thing applies to the average percent correct scores of Jordan and West Bank as obtained from earlier report by (Ahlawat et al., 1992).

percent correct scores of Jordan and West Bank as obtained from earlier report by (Ahlawat et al., 1992).

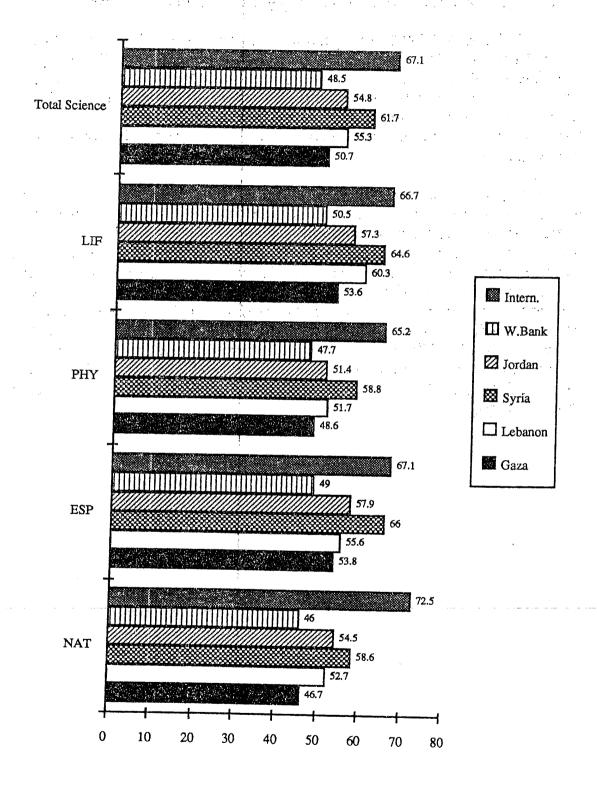
A graphic display of average percent correct scores on Science Test components along with *IAEP II* study population is given in Figure 4.

Table 16 includes the average percent correct scores, Standard Error and statistical significance on the three major skills measured, namely: Knowledge of facts, concepts and principles (KNOW), Application of knowledge to solve simple problems (APP) and Integration of knowledge and principles to solve more complex problems (INT). Furthermore, Table 16 also shows the average percent correct score of the *IAEP II* population on these skills.

It is quite evident that there is a general weakness among all *UNRWA* 8th grade students in the Integration skill where the highest average percent correct score of 51.1 was obtained by Syria which is about 15 points less than the International norm of 66.3. However, upon looking at the average percent correct scores on the other two skills, one can see that results were quite acceptable on the Knowledge skill in most of the fields as compared to the international norm, while the Application skill still remains as a general area of weakness in all fields except for Syria with its average percent correct score of 62.3 as compared to the International norm of 65.5. A graphical representation of the average percent scores on the three main skills of the different fields along with that of the *IAEP II* study population is given in Figure 5.

Figure (4)

Average Percent Correct Scores of the 8th Grade Students of the Different Samples on the Science Test and its Content Subscales



经经验的基础

Average Percent Correct, Standard Error and Statistical Significance Between Gaza, Lebanon, and Syria As One Set and Bewteen Jordan and West Bank As a Separate Set on the Science Test and its Skill Subscales

<u> </u>	····			Ave	rage Per	cent Cor	rrect	
Varial	ble	Statistic	Gaza N=175	Lebanon N=128	Syria N=226	Jordan N=317	W.Bank N=180	INT. AVG.
KNOW	(21)	Sig Mean S.Error	56.7 1.2	65.8 1.3	G <l<s *<br="">70.3 0.9</l<s>	63.1 0.9	*** 56.9 1.3	70.5
APP	(32)	Sig Mean S.Error	51.1 1.1	55.2 1.1	G <l<s* 62.3 1.0</l<s* 	54.3 1.0	*** 47.4 1.2	65.5
INT	(19)	Sig Mean S.Error	43.2 1.2	44.0 1.2	G,L <s* 51.1 1.1</s* 	46.2 1.0	*** 41.1 1.2	66.3

Notes:

1. (Jordan and WB data added from earlier study: Ahlawat et al., 1992).

Figure (5)

Average Percent Correct Scores of the 8th Grade Students of the Different Samples on the Cognitive Skill Subscales of the Science Test

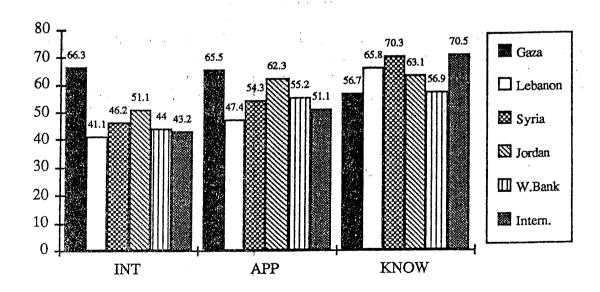


Figure 5 above shows a general weakness among all *UNRWA* 8th grade students in the Integration skill (INT) while relative performance on the Knowledge skill (KNOW) was the best, in each field.

Table 17 includes 10 combinations of the four topics and the three skills measured (Note that Earth and Space Sciences and Nature of Science were measured on two skills each). Table 17 also includes the average percent correct score, Standard Error, statistical significance and the International norm for each combination.

If we compare the performance of 8th grade students in Gaza, Lebanon and Syria on these 10 combinations internally among the 10 subscales a general weakness in the performance of UNRWA students on Physical Sciences Integration (PHY INT) followed by Life Sciences Integration (LIF INT) and Nature of Science Integration (NAT INT) is evident from Table 17 and Figure 6. Whereas, the performance of 8th grade students in Gaza, Lebanon and Syria on the remaining combinations stands at a better footing especially on Life Sciences Knowledge (LIF KNOW), Physical Sciences Knowledge (PHY KNO) and the Earth and Space Science Knowledge (ESP KNO) on which they are at their best.

Further comparing the performance of 8th grade students in each of Gaza, Lebanon and Syria, on these 10 combinations with the International norm, it can be seen that on the variable "LIF KNOW" the mean of students in Gaza and Lebanon is 12 and 0.8 points less respectively than the International norm of 69.0, while the mean of students in Syria is one point higher than that of the International norm. In fact mean scores of students in Syria were quite close to the International norm on other variables such as (PHY APP) (ESP KNO), (ESP APP) with a difference of one point only.

UNRWA students performance in Jordan and West Bank, evidently, depicts a weakness on all combinations with specific reference to Life Sciences Application and Integration skills (LIF APP and LIF INT), Physical Sciences Application and Integration skills (PHY APP and PHY INT) and Nature of Science Integration skills (NAT INT).

A graphic display of average percent correct scores along with *IAEP*II study population on the Science Test's combination subscales is given in Figure 6.

**Table (17)** Average Percent Correct, Standard Error and Statistical significance Between Gaza, Lebanon, Syria As One Set and Between Jordan and West Bank As a Separate Set on the Science Test and its Subscales

ne de de			Ave	rage Per	cent Co	rrect	
Variable	Statistic	Gaza N=175	Lebanon N=128	Syria N=226	Jordan N=317	W.Bank N=180	INT. AVG.
LIF KNOW (16)	Sig Mean S.Error	56.8 1.3	68.2 1.4	G <l,s.* 70.0 1.0</l,s.* 	63.6 0.9	*** 57.2 1.4	69.0
LIF APP	Sig Mean S.Error	51.0 1.6	48.4 2.0	GL <s* 56.1 1.4</s* 	47.4 1.4	*** 37.7 1.4	62.2
LIF INT (03)	Sig Mean S.Error	42.3 2.1	41.4 2.2	G,L <s* 51.0 1.8</s* 	43.4 1.5	** 37.0 2.0	66.7
PHY KNO (04)	Sig Mean S.Error	55.4 1.9	56.6 2.2	G,L<\$* 68.0 1.4	60.6 1.5	*** 53.1 1.9	73.4
PHY APP (14)	Sig Mean S.Error	50.2 1.2	56.2 1.4	G <l<s* 62.2 1.1</l<s* 	53.9 1.0	*** 50.8 1.4	63.6
PHY INT (08)	Sig Mean S.Error	42.4 1.5	41.2 1.8	G,L <s* 48.4 1.5</s* 	42.5 1.5	39.5 1.6	62.5
ESP KNO (01)	Sig Mean S.Error	61.7 3.7	64.8 4.2	G,L <s* 78.3 2.7</s* 	65.6 2.7	57.8 3.7	80.1
ESP APP (08)	Sig Mean S.Error	52.9 1.6	54.5 1.5	G,L <s* 64.5 1.4</s* 	59.9 1.3	*** 47.9 1.5	65.5
NAT APP (04)	Sig Mean S.Error	51.3 2.1	62.9 2.2	G <l,s* 67.8 1.8</l,s* 	61.4 1.6	*** 49.3 2.4	77.8
NAT INT (08)	Sig Mean S.Error	44.4 1.5	47.7 1.5	G,L <s* 54.0 1.4</s* 	51.0 1.2	*** 44.3 1.6	70.5

Notes:

1. (Jordan and WB data added from earlier study, Ahlawat et al., 1992).

2. \* Statistically significant at (p < .05)

\*\* Statistically significant at (P < .01)

\*\*\* Statistically significant at (P < .001)

Figure (6)

Average Percent Correct Scores of the 8th Grade Students of Different Samples on Content by Skill Subscales of the Science Test

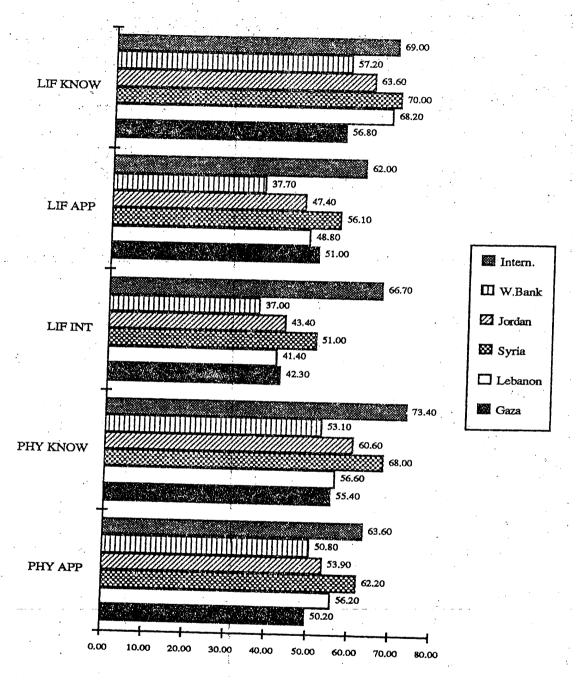
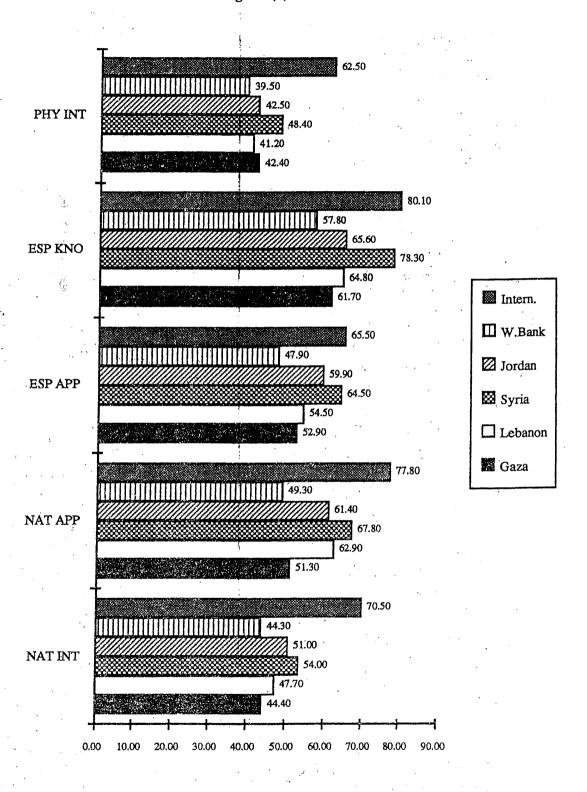


Figure 6 shows that *UNRWA* students' performance on Life Sciences Knowledge (LIF KNOW) in relation with their performance on other scales ranks first followed by Physical Sciences Knowledge (PHY KNOW), Life Sciences Application (LIF APP) and Life Sciences Integration (LIF INT). A study of such indicators could assist in the future model of the unified curricula for the Palestinian students.

Figure (6) Continued



#### Gender and Achievement in Science

# I Comparative Science Achievement of Male and Female Students in Gaza, Lebanon, Syria, Jordan and West Bank

The differences studied in the previous section were based upon groups of *UNRWA* students pooled together without consideration of their gender. Findings of numerous empirical studies tell us, almost for certain, that gender differences in students achievement do exist.

Table 18 presents the mean percent correct score (M), Standard Deviation (SD) and statistical significance between male and female students' achievement on each component of the science test including the total Science Test score separately for each field namely: Gaza, Lebanon, Syria, Jordan and West Bank.

Table (18)

Average Percent Correct (M), Standard Deviation (SD) and Statistical Significance Between Male and Female Student Achievement on Science Test in Each Field

	7	G	aza	Leb	anon	S	yria	Jor	dan	W.	Bank
Contents	STA	Male N=87	Female N=88	Male N=62	Female N=66	Male N=114	Female N=112	Male N=160	Female N=157	Male N=76	Femal
Total Sc.	<del>M</del>	52.1	49.3	58.5*	52.3	64.5*	58.8	58.5*	50.9	47.0	49.7
	SD	14.6	11.6	10.6	10.1	14.0	12.3	16.3	12.8	14.9	14.3
LIF	M	53.6	53.6	62.6	58.2	66.7*	62.5	60.5*	54.0	47.8	52.4
ξ.	SD	16.0	13.2	13.7	12.9	14.3	13.3	16.7	20.0	16.7	14.6
PHY	M	30.8*	46.4	33.1*	48.4	62.8*	54.8	55.9*	46.8	47.8	47.6
	SD	15.0	13.4	14.4	12.4	15.7	13.4	17.2	14.7	16.7	15.7
ESP	M	58.4*	49.4	62.4*	49.3	70.3*	61.7	63.2*	52.4	47.5	50.1
~	SD	22.0	18.6	15.1	16.8	21.2	18.9	22.8	19.4	20.1	20.2
NAT	M	46.8	46.6	54.7	50.9	59.4	57.7	56.5	52.4	43.1	48.1
	SD	21.0	18.0	17.9	15.1	19.5	19.7	21.8	18.9	22.2	21.2
KNOW	M SD	58.1	35.4	69.2*	62.6	72.9*	67.6	66.7*	59.5	53.8	59.2
APP	M M	17.4 53.4*	14.9	14.4	14.6	14.3	14.6	21.5	14.9	19.2	16.9
AFF	SD	15.9	48.9 12.5	59.9*	50.9	65.7*	58.9	58.5*	50.1	47.2	47.6
ĪNT	$\frac{13D}{M}$	43.2	43.2	12.3	11.0 43.5	15.6	13.8	18.2	14.4	16.0	15.8
1111	SD	16.5	14.4	44.5 13.7	43.5 13.0	53.2 18.0	49.1 15.9	49.5* 18.7	42.9 15.2	39.1	42.6
LIF KNO	M	38.0	33.5	71.2*	65.3	72.5*	68.1	67.0*	60.2	17.6 55.2	15.9 59.7
	SD	18.3	15.0	15.3	15.7	15.3	15.6	17.6	14.8	21.1	17.7
LIF APP	M	50.0	51.9	52.7	45.2	58.0	54.2	49.7	45.0	34.9	39.7
	SD	20.6	21.7	21.8	22.6	20.7	21.2	26.4	21.2	18.1	18.8
LIF INT	M	37.2	47.3	36.6	46.0*	52.9	49.1	47.9*	38.9	34.6	38.8
	SD	26.6	29.3	25.4	24.6	28.7	26.1	28.2	2.08	25.8	27.5
PHY KNO	M	57.5	53.4	61.3*	52.3	71.9*	64.1	65.0*	56.1	49.3	55.8
~~~	SD	24.4	25.2	25.5	23.9	-19.7	22.2	26.1	26.1	23.5	26.6
PHY APP	M	52.6*	47.8	61.1*	51.6	66.8*	57.5	58.9*	48.7	53.3	49.0
***************************************	SD	17.3	14.0	16.7	14.8	16.7	15.0	17.5	15.5	19.6	17.6
PHY INT	M	44.4	40.3	41.5	40.9	51.1	45.6	46.3*	38.7	37.3	41.1
ESP KNO	SD M	19.7	20.7	21.2	20.0	23.3	20.6	24.0	19.2	24.2	20.0
ESP KNO	SD SD	62.1 48.8	61.4 49.0	69.4	60.9	83.3	73.2	68.8	62.4	50.0	63.5
ESP APP	M	57.9*		46.5 61.5*	49.2 47.9	37.4		46.5	48.6	50.3	48.3
SOL ALI	SD	22.7	18.0	15.7	16.6	68.6 21.8	60.3 19.5	62.5*	51.2	47.2	48.4
NAT APP	M	52.0	50.6	63.3	62.5	$\frac{21.8}{67.5}$	68.1	23.5	20.5	20.4	20.7
	SD	28.1	28.9	27.4	23.3	26.4	29.0	31.0	27.6	44.1 31.3	53.1 31.7
VAT INT	M	44.3	44.6	50.4	45.1	55.4	52.6	53.9*	48.6	42.6	45.6
ı	SD	22.1	18.2	16.9	16.7	22.1	19.8	21.3	20.0	22.7	20.8

### Gender Comparisons in Gaza

As shown in Table 18, on 5 science test components out of the 18, the mean score of male students is statistically significantly higher than that of the female students at the  $\alpha = 0.05$  level. The 5 significantly discriminating variables are: Physical Sciences topic (PHY), Earth and Space Sciences topic (ESP), Application skill (APP), Application skill in Physical Sciences (PHY APP) and Integration skill in Nature of Science (NAT INT). Moreover, in terms of absolute difference between the means of the two groups, on 7 out of 18 variables, male students' scores are slightly higher than female students' scores. On two subscales namely, Life Sciences content and Integration skill, the mean scores of male and female students are the same, while on other two subscales, Nature of Science and Nature of Science Integration (NAT & NAT INT) mean scores of males and females are nearly the same. On the remaining two subscales, namely (LIF APP & LIF INT) female students' mean scores are slightly higher but only on one of them (LIF INT) the difference is statistically significant at the  $\alpha = 0.05$  level.

### Gender Comparisons in Lebanon

Means and Standard Deviations of male and female students' scores on each science test component including the total science test score are shown in Table 18. On 9 out of 18 variables including the total test score, differences are statistically significant in favor of male students at the  $\alpha=0.05$  level. Furthermore, in terms of absolute differences between means of the two groups, on 17 out of 18 variables, male students' scores are slightly higher than female students' scores. Similarly, as has been the case in Gaza, scores of female students in Life Sciences Integration were statistically significantly higher than those of their male cohorts at the 0.05 Alpha level.

## Gender Comparisons in Syria

Upon comparing male and female means in the third column from the right in Table 18, on 9 out of 18 variables including the total Science Test score, differences were statistically significant in favor of male students at the  $\alpha=0.05$  level. Female students in Syria did not perform significantly better than their male cohorts on any of the 18 variables. However, in terms of absolute difference between means of the two groups, male students slightly outscored females on 17 out of 18 variables.

#### Gender Comparisons in West Bank and Jordan

The Last two columns from the left in Table 18 represent average percent correct score on each component of the Science Test, its Standard Deviation and the statistical significance of the difference in the performance of male and female students in Jordan and West Bank as reported in an earlier study (Ahlawat et al., 1992).

In Jordan, on 14 out of 18 variables including the total test score, differences were statistically significant in favor of male students at the  $\alpha = 0.05$ .

However, in the West Bank, neither on total Science Test score nor on any of the 17 subtest scores did the difference between male and female scores reach statistical significance at the 0.05 Alpha level.

Statistical analyses of the data confirm the fact that in Gaza and West Bank the male/female achievement profile has more similarity in the sense that mean scores of male and female students tend to deviate from the norm found in Jordan, Lebanon and Syria where differences were statistically significant in favor of male students. A plausible explanation seems to be that in both Gaza and West Bank students' science performance, irrespective of the sex of the students, is across-the-board so poor that the floor effect has probably masked all the potential differences that could have emerged due to sex of the students.

# II Comparative Science Achievement Between Male and Female Students in the Entire Sample

Table 19 presents average percent correct score (M), Standard Deviation (SD) and Standard Error of mean (SE) between male and female student achievement in the entire sample covering Gaza, Lebanon and Syria.

Upon examining the differences between pairs of means and their statistical significance at  $\alpha = 0.05$ , it is quite evident that on 11 out of 18 components including the total Science Test score male students outscored the females. Also, in terms of absolute difference in means on each of the Science Test components, on <u>one</u> out of 18 variables, namely Life Sciences Integration female students scored slightly better than the male students; while on the remaining 17 variables including the total Science Test score, male students outscored the females.

Table (19)

Average Percent Correct (M), Standard Deviation (SD) and

Standard Error (SE) of Mean and Statistical Significance of the Difference Between

Male and Female Student Achievement in Total Sample of Gaza, Lebanon & Syria

Variable	STAT	Male N=263	Female N=266
Total Science	M	59.0	54.1*
	SD	14.5	12.3
(72)	SE	0.9	0.8
LIF	M	61.4	58.5*
	SD	15.8	13.7
(25)	SE	1.0	0.8
PHY	M	57.0	50.4*
	\$D	16.0	13.7
(26)		1.0	0.8
ESP	M	64.5	54.6*
	SD	20.9	19.2
(09)		1.3	1.2
NAT (C)	M	54.2	52.3
	ŞD	20.3	18.7
(12)	SE	1.3	1.1
KNOW	M	67.2	62.3*
	SD	16.7	15.5
(21)	SE	1.0	1.0
APP	M	60.3	53.6*
	SD	15.9	13.5
(32)	SE_	1.0	0.8 45.8
INT	M	47.8	
	SD	17.2	14.9
(19)	SE	1.1	0.9
LIF KNOW	M	67.4	63.2*
4.0	SD	17.6	16.3 1.0
(16)	SE	1.1 54.1	51.2
LIF APP	M		21.9
(00)	SD	21.2 1.3	1.3
(06)	SE		47.7
LIF INNT	M	43.9	26.8
(00)	SD	28.3	1.6
(03)	SE	1.7	57.6*
PHY KNO	M	64.6	24.2
<b>6</b> 0.0	. SD	23.6	1.4
(04)	SE	1.5	52.8*
PHY APP	M	60.8	15.1
44.45	SD	17.9	0.9
(14)		1.1	42.7*
PHY INT	M	46.6 22.0	20.6
<b>/00</b>	SD	22.0	1.2
(08)		1.4	66.2
ESP KNO	M	73.0	47.4
	SD	44.5 2.7	2.9
(01)			33.1*
ESP APP	M	63.4	
	SD	21.3 1.3	19.3 1.2
(08)			60.9
NAT APP	M	61.4	
40.45	SD	27.9	28.5
(04)		1.7	1.8
NAT INT	M	50.5	48.1
	SD	21.5	18.9
(08)	SE	1.3	1.2

<sup>\*</sup> Significant Difference at  $\alpha = 0.05$ .

# AREAS OF WEAKNESS AND STRENGTH

The use of objective measures may do much to assist educationists to make their planning more scientific and teachers to make their teaching more effective. Information obtained by means of the standardised tests should be used to guide the efforts of the teachers and administrators to know the areas of strength and weakness in order to target their limited time and resources. On the other hand, the main issue in the interpretation of test findings is to discern their value for future educational treatment as the test results are partly diagnostic by themselves. In this part, we shall highlight some selected test findings in order to emphasise the need and facilitate the plans for appropriate remedial measures.

# Areas of Weakness and Strengh in Mathematics

Table 20 below represents the test findings in terms of the number of test items with success rates of less than 30%, above 30% but not greater than 50%, and over 50%.

This is to say that we have classified all the test items measuring each of the three cognitive skills into three categories with respect to the range of difficulty index of each item. The item that is passed (correctly answered) by less than 30% of the students in a given sample belongs to the "< 30%" category, the high difficulty range. The item that is passed by 30% through 50% of the students is assigned above average difficulty range "30-50" - category. The items that are passed by more than 50% of the given sample "> 50" - category, the more acceptable difficulty range.

Table (20)

Number of Test Items Under Success Percentages

	CU (25 Questions)			(27	PK Questic	ns)	PS (24 Questions)			
% Passing	< 30	30 - 50	> 50	< 30	30 - 50	> 50	< 30	30 - 50	> 50	
Gaza	7	9	9	14	8	5	12	. 6	6	
Lebanon	5	9	11	11	9	7	9	9	6	
Syria	6	6	13	9	11	7	10	6	8	
All Sample	4	8	13	9	12	6	11	6	7	

Each row of Table 20 represents the number of items from each of the three cognitive skills falling in each of the three difficulty ranges according

to percentage-range of students from a given field sample answering the items correctly.

To illustrate, the number 7 in the first cell of the matrix of numbers, means that 7 out of the 25 on Conceptual Understanding skill items (i.e., 28% of the CU items) were answered correctly by less than 30% of the Gaza sample. In other words, 28% of the 25 items measuring the conceptual understanding of Math concepts were found very difficult by the 8th grade students in the Gaza field.

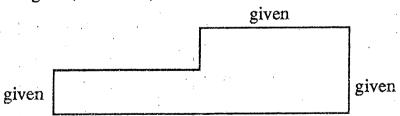
Generally speaking, it is important to note that, in all three fields, the worst performance was in the area of Procedural Knowledge, though the success percentages on the other two skills were not pleasing either.

All that has been described so far, was aimed to provide a general representation of the magnitude and scope of weaknesses in Math. However, upon looking at Tables 21, 22, 23, and 24, which represent the percentage of correct answers on test items in each part of the test, it can be clearly seen that the 8th grade students are particularly weak in the content and skills specifically tapped by the following set of items:

- 9, 11, 12, 14, 19 9, 15, 17, 19
- 4, 18
- 6, 15

Table 21, shows an evident weakness in the performance of students in each field on items 9, 11, 12, 14, and 19. For instance, only 17% of the entire sample could answer item 9 of Part I, correctly.

Q 9 requires students to compute the perimeter of an irregular rectangle of the following shape in centimetres of which the length and width are given, as shown, in centimetres.



This test item belongs to the combination Measurement-Problem Solving. Only 17% students in the entire sample answered this item correctly. It measures the students' ability to apply knowledge in solving simple measurement problems. Weakness may be attributed to the assumption that the students could not give the relation between the two

opposite sides of the rectangle when one of them is broken into two line segments.

Q 11 and Q 12 used a two-dimensional graph made of equal-sized smaller rectangles in which the horizontal axis represents the time and the vertical axis represents the distance travelled in kilometres. Students are required to read from the graph the distance travelled in a given time and the duration of the time during which the traveller stopped.

Table (21)
Percentage of 8th Grade Students Passing Each Item in Part I of the Math Test in Each Field and the Entire Sample

		Gaza		Lebanon				Syria		Entire Sample			
Item		Female	Total	Male	Female		Male	Female	Total	Male	Female	Total	
No.	N=88	N=87	N=157	N=63	N=68	N=131	N=111	N=114	N=225		N=269	N=531	
1	58	62	60	75	47	60	74	78	76	69	65	67	
2	51 .	51	51	60	50	.55	58	56	57	56	53	54	
3	64	55	59	87	74	80	83	64	. 74	77	64	71	
4	30	20	25	48	26	37	.44.	26	35	40	24	32	
5	53	53	53	51	50	50	76	44	60	62	48	55	
6	42	69	55	68	65	66	68	57	62	59	63	61	
7	25	29	27	30	24	27	35	27	31	31	27	29	
8	19	32	26	29	26	27	39	35	37	30	32	31	
9	15	08	11 <sup>.</sup>	24	09	16	33	12	23	25	10	17	
10	39	39	39	57	38	47	51	42	47	48	40	44	
11	0.5	10	07	24	04	14	19	08	13	15	08	11	
12	20	07	14	37	15	26	29	13	21	28	12	20	
13	47	43	45	73	50	61	71	57	64	63	51	57	
14	02	02	02	08	04	06	. 08	04	06	06	04	05	
15	39	34	37	52	47	50	. 57	55	56	50	46	48.	
16	22	26	24	32	16	24	42	37	40	33	28	31	
17	24	33	29	14	22	18	22	25	24	21	27	24	
18	32	20	26	37	34	35	38	32	35	35	- 29	32	
.19	10	15	13	08	04	06	10	15	12	10	12	11	

These two test items belong to the Statistics Procedural Knowledge combination. The students of 8th grade are not accustomed to solve problems related to line graphs as line graphs are not emphasised in the 8th

grade, graphs emphasized in the preceding grades though include BAR and PIE graphs. Weakness in this area, therefore, could be attributed to lack of curricular emphasis.

As for Q 14 it presents a figure representing the cross-sections of a pyramid made of small cubes. Students are required to figure out the number of cubes needed to build it.

This test item belongs to the combination Algebra-Problem Solving. It is related to visual thinking. Students in the countries under study are not exposed to this type of problems.

In Q 19 students are asked to compute the total surface area of the diagram of a regular cube in cubic centimetres. The identical magnitude of the length, width and heights of the cube is clearly shown in the cubic diagram.

This item belongs to the combination Measurement Procedural Knowledge. The success percentage on this item for the entire sample was only 11%. Such an item measures the students' knowledge of formulas and ability to apply them. Weakness may be attributed to the weakness of textbooks in emphasising measurement content. To improve students' performance in this area they have to be trained to differentiate between area and volume through building the solid figure from 6 squared cardboard pieces.

Table 22 Clearly shows weak performance according to the criterion of less than 30% correct response in the whole sample on items 9,15,17, and 19.

Q 9 belongs to the combination of Data Conceptual Understanding. It measures the students' understanding of the concept of probability in addition to that of the comparison between rational numbers. This topic is not included in the syllabuses of the 8th grade and below. Weakness is, therefore, expected.

Table (22)

Percentage of 8th Grade Students Passing Each Item in Part II of the Math Test in Each Field and the Entire Sample

	Gaza			Lebanon			Ι	Syria		Entire Sample		
Iten No.	i Male N=88	Female N=87	Total N=175	Male N=63	Female N=68	Total N=131	Male N=111	Female N=114	Total N=225	Male N=262	Female N=269	Total N=53
1	61	70 ~	66	79	71	75	91	72	** 81	78	71	75
2	69	71	70	65	60	63	. 77	71	74	72	. 68	70
3	45	44	45	48	47	47	60	53	56	52	48	50
4	47	57	52	65	62	63	77	70	73	64	<sup>′</sup> 64	64
5.	51	62	57	67	59	63	- 66	68	67	61	64	62
6	59	61	60	81	65	73	84	73	78	75	67	71
7	57	72	65	75	63	69	76	69	72	69	69	69
8	32	38	35	46	31	38	48	32	40	42	33	38
9	14	21	17	22	07	15	33	18	26	24	16	20
10	33	18	26	54	24	39	50-	39	44	45	28	37
11	68	83	75	71	54	63	. 86	84	85	76	76	76
12	30	28	29	44	26	35	38	26	32	37	27	32
13	55	61	58	76	66	71	66	67	66	65	65	65
14	45	51	48	56	43	49	59	61	60	54	53	53
15	17	20	18	37	28	32	28	21	24	26	22	24
16	26	33	30	41	35	38	43	42	43	37	38	37
17	16	23	19	41	25	33	25	16	20	26	20	23
18	26	-33	30	. 43	26	34 :	26	24	25	30	28	29
19	. 22	17	19	38	28	33	23	22	23	26	22	24

As for Q15 it belongs to the Measurement-Problem Solving combination. The success percentage on this item is 24% for the entire sample. It measures students' ability to apply knowledge in solving problems. Weakness may be attributed to the fact that neither the teachers nor the textbooks train the students to solve problems by trail and error. Part of the weakness may be attributed also to the general weakness in measurement.

Q 17 belongs to the area of Geometry-Problem Solving. On this item the success percentage for the entire sample is 23. It measures the students' spatial and visual ability in solving geometrical problems. This type of skill is emphasized neither by the curriculum nor by the teachers. Because of this, weakness is expected. Students' achievement may be improved by training the teachers in the preparation and implementation of exercises effective in improving students' visual and spatial ability.

Q 19 belongs to the Number-Conceptual Understanding combination. The success rate on this item for the entire sample is 24%. It measures the students' understanding of concepts and their application in solving problems and giving the right answer. This topic is in the 6th grade textbooks and emphasized in upper grades when teaching operations on fractions. This kind of weakness is not expected and may be attributed to the fact that many students forget formulas and procedures.

Table 23 gives a clear indication of weakness in students' performance on items 4,18 as the percentage of correct answers for the total sample was less than 20.

Table (23)

Percentage of 8th Grade Students Passing Each Item in Part III of the Math Test in Each Field and the Entire Sample

			Gaza			Lebano	n		Syria		En	Entire San	
	Item No.	Male N=88	Female N=87	Total N=175	Male N=63	Female N=68	Total N=131	Male N=111	Female N=114	Total	Male	Female N=269	Total N=531
ŀ	1	61	70	66	N=0.3 56	53	N=131 54	73	71	N=225	N=262 65	66	66 N=551
											03		
1	2	62	79	71	76	75.	76	67	67	67	68	73	70
	3	65	64	65	75	69	. 72	81	83	82	74	74	74
	4	09	07	08	17	06	11	23	10	16	17	08	12
	5	35	36	35	70	59	64	50	62	56	50	53	51
	6	31	26	29	30	19	24	35	35	35	32	28	30
1	7	27	31	29	38	22	30	41	34	37	35	30	33
	8	40	41	41	27	38	33	- 56	43	49	44	41	42
	9	.37	. 52	45	52	31	41	51	54	53	47	48	47
	10	31	24	27	54	29	41	49	44	46	44	34	39
	11	32	38	35	57	31	44	43	38	40	43	36	39
	12	60	67	63	44 .	40	42	40	. 54	47	48	55	51
	13	30	18	24	38	07	22	27	18	22	31	. 15	29
ŀ	14	23	33	28	62	29	45	48	40	44	43	35	39
	15	35	31	33	37	19	28	33	25	29	35	26	30
	16	23	28	25	22	12	17	30	21	25	26	21	23
	17	22	46	34	25	40	33	36	34	35	29	39	34
	18	17	15	16	08	19	14	12	04	08	13	12	12
	19	32	29	30	22	15	18	18	18	18	24	21	22

Q 4 in Part III of the Math Test belongs to the combination of Measurement-Conceptual Understanding. The success rate on this item for the whole sample is 12%. It measures the students' understanding of concepts and how should this understanding lead to the correct answer. Weakness may be attributed to the lack of understanding of the concept of length. They are not aware of the fact that the number of units increases as the length of the unit of measurement becomes shorter and vice versa. This main concept must be emphasized thoroughly as the teacher begins teaching the students the concept of length and measurement. In this case, estimation of the measure is appreciated.

Q 18 belongs to the Measurement-Problem Solving combination. The success rate on this item is 12% for the entire sample. The item is designed to measure the students' ability to apply knowledge in solving problems.

Weakness may be attributed to the students' inability to know that the four closed areas at the four corners of the square are equal and to their weakness in dealing with decimal numbers. Achievement may be improved if the students solve a set of direct exercises on the four operations on the different types of numbers.

Table 24 shows that items 6 and 15 of Part IV have been answered correctly by less than 20% of the entire sample.

As for Q 6, this test item belongs to the Number-Conceptual Understanding combination. The success rate on this item of the entire sample (barely 6%) is simply deplorable. Weakness may be attributed to the misconception of decimal fractions, and the value of a number in a given decimal place. Such comparisons must be thoroughly emphasized when teaching decimal fractions.

Q 15 belongs to Number-Procedural Knowledge. The success rate on this item is 19%. It is designed to measure the students' ability to do calculations. Weakness may be attributed to the students' inability to perform the division operation of a whole number by a rational number. The students learn this operation in the 5th and 6th grades.

Table (24)

Percentage of 8th Grade Students Passing Each Item in Part IV of the Math Test in Each Field and the Entire Sample

	T	Gaza			Lebano	n		Syria		Entire Sample			
Item	Male	Female N=87	Total	Male	Female N=68	Total .	Male	Female N=114	Total	Male N=262	Female N=269	Total N=531	
No. 1	N=88 55	62	N=175 58	N=63 62	65	N=131 63	N=111 78	77	N=225 78	66 66	69	68	
1				l							39	49	
2	49	36	42	67	40 .	53	64	42	53	60	39 .	47	
3	39	40	39	22	51	37	32	40	36	32	43	37	
4	19	32	26	35	12	23	29	12	20	27	19	23	
5`	50	72	61	76	63	69	78	59	68	68	64	66	
6	12	。01	07	. 08	06	1.07	06	02	04	09	03	06	
. 7	41	32	37	40	24	<sup>-</sup> 31	34	18	26	38	24	31	
8	35	40	38	49	31	40	52	39 <sub>‡</sub>	.46	46	38	42	
9	39	.66	52	48	41	44	77	72	74	57	62	60	
10	43	38	41	70	57	63	65	58	61	59	51	55	
11	37	43	40	75	49	61	71	62	67	61	52	56	
12	19	28	23	46	46	46	38	- 43	40	34	39	36	
13	23	23	23	62	28	45	54	32	43	45	38	37	
14	18	22	- 20	14	28	21	30	26	28	22	25	24	
15	16	17	17	19	06	12	29	.19	24	22	15	19	
16	27	36	31	33	22	27	48	36	42	37	32	35	
17	35	23	29	21	15	18	22	15	18	26	17	22	
18	27	37	32	17	24	21	22	13	17	23	23	23	
19	26	22	24	29	16	22	22	08	15	25	14	20	

To sum up, performance on the IAEP II Math Test is evidently the weakest in Measurement topic and Procedural Knowledge skills. Improving the current situation entails exerting more efforts and emphasis in teaching this topic and the procedural knowledge skills involved. Weakness in the areas of content and skills, not emphasized in the curriculum and not covered by the students, is understandable, but weakness in the content area that constitutes a major part of the syllabuses is hard to be justified.

Teachers of the first six grades should be well-advised to emphasize practice on the four operations on the different types of numbers in addition to some basic ideas on factorization and divisibility.

Students in the 7th and 8th grades would be better-off in a problemsolving classroom environment where teachers could give them more opportunities to solve exercises and problems using equations in the area of measurement.

# Areas of Weakness and Strength in Science

Table 25 presents the number of Science Test items answered correctly by less than 30% of the students in each field classified into the categories defined by each combination of content and skill. Table 26 presents the items that were answered correctly by more than 30% but less than 50% of the students, in the same fashion as Table 25.

Table 25 lists the number of items, from each of the 10 content-skill areas represented in the Science Test, that were found very difficult by the students in each *UNRWA* field. An item that was answered correctly by less than 30% of the given sample of students was classified as very difficult for that field.

Likewise, Table 26 presents the items of the above average range of difficulty, that is the items that were answered correctly by more than 29% but less than 50% of the students in each field sample.

Table (25)

Number of Test Item With Less Than 30% Correct Answers in Each Field Classified by Content and Skill Category Measured by the Science Test

Content and	Total No. of Items on IAEP			t Items wit Correct Ar	th less than iswers	
Skill	Science Test	Gaza	Lebanon	Syria	W.Bank	Jordan
Life Know	16	3			2	2
Life App	6	1	1	-	2	_
Life Int	3	-	1	-	1	1
Phy Know	4	1	1	-	1	_
Phy APP	14	2	1	_	2	1
Phy Int	8	2	2	1 .	-3	3
ESP Know	1	-	-	-	-	-
ESP App	8	2	3	-	2	1
Nat App	4	-		. –	_	' <del></del>
Nat Int	8	1	2	· -	-	_

Adding up the items under Gaza column of Table 25 clearly shows that in Gaza 12 items out of 72 (about 17% of the Science Test items) were answered correctly by less than 30% of the students. In Lebanon and Syria the number of such items was 11 and 1 respectively.

Table (26)

Number of Test Items in the Range of 30% - 50% Correct Answers in Each Field Classified by Content and Skill Category Measured by the Science Test

Content and	Total No. of Items on IAEP		No. of Test (30% - 50	Items in the (%) Correction	he Range o t Answers	
Skill	Science Test	Gaza	Lebanon		W.Bank	Jordan
Life Know Life App Life Int	16 6 3	3 2 2	1 2 1	2 3 2	2 3 2	2 4 1
Phy Know Phy APP Phy Int	14 14 8	1 7 -	2 4 4	2 5 2	1 6 2	6 2
ESP Know ESP App	1 8	2	1	2	2	3
Nat App Nat Int	4 8	1 4	1 3	1 4	2 4	1 4

Table 26 shows that 22 out of 72 test items (about 31% of all the test items) were answered correctly by 30% through 50% of the students in Gaza, while the number of such items was 19 (26%) in Lebanon and 23 (32%) in Syria.

2016 230

The description of Gaza students' achievement on the Science Test viewed from the content and skills perspective clearly indicates that their performance is the lowest in Life Sciences Integration skill and Physical Sciences Application and Integration skills. Next is Life Sciences and Earth and Space Sciences Application skill, Physical Sciences Knowledge skill, and Nature of Science Application and Integration skills.

In Lebanon weak performance is quite discernible in Life Sciences Integration, Physical Sciences Knowledge, and Integration skills, and Nature of Science Integration skill.

In the Syrian field, weak performance is quite obvious in Physical Sciences Knowledge, as well as, the Integration skill in both Life Sciences and Nature of Science.

In the West Bank, performance is drastic in Life Sciences Application and Integration skills and Physical Sciences Knowledge, application and Integration skills. Moreover, weak performance is quite clear in the Application skill of Earth and Space Sciences and Application and Integration skills of Nature of Science.

Similarly, in Jordan, poor performance is clear in Life Sciences and Physical Sciences Application and Integration skills, next is Earth and Space Sciences Application skill and Nature of Science Integration skill, both of which stand on the same footing in terms of their weakness.

What has been presented with respect to achievement in Science so far, reflects the relative performance of students in different areas of content and skills in each field. What follows is an examination of male and female students' performance on individual items and identification of the specific areas of weakness defined by the more specific skill and content tapped by particular test items. In this respect, Tables 27, 28, 29, and 30 represent the percentage of students in each group answering each item correctly, respectively, in Part II, Part III, and Part IV of the Science Test.

It can be clearly seen from Table 27 that there is an obvious weakness in student performance on items 17, 18 and 14 of the Part I.

Item 17 measures the student's ability to design experiments in order to use her/his knowledge in Biological Sciences to come to a conclusion. In actual practice neither the teachers nor the curricula emphazise such skills. Consequently, the students have not been exposed to teaching / learning situations that develop experimentation skills in them.

Table (27)

Percentage of 8th Grade Students Passing Each Item in Part I of the Science Test in Each Field and the Entire Sample

		Gaza		<u> </u>	Lebanoi	n	T	Syria	·	Entire
Item No.	Male N=87	Female N=88	Total N=175	Male N=62	Female N=66	Total N=128	Male N=114	Female N=112	Total N=226	Sample N=529
1	74	73	. 73	90	73	81	82	62	72	74
2	61	64	62	74	64	69	67	69	68	66
3	47	37	42	47	42	45	53	40	46	45
4	79	82	81	71	56	63	82	81	81	77
5	70	58	64	58	56	57	73	68	70	65
6	60	. 68	64	44	52	48	51	59	55	56
7	<b>75</b> 7	80	77	66	74	70	82	79	81	77
8	87	78	83	68	61	64	75	79	77	76
9.	86	78	82	81 :	82	81	80	89	85	83
10	38	20	29	63	48	55	70	54	62	50
11	29	58	43	47	47	47	56	62	59	51
12	48	50	49	74	70	72	72	69	70	64
13	43	57	50	35	45	41	60	53	56	50
14	38	31	34	32	20	26	43	42	42	36
15	26	20	23	37	20	28	68	- 52	60	40
16	26	31	29	- 55	55	55	57	43	50	44
17	28	36	32	21	18	20	43	33	38	32
18	43	28	35	29	29	29	32	35	33	33

Further more, Q18, in the same vein of experimentation in science, presents a graphic record of experimental data and requires students to interpret the graph and draw valid conclusions. This item measures the ability of the students to integrate knowledge about the concept of boiling with that of their ability to read and interpret a heating curve. A reasonable explanation as to the cause of the weak performance (33% correct on this item from the total sample) is that the students knew that water boils at 100° yet they failed to integrate this knowledge about the concept of boiling with its representation on the graph. In fact, science textbooks, in general, lack this sort of graphical representations of the scientific concepts, as a result, the students fail to develop appropriate skills needed to interpret them. We all know that teachers faithfully stick to the textbooks.

Students' performance on part III of the Science Test, obviously, reflects a general weakness on concepts and skills tested by majority of the items. Items 11, 18 and 14 cause a major concern.

Item 11 in Part III is designed to measure students' ability to integrate understanding of the Nature of Science so as to be able to choose the correct experiment to test a hypothesis. 27 percent correct response in the entire sample is clear evidence of students' weakness in this area of knowledge and skill. The explanation given earlier for the previous items holds true here too, i.e., neither the teachers nor the science textbooks - in general - adequately present this type of skill.

Table (28)

Percentage of 8th Grade Students Passing Each Item in Part II of the Science Test in Each Field and the Entire Sample

					ach Pier				···	
		Gaza		<u> </u>	Lebano			Syria		Entire
Item No.		Female N=88	Total N=175	Male N=62	Female N=66	Total N=128	Male N=114	Female N=112	Total N=226	Sample N=529
l i	78	64	71	79	71	75	79	71	75	74
2	74	70	72	89	95	92.	93	91	92	85
3	62	61	. 62	69	61	65	83	73	78	70
. 4	62	60	61	73	79	76	73	. 81	. 77	71
5	78	69	74	84	- 79	81	92	84	88	82
6	77	87	82	76 -	59	67	82	77	80	· 78
7	69	75	72	85	79	82	88	88	88	81
. 8	55	51	53	71.	45	58	79	65	72	62
9	90	78	84	87	65	76	. 93	87	90	85
.10	62	61	62	77	62	70	85	80	83	73
11	78	82	80	95	95	95	96	99	98	91
12	45	45	45	63	50	<b>5</b> 6	66	44	55	. 52
13	45	40	42	79 :	50	64	89	92	90	68
14	67	65	66	85 .	70	77	82	78	80	75
. 15	49	58	54	79	68	73	66	79	72	66
16	37	50	43	50	70	60	61	73	67	58
17	61	70	66	65	71	68	57	69	63	65
18	41	33	37	47	44	45	53	57	55	47

Table (29)

Percentage of 8th Grade Students Passing Each Item in Part III of the Science Test in Each Field and the Entire Sample

		~ ~	ence re		· .					Y3-45
-		Gaza			Lebanor			Syria		Entire
Item	Male	Female N=88	Total N=175	Male N=62	Female N=66	Total N=128	Male N=114	Female N=112	Total N=226	Sample N=529
No.	N=87 87	81	84	89	64	76	96	83	. 89	84
2	67	53	60	84	73	78	81	75	78	72
3	56	52	54	63	35	48	64	41	53	52
4	40	35	38	32	32	32	59	46	53	43
5	47	48	47	45	38	41	61	48	54	49
6	52	48	50	66	52	59	74	59	66	59
. 7	44	43	43	58	36	47	42	.34	38	42
8	46	33	39	61	56	59	46	41	43	46
9	47	56	51	39	50	45	54	41	48	48
10	43	32	37	39	36	37	66	59	62	48
11	23	10	17	21	12	16	48	35	42	27
12	64	37	51	53	32	42	65	35	50	48
13	32	47	39	35	54	41	39	34	36	38
14	26	28	27	34	23	28	33	. 30	32	29
15	18	24	21	53	58	55	39	29	34	35
16	23	26	25	63	65	64	47	38	43	42
17	47	40	43	27	38	33	36	29	33	36
18	22	24	23	29	33	31	35	22	29	27

Item 18 of Part III measures students' ability to integrate knowledge of the concepts of physical and chemical changes that go with the processes of burning, rusting, rotting and melting. A 27 percent correct response from the total sample of 8th grade students in Gaza, Lebanon and Syria on such an item is quite astonishing. The Science curricula in all fields emphasize the concepts of physical and chemical changes. The science textbooks also include activities for the students to enable them to distinguish between both concepts. A possible explanation to the failure of the students to perform better on such an item is the failure of the teachers to create teaching / learning situations that stimulate students to use their scientific knowledge in explaining real life situations. Furthermore, the teachers seem to persist with their rather undesirable tendency to teach physical and chemical changes as isolated concepts without linking them with other concepts such as burning, rusting, rotting, or melting.

Table 30 also reveals a general weakness of performance on, at least, half of the items in Part IV, as indicated by the percentage of correct answers from the entire sample. Items 18, 11, 12 and 17 are good examples as they have the least percentage of correct answers.

Item 18 in Part IV of the Science Test measures the ability to apply knowledge of barometric pressure VS height on the given figure. No doubt that the students do know the factors which have an effect on atmospheric pressure, height being one of them, but they do not know how to apply this knowledge in a specific situation and eventually they couldn't find out at which location would the atmospheric pressure be the highest.

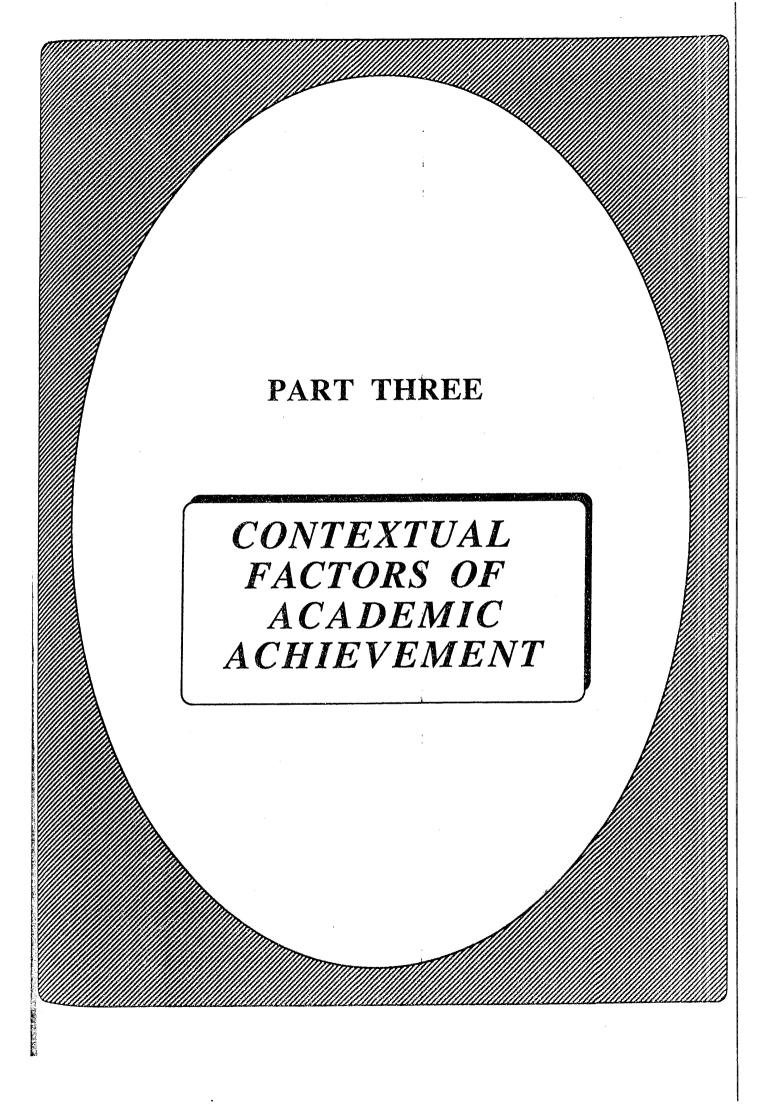
Table (30)

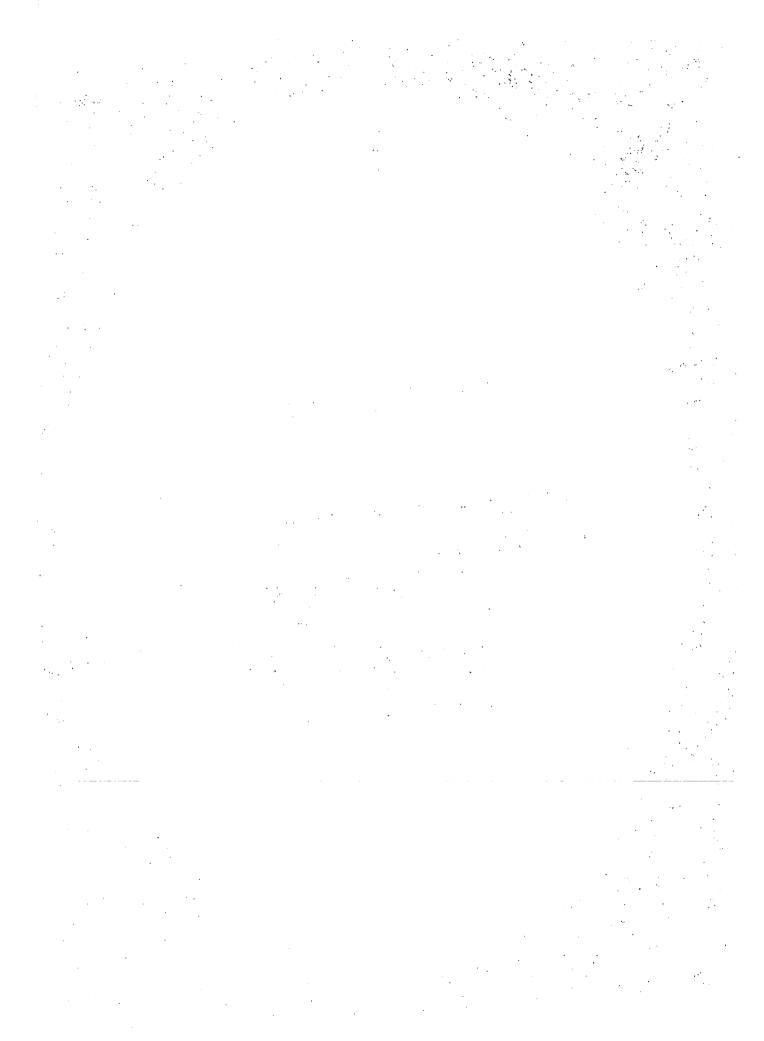
Percentage of 8th Grade Students Passing Each Item in Part IV of the Science Test in Each Field and the Entire Sample

		Gaza		T	Lebano	1	T	Syria		Entire
Iten		Female	Total	Male	Female	Total	Male	Female	Total	Sample
No		N=88	N=175	N=62	N=66	N=128	N=114	N=112	N=226	N=529
1	77	62	70	84	76	80	83	79	81	77
2	71	67	69	74	73	73	94	81	88	78
3	47	48	47	55	67	61	79	64	72	61
4	53	49	51	48	59	54	66	59	62	57
5	53	50	51	97	89	93	94	96	95	80
6	75	66	70	81	- 61	70	70	70	59	. 68
7	46	30	38	48	44	46	36	29	33	38
8	59	60	<b>5</b> 9.	44	47	45	66	62	64	58
9	38	36	37 <sup>*</sup> .	47	30	38	54	37	46	41
10	48	43	46	61	64	62	62	55	59	55
11	33	19	26	24	36	30	44	32	38	32
12	26	20	23	24	24	. 24	52	45	48	34
13	62	58	60	69	55	67	64	63	64	63
14	-32	41	37	42	32	37	. 38	52	45	40
15	62 -	33	47	37	6	21	49	48	49	42
16	32	42	37	69	58	63	67	60	63	55
17	30	24	27	44	14	28	48	45	46	36.
18	32	12	22	29	23	26	32	31	32	27

Item 12 of Part IV measures students' ability to integrate their knowledge about the concepts of liquid, solid and gas with that of the nature of the motion of molecules and decide the state of matter with molecules apart and having weak forces between them. A 34 percentage of correct answers from all respondents in the entire sample could be due to the lack of emphasis on integration skills on the part of the teachers, as well as, the absence of such an approach in the science curricula to deal with the phases of matter. Almost all science curricula in the basic cycle discuss phases of matter from a simple descriptive perspective without going into details as to the behaviour of molecules or the nature of forces binding these molecules in each of the three phases of matter.

To sum up, the weak performance of the 8th grade UNRWA students on the IAEP II Science Test raises the issue of educational accountability to answer the question: who or what is to be held accountable for making the students achieve the expected learning outcomes? Is it the students themselves? Their home life? Parents? Teachers? The school? Or the instructional program? The issue is quite controversial and it goes far beyond what happens in the classroom, and therefore it needs a thorough investigation in its own right.





÷ . . . . . .

# Chapter VI

# HOME BACKGROUND, CLASSROOM ACTIVITIES, STUDENT ATTITUDES AND ACHIEVEMENT

# The Family Background, Classroom Activities, And Attitudes Of The Math Sample Students

Home and family background variables such as family size, study environment, parental aspirations, parental interest and involvement, proxy measures of scioecnomic status, and psychological climate are often cited in educational research literature as correlates of students' academic achievement.

To study the relationship between these peripheral variables and student achievement in Mathematics, several statistical analyses were conducted including frequencies and linear correlation.

The 24 items of the student family background questionnaire are described and their status discussed on the basis of frequency and percentage of responses in each response category in the following sections of this report.

# Home language

The first question in the students' family background questionnaire sought information regarding the main language spoken at home. The question and the frequency and the percentage of responses in each category are given below in Table (M1).

Table (M1)
Q1. What language is usually spoken in your home?

Label	Value	Freq.	Percent
Missing	0	21	4.0
Missing Arabic	1	492	92.7
English	2	11	2.1
Other	3	7	1.3
Total		531	100

It is clear from the column headed "Percent" in the above table that 92.7% of the 8th grade students reported that Arabic was usually spoken in their homes while 2.1% reported English and 1.3% reported other languages.

Table 31 presents the linear correlation coefficients of home background, classroom activities and students' variables with Math achievement. Because language spoken in the home, (Q1) and whether Math is for boys or for girls, or for both (Q16) are strictly nominal variables in which ordering the response categories does not make sense, the rows representing variables Q1 and Q 16 in Table 32, therefore, are left blank. Nevertheless, it is of interest to know whether or not students' achievement in mathematics varies with respect to the language (Arabic, English or other languages) usually spoken in the home. Recognising the fact that Arabic is spoken in the homes of 93% of the 8th grade students in the entire Math Test sample of the three fields while, English is spoken in the homes of 2% of them, other languages in the homes of 1.3% of them and 4% of them did not respond; we compared the Math Test score means of the 4 groups via Oneway ANOVA (unequal cells).

The null hypotheses of the equality of the four group means could not be rejected at the  $\alpha = 0.05$  level of statistical significance.

#### Brothers and sisters

Number of brothers and sisters was taken as a measure for the family size. Information obtained from the Math sample is given below in table (M2).

Table (M2)

Q2. How many brothers and sisters do you have? Don't count your self.

Label	Value	Freq.	Percent
Missing	0	2	0.4
None	1	8	1.5
One	2	. 9	1.7
Two	3	36	6.8
Three	4	49	9.2
Four	5	59	11.1
Five	6	80	15.1
6 or more	7	288	54.2
Total		531	100

The "Percent" column of table M2 clearly shows that about 54% of the families have more than 6 children. This variable is one of the few reliable variables that are good predictors of academic achievement of the students.

Row 2 of Table 32 shows a consistently negative but non-significant linear relationship between Q2 and achievement in Math. Accumulation of more than 54% of the families in a single category (6 or more) has artificially curtailed the normal distribution and severely restricted the variance, and consequently, constrained the correlation coefficient from reaching its real value.

Table (31)

# Linear Correlation Coefficients of Home Background, Classroom Activities and Student Attitudes with Math Achievement Components in Gaza, Lebanon and Syria

L						A					
النسنا	.1777 **	2311 **	.2418 **	.1747 **	2329 **	.2390 **	.1544 **	2032**	.2554 **	My parents want me to do well in Math.	12,
	.1340 *	.0918	.0850	.1255 *	.0961	.0945	.0430	.0899	.1128	It is important to know some Math to get a good job.	11.
	.0465	.0089	0259	.0253	.0122	0271	.0034	.0162	.0095	Mathematics is useful in solving every day problems.	10.
	.0457	.0174	0108	.0171	0039	.0038	.0452	.0141	.0184	Do you ever use computer for school work?	
<u>Ł</u>	0563	0979	1015	0554	1062	0864	.0343	1032	0974	Do you ever use calculator in school?	-
	.1202 *	.1360 *	.0958	.1218 *	.1074	.1664 **	.0721	.0836	.1327 *	Do you have a calculator?	7.
(63)	.0936	.1154	.1239 *	.0727	.1190	.1482 *	.0969	.0909	.1246 *	How much time do you spend each day on homework?	5
1	.1245 *	.1236*	.1240 *	.1264	.1018	.1059	.0599	.1380 *	.1380 *	How often do you read on your own for fun?	.51
1	.1242 *	.1361 *	.1537 **	.1030	.1082	.2162 **	.0860	.1162	.1538 **	How many hours do you spend watching TV at home?	2 <del>1</del> .
L	.1569 **	.1822 **	.1668 **	.1593 **	.1462 *	.1780 **	.9000	.1746 **	.1897 **	How many books are there at your home?	
	0943	0788	0605	0996	0532	1107	.0363	0507	0862	How many brothers and sisters do you have?	2.
										What Language is usually spoken at your home?	i.
	Z	PK	CU	ALG	DAT	GEO	MEA	NUM	T.Math	Question	Q.No
•		-						-			

(63)

Table (31) Continued

:	7.		3	3	71	3 5	10	\$	17.	16.	5	1	14	Q.No
homework?	Do you get help at home to do your	Do you talk with others at home about Mathematics?	THE SPAIR TO HOMEWORK EVERY WEEK	quiz?	objects and instruments in Math lesson?	small groups during the Math class?	yourself during the Math class?	giving a Math lesson?	How often do you listen to the teacher	With which of the following statements	1 am good at Mathematics.	important as getting answer.		1
	1124	.1013	.1157	.0202	1227 *	1392 *	0304	2020	CSON		.2067 **	0/50	2036	T.Math
	0921	.1167	.1031	.0117	0872	1024	.0262	08/0			.1930 **	0/04	2340 **	NUM
1010	- 0181	0933	.0385	.0211	1045	0674	.0131	0925			.1601 **	0558	1872 **	MEA
CCOT	1055	.0539	.1288 *	.0103	-:0736	1499 *	.0343	0696			.1322 *	0868	2871 **	GEO
1611'-	1101	.0730	.0930	.0009	1184	1174	.0032	1095			.1228 *	0477	1781 **	DAT
-:1145		.0585	.0938	.0355	1285*	1303 *	.0545	0426			.2017 **	0401	2491 **	ALG
0985		.1110	.0727	.0257	1251 *	1461 *	.0283	- 0864			.1374 *	0882	2644 **	CU
1076		.0788	.1371 *	.0403	1284 *	1266 *	.0392	0883			.2274 **	0870	2804 **	PK
0940		.0871	.0892	.0341	0664	0994	.0084	0804	:		.1772 **	0147	2063 **	Z

#### Number of books at home

Question 3 of the Student Questionnaire asked about the number of books in the home, again a proxy indicator for home environment and socio-economic status. Frequency and percent of responses in each category are given in Table (M3).

Table (M3)

Q3. How many books are there at your home?

Label	Value	Freq.	Percent
Missing	0	2	0.4
0 - 10	1	67	12.6
11 - 24	2	151	28.4
25 - 100	3	160	30.1
> 100	4	151	28.4
Total		531	100

About 28% of the 8th grade students reported to have more than 100 books at home. Looking at Q 3 in Table 31 one can see that there is a significant correlation between the number of books in the home and the total score on the Math Test.

The number of books in the home evidently reflects the cultural background, academic interest and value system of the family; all these factors are known to contribute toward creating a conducive learning environment at home. Number of books at home being an indicator of the socio-economic status of the family is, usually, positively correlated with student achievement.

# Reading for fun

Question 5 inquires about students' habit of reading for self enjoyment. The response frequency and percentage in each category are given in Table (M4).

Table (M4)
Q5. How often do you read on your own for fun outside of school?

Label	Value	Freq.	Percent
Missing	0	5	0.9
Every day	4 :	117	22.0
1 - 2 a week	3	234	44.1
1 - 2 a month	2	87	16.4
Never	1	88	16.6
Total		531	100

Row 5 in Table 31 shows that reading for fun has a statistically significant positive correlation with the performance of the students on the Math Test.

This may be attributed to the assumption that reading for fun fosters academic interest and widens the conceptual horizons of the students.

#### Time spent on homework each day

Appropriate type of homework given to students in right doses is a potentially effective means of teaching. It can be used to encourage self-learning, enhance motivation, and foster positive attitude and subject-related self-concept among the students. Question 6 of the Part V of the Math Test enquired about the time students usually spent each day on homework in all school subjects. Table (M5) presents the students responses in each category.

Table (M5)

Q6. How much time do you usually spend each day on homework for all school subjects?

Label	Value	Freq.	Percent
Missing	0	5	0.9
Have no home work	1	5	0.9
Half hour or less	2 .	42	7.9
1 hour	3	115	21.7
2 hours 3 hours or more	4	182	34.3
	5	182	34.3
Total		531	100

It is clear from Table M5 that about 8% of the students spent half an hour for doing homework for all subjects each day. About 22% spent one hour and about 34% spent 2 hours, and also about 34% spent 3 hours or more on homework each day.

Looking at Q6 in Table 32, we find a positive significant correlation between time spent on doing homework and the performance of the students on the Math Test.

# Watching television at home on school days

There are two more questions in this questionnaire that are concerned with the amount of time students spent on particular activities. Question 4 asks about time spent watching TV at home on a school day and Q22 asks about time spent each week on mathematics home work. The percentage of responses to each individual item in each category are given in Table (M6).

Table (M6)
Q4 and Q5: Percentage of Response in Each Category

No.	Question	None	< 1	2 hrs	3 hrs	4 hrs	5 hrs	> 6
4.	Hoe many hours do you spend watching TV at home?	5,1	40.1	33.9	10.5	6.8	1.9	1.3
22.	Time spent for homework every week.	1,3	42.4	28.6	8.1	5.1	3.2	9.4

Row 4 in Table 31 shows a statistically significant positive relationship between the number of hours students spent watching TV at home and their performance on the Math Test.

Given that research findings from the industrialized countries indicate deleterious effect of TV watching on students' academic achievement, UNRWA fields' findings may seem contradictory, to some, even heretic. But, considering the moderation with which UNRWA 8th Grade students watch TV, the obtained results make perfect sense. It should be noted that 69% of the students spend 2 hours. or less on TV while 40% of all students spend less than one hour on TV. Three hours or less, including 5% who do not watch TV at all, accounts for 90% of the sample. In this sense, amount of TV watching within a well defined range of moderation (say, from 0 - < 3 hours) may provide necessary recreation and recharge the students for other work.

Moreover, if the TV programmes watched by students have some knowledge and or mind-expanding ingredient then TV watching in moderation may have positive effect on students' conceptual/ intellectual growth and enhance their innovative and creative capabilities.

Q 22, time spent every week on Math homework, again, has positive, significant but weak relationship with students' Math Test score.

# Computer / Calculator

In the age of technology, instructional technology has been developing rapidly and the concept of computer literacy has been gaining ground in many countries' education systems. The students' questionnaire included three computer / calculator related questions. Each question required "Yes" or "No" answer to these questions and to two other questions which are included in Table (M7).

Table (M7)
Percentage of Yes / No Responses to Q7, Q8, Q9, Q23, and Q24

Q.No	Questions	Per	cent
		Yes	No
7.	Do you have a calculator?	59.7	40:1
8.	Do you ever use calculator in school?	15.4	84.2
9.	Do you ever use computer for school work?	3.2	96.2
23.	Does anyone at home ever talk to you about what you are learning in Math Class?	78.3	19.6
24.	Does anyone at home help you with your Math homework?	46.5	51.4

Responses on Q7 show that 59.7% of the sample said "Yes" and 40.1% said "No". Possessing a calculator is one thing but using it in school is another. Students are not allowed to use calculators at schools except in Syria and at the 8th grade level only. This is why the response for using the calculator at school was 15.4%.

Looking at Q7 in Table 31, it can be seen that having a calculator has a significant positive correlation with achievement on the Math Test but using calculators at school has a negligible negative correlation with the total Math Test score.

As to family involvement and help in Math homework, questions numbered 23 and 24 in Table M7 tap the dimension of interest and support of the family toward learning Mathematics. Table M7 shows that 78.3% of the sample stated that someone at home enquired about what they learn in Mathematics class and 46.5% of the students admitted that they were helped by someone at home in their Math homework.

Looking at these two questions in Table 31, there is a positive but non-significant correlation between Q 23 and the students performance on the Math Test. The correlation coefficient between Q 24 and the Math Test score is negative but again statistically nonsignificant.

This is, perhaps, because only the weaker students need help with their Math homework.

# Mathematics for girls or for boys or for both equally

Question 16 explored students' opinion about whether Mathematics was more for boys or more for girls or about equally for both. Distribution of students' responses is presented in Table (M8).

Table (M8)

Q16. With which of the following statements about Mathematics do you agree?

Label	Value	Freq.	Percent
Missing	0	10	1.9
Math is more for boys than for girls.	1	86	16.2
Math is more for girls than for boys.	2	57	10.7
Math is for boys and girls about equally.	3	378	71.2
Total	7	531	100

It is clearly seen from the above table that 71.2% of the sample agreed that Math is for both boys and girls, 16.2% agreed that math is more for boys, and 10.7% agreed that math is more for girls.

Because this is a pure nominal, categorical, variable, Oneway ANOVA was conducted to investigate if there were any differences in the Math Test scores of the three groups of students each holding a specific opinion about gender stereotype and Mathematics. The F-test was not found significant at the  $\alpha$ = 0.05 level.

#### Attitude toward mathematics

Questions 10 through 15 of the Students' Questionnaire composed the attitude subscale. Each item was rated as "Strongly Agree" to "Strongly" Disagree" on the 5 - point rating scale. Distribution of percentage of responses on each item is given in table (M9).

Table (M9)

Percentage of Responses in Each Category of the Items 10-15

No.	Question	SA	A	NE	D	SD
10.	Math is useful in solving every day problems.	14.1	47.6	20.7	10.9	6.2
11.	It is important to know some Math to get a good job.	32.0	54.8	7.3	1.3	2.8
12.	My parents want me to do well in Math.	67.4	27.3	1.9	1.1	1.7
13.	Learning Math is mostly memorising	14.3	31.1	14.5	24.3	15.3
14.	Knowing how to solve a problem is as important as getting the answer.	40.7	44.8	5.8	.5.6	2.4
15.	I am good at Math	29.4	48.6	13.0	4.0	3.4

From table M9 it can be seen that 61.7% of the sample agreed or strongly agreed with the opinion that Math is useful in solving everyday problems (Q10), 20.7% were neutral and 17.1% disagreed or strongly disagreed with the statement.

As for Q11, 86.8% of the sample agreed or strongly agreed with the statement that Math is important for getting a good job.

In response to Q12, 94.7% of the sample agreed or strongly agreed with the statement "My parents want me to do well in Math". It is well-known that students' attitude toward a subject is positively correlated with their achievement in that subject. The statistically significant positive correlation of Questions 10, 11, and 12 with performance on the Math Test supports this assertion.

As for Q13, 45.4% of the sample agreed or strongly agreed that Math is mostly memorising, 14.5% were neutral and 39.6% strongly disagreed or disagreed with the statement. Highly significant negative correlation of Q13 with students' Math Test scores in Table 32 indicates that the brighter the student the stronger the disagreement with the statement because the brighter students know that there is more to mathematics learning than merely memorising.

As for Q14, 85.5% of the sample agreed that knowing how to solve a problem is as important as getting the answer. Irrespective of this, the correlation between this statement and the performance of the students on the Math Test was consistently negative but statistically non-significant.

As for Q15 (I am good at mathematics), 78% of the sample stated that they were good at math. Again, consistently, highly significant positive relationship between Q15 and Math achievement in Table 31 reconfirms the commonly held assumption of positive relationship between subject-specific self-concept and achievement in the subject.

#### Practices and activities in mathematics classroom

There has been growing awareness among educators of the findings of educational research that effective use of various classroom activities and testing practices can enhance students learning and achievement. Activities such as problem solving in small groups, doing exercises by oneself, working with objects and instruments have been found conducive to learning and hence these activities are considered indicators of quality teaching. In the same vein, tests and quizzes are not merely tools of assessment, they have been used by good teachers in numerous classrooms throughout the world as efficient means of effective instruction. Questions 17 through 21 belonging to this group explored the frequency of each classroom activity. Percentage of students' responses in each category of each item is given in Table (M10). Occurrence of each type of activity is reported in five discrete categories, ranging from "Never" through "Every day".

Table (M10)

Percentage of Responses in Each Category of Items 17 through 21

<u> </u>			Res	onse per	Response percent									
No.	Question	Never	Less than once a week	Once a week	2/3 times a week	Every day								
17.	How often do you listen to the teacher give a Math lesson?	0.9	0.8	1.9	37.1	57.1								
18.	How often do you do Math exercises by yourself in the class?	1.7	3.2	4.7	55.6	33.5								
19.	How often do you solve questions in small groups in the Math class?	11.7	8.7	17.3	43.9	17.1								
20.	How often do you use Mathematics forms and	28.8	14.1	14.7	33.9	7.2								
21.	instruments in math class?  How often do you take a math test or quiz?	4.1	30.5	30.9	28.8	4.4								

From Table (M10), it is clear that 57% of the sample listened to the teacher give a Math lesson every day. The effect of this listening on the students' performance on the math test components is shown in Table 31. The correlation is consistently negative but statistically nonsignificant.

Question 18, which is concerned with students doing mathematics exercises by themselves during Mathematics class, is shown, from Table 31 to be consistently positively correlated with performance on the Math Test, but the correlation did not reach statistical significance at the  $\alpha$ = 0.05 level.

The negative and positive trends in relationship respectively of Q17 and Q18 with students' performance in Mathematics (despite the statistical nonsignificance) tell a significant didactic tale. That is, listening to the teacher give a Math lesson indicates the practice of traditional lecture method of teaching on the part of the teacher and passive learning on the part of the students; whereas, doing math exercises by themselves is indicative of active self-learning approach.

As for Q19, 70% of the sample reported that they were solving math problems in small groups. The correlation coefficient between this statement and the students' performance on the Math Test is negative and significant. This may mean that teachers are not using the small-group-problem solving method effectively in their classes.

As for Q20, 41.1% of the sample stated that they use geometrical instruments in math lessons quite often (2/3 times a week and every day). This statement is negatively and significantly correlated to the pupils performance on the Math Test.

Question 21 is related to the number of tests or quizzes given to the students. Tests are disliked by both teachers and students. They have the potential of being devastating if misused and beneficial if used properly. Table 31 shows that, Q21 is positively correlated to the pupils performance on the Math Test, but the correlation is not significant. If tests and quizzes were used properly as a means of instruction then one would expect a stronger positive relationship between frequency of quizzing and test performance of students.

# Home Background, Classroom Activities, Student Attitudes and Achievement in Science

As it was mentioned earlier there are factors which have an impact on students' learning other than school variables. Family and out-of-school activities play important roles in promoting in-school success. Some aspects of home-life, such as number of books in the home and family size are often cited as indicators of social and economic advantage and in the *IAEPII* study these variables have been related to science achievement in predictable ways. Perhaps parental involvement, which can influence a child's academic performance regardless of a family's socio-economic status, is another variable that is worth consideration.

What students do with their time after school seems to be another important factor that affects academic performance. In many *IAEPII* study populations, performance on the Science Test was positively associated with amount of time spent on leisure reading and homework in all school subjects and negatively with amount of time spent watching television (Lapointe et al., Learning Science, p. 74).

The way the students perceive the nature, usefulness and characteristics of science and their attitude toward the subject constitute additional variables that influence students achievement.

To study the relationship between every one of these factors and student achievement in Science several statistical analyses were conducted including frequencies, Oneway Analysis of Variance, and linear correlation. The 22 items of the students' family background questionnaire are described on the basis of frequency and percentage of responses in each response category in the following sections of the report. First, we will present the status of 8th grade students in the entire sample as a whole on each variable, then we will present and discuss the linear relationship of each variable as well as of subscales composed by groups of items with student achievement in total science test and its various subtests measuring students' achievement in each content area and each cognitive process skill

# Home language

The first question in the students' family background questionnaire sought information regarding the main language spoken at home. The question and frequency and percentage of responses in each category are given in Table (S1).

Table (S1)
Frequency, and Percentage of Responses in Each Language Category

Label	Value	Freq.	%
Arabic	1	500	94.5
English Other	2	7	1.3
Other	3	9	1.7
Missing	0 ·	13	2.5
Total		529	100

It is clear from column headed "%" in Table (S1) that around 95% of eighth grade students taking the Science Test reported that Arabic was usually spoken in their homes while 1.3% of them reported English and 1.7% reported other languages besides Arabic or English and the rest did not report anything.

#### Brothers and sisters

Number of brothers and sisters was taken as a measure of family size. The information obtained from the science sample is shown in the following Table (S2).

Table (S2)
Frequency and Percentage of Responses in Each Response Category

	Label	Value	Freq.	%
None		1	8	1.5
One	•	2	9	1.7
Two		3	23	4.3
Three		4	51	9.6
Four		5	70	13.2
Five	•	6	82	15.5
Six or More	Α.	7	283	53.5
Missing		0	3	0.6
Total			529	100

As can be computed from the Percent Column in Table (S2), only 7.5% of the students have 2 or less brothers and /or sisters and 9.6% have 3 brothers and sisters, while 53.5% of the students reported having 6 or more brothers and /or sisters.

In order to understand the influence of home characteristics and classroom practices, we study their status in the context of their relationship with the Science components.

Table 32 presents the linear correlation coefficients of each question of the science student questionnaire with all the components of the Science Achievement Test.

Row 2 of Table 32 shows that family size consistently tends to be negatively correlated with achievement in Science. Because of the concentration of about 54% of the cases in the 7th category (six or more) the distribution of scores has been drastically curtailed. This has artificially restricted the correlation from reaching its true value. In *IAEPII* study (Learning Science, Lapointe, 1992 p.65) in 15 countries the negative correlation was statistically significant.

Table (32)

Linear Correlation Coefficients of Home Background, Classroom Activities and

Student Attitudes with Science Achievement Components
in Gaza, Lebanon and Syria

	11.	10.	9.	œ	7.	6	Ċ.	.4	'n	2.	1.	Q.No
	Facts and ideas about science cannot be questioned or changed.	Scientists discover new facts and ideas about science by making observations and doing experiments	My parents are interested in Science.	It is important to know some science in order to get job.	Much of what you learn in Science class is useful in everyday life.	How much time do you spend each day on homework for all school subjects?	How often do you read on your own for fun?	How many hours do you spend watching TV at home on school day?	How many books are there at your home?	How many brothers and sisters do you have?	What language is usually spoken at your home?	Question
	.03	.15*	08	.08	.14*	.12*	.04	.17 **	.10	05		T.Sc
	.07	.11	10	.08	.10	.12 *	01	.17 **	.07	06		LIF
	.00	.12*	00	.05	.12 *	.11	.05	.15**	.08	.02		РНҮ
	.02	.07	11	.02	.11	.06	.06	.12*	.10	.02		ESP
	.05	.17 **	10	.09	.15 **	.08	.03	.09	.08	09		NAT
	.03	.12 *	13 *	.05	.09	.12	.01	.15 **	.07	06		KNO
<u> </u>	.03	.15**	05	.10	.16 **	.10	.06	.18 **	.07	03		APP
*	02	.10	04	.04	.10	.10	.03	.07	.12 *	04		INT

Table (32) Continued

<i>L.L.</i>	33	21.	ţ	20	19	18.	1/.	1		16.	Į.	1.5	14.	-	13	į	12	Q.No	
TV program about science at home?	your science home work?	Does any one at home eyer hale	about what you are learning in science class?	doing science homework each week?	quiz in school?  About how much time do you good.	How often do you take a science test or	How often do you watch a film, video or	students in school?	iments by your self	von do	science experiments?	give a science lesson.	How often do you listen to your teacher	about science do you agree.	With which of the following statements	T WILL GOVE III SOLICIECE.	I am good in Science	Question	
01	11	4.4	01	02	3 5	-03	04		08	00	.00		.11		***************************************	.12,	10	T.Sc	
01	.14*		02	04	0/	07	03		04		.03	- 1	.111*			.0/	2	HIT	
01	05		01	02	.01	01	06		04		.02	,				i		РНҮ	
2	02		01	.00	01		05		08		.00	Ċ	07		***************************************	.04		ESP	
04	12*		00 ;	.00	00		.01		13 *		10		03			.17 **		NAT	
00	14*		03	02	04		<u>-</u> .01	-	04	-	.04	. 1	11			.08		KNO	
01	08		.01	02	02		-04		07		02	·Ç			-	.11		APP	
02	05		02	02	02		- 05		09		02	.08		·		.12 *	`	INI	

#### Number of books in the home

Question 3 of the Students Questionnaire asked about the number of books in the home which is also an indicator of parents' education and study environment in the home. Frequency and percentage of response in each category is given in Table (S3).

Table (S3)

Frequency and Percentage of Responses in Each Response Category.

Q3: How many books are there in your home? (Do not count newspapers, magazines, or comics)

Label	Value	Freq.	%
0 - 10 Books	1	55	10.4
11 - 24 Books	2	187	35.3
25 - 100 Books	3	161	30.4
> 100 Books	4	122	23.1
Missing	0	4	.8
Total		529	100

About 23% of 8th grade students who took the Science Test reported to have more than 100 books (excluding journals, magazines and newspapers and school books) at home, 45.7% of the students with less than 25 books and over 10% reported having ten books or less. More than 53% of the students in the West Bank reported having less than 25 books in their homes and more than 13% of all students in the West Bank reported having from 10 to no books at all in their homes. (Ahlawat et al., 1992).

As a matter of fact, number of books at home does have a positive influence on students' achievement. This can be easily verified from the significant positive correlation between these two variables found in every one of the 20 IAEP countries (Learning Science, Lapointe et al, 1992, P.65). However, Row 3 of Table 32 shows a positive but non-significant relationship between number of books at home and all the components of the Science Test except one viz., the Integration component of the Science Test where the correlation statistically significant.

#### Parental involvement and student achievement in science

Items 9 (My parents are interested in Science), 20 (Does any one at home ever talk to you about what you are learning in science class?) and 21 (Does any one at home ever help you with your science homework?) indicate parental involvement and home support. Information about item 9 is presented in Table (S4) and about items 20 and 21 is presented in Table (S5).

Table (S4)

Percentage of Responses in Each Category of the Items 10-15

Q.	Question	SD	D	UN	TA	T*SA
No.		~ ##	, T			
7.	Much of what you learn in science class is useful in everyday life.	3.2	0.6	6.0	34.1	35.3
8.	It is important to know some science in order to get a good job.	2.6	2.3	10.0	52.3	31.9
9.	My parents are interested in Science.	4.2	12.1	21.7	38.4	22.5
10.	Scientists discover new facts and ideas about science by making observations and doing experiments.	1.1	1.9	6.0	29.5	61.2
11.	Facts and ideas abut science cannot be questioned or changed.	4.2	9.1	15.5	37.4	32.9
12.	I am good in Science.	1.7	2.1	9.6	58.0	27.6

Table (S5)
Response Percentage in Each Category for Each Parental Involvement Item

Q.No	Questions	Percent		
		Yes	No	
20.	Does anyone at home ever talk to you about what you are learning in Science homework?	77.9	19.7	
<b>21</b>	Does anyone at home ever help you with your Science homework?	38	59	

Question 9 in Table (S4) shows that 38% of the students are either not sure or disagree with the statement (My parents are interested in Science) signifying the fact that their parents are not interested in science. However, 61% of the respondents reported that their parents were interested in science. This puts UNRWA in the third position from the top among the 20 countries, Spain with an average percent correct score of 68 on the Science Test being the first, and China with a 67 average percent correct the second on this measure. It is worth noting that the highest achiever, Korea, (78% correct) on the IAEPII Science Test had only 28% students reporting their parents were interested in science.

In the UNRWA fields sample, parents' interest in Science showed a significant negative correlation on the knowledge component in the Science Test (Table 32). At the international level, in 8 countries out of 20, there was a significant positive correlation between parents' interest in Science and students' performance on the Science Test. In the other 12 countries the relation between parents' interest in Science and students' performance on the Science Test was not significant.

When students were asked if someone at home talked to them about what they learnt in the science class, around 78% of the responses were Yes. As shown in Table (S5), parents were more likely to ask their children about their science class (Q20) than help them with their science homework (Q21). Only 38% of the students reported receiving help with their science homework.

Row 20 of Table 32 indicates no significant correlation with any of the science test components, whereas, row 21 of Table 32 shows a significant negative correlation between receiving help in science homework and achievement in Life Sciences and Nature of Science topics and the Knowledge skill.

Another home related factor in addition to parental involvement and help with science homework concerns the availability of facilities to watch video or television programs about science. Question 22 tackled this matter. Table S6 shows student responses ranging from "Never" to "Everyday". About 39% of the students reported watching a film, video or television program about science 2 or 3 times a week. This variable has no significant relation with performance on the Science Test.

Table (S6)

Response Percentage in Each Category for Each Science Learning Activity

		Response percent				
Q.No.	Question	Every day	2 or 3 times a week	Once a week		Never
14.	How often do you listen to the teacher give a Science lesson?	. 38.8	58.8	2.3	0.2	0.4
15.	How often do you watch your teacher to do Science experiments?	5.1	55.0	22.3	12.7	4.3
16.	How often do you do science experiments by yourself or with other students in class?	3.4	21.4	20.8	22.9	30.8
17.	How often do you watch a film, Video, or TV program about science in schools?	7.2	22.7	20.8	19.3	29.5
18.	How often do you take a science test or quiz in school?	3.2	35.5	31.6	26.1	2.8
22.	How often do you watch a film, Video or TV program about science at home?	10.8	38.9	26.5	13.8	7.0

# Watching television and science achievement

Some television programs are clearly targeted at developing the academic abilities of children and some countries provide more of this programming than do others. However, for many students the content of the TV has little academic value and consumes much of their time which could have been devoted to other more rewarding activities. Question 4 of the Students' Questionnaire collected such information.

Table (S7)

Percentage of Students in Each Category of Watching TV at Home on a School Day and Spending Time Doing Homework Each Week?

No.	Question	None	1 Hr	2 Hrs	3 Hrs	4 Hrc	5 Hrs	6 Hrs
	How many hours do you spend watching TV on a school day?	5.9	38.4	35.2	11.7	6.6	1.3	0.9
	About how much time do you usually spend each week on Science homework?	1.7	42.7	32.9	8.5	4.7	3.6	3.4

Table (S7) clearly shows that only 2% of the respondents reported watching TV 5 Hrs. or more on a school day. This corresponds to the lowest measure among the 20 countries in the *IAEPII* international study as reported by 2% of the responses of China and 4% of responses of France.

Row 4 of Table 32 shows that time spent on watching TV is positively correlated with students performance on the Science Test. Especially students who reported watching TV more frequently tended to score higher on Life Sciences, Physical Sciences, Earth and Space Sciences content and on the Knowledge and Application skills.

On the international level, in 10 out of the 20 *IAEP* countries, the correlation, between time spent on watching TV with achievement was negative, in one country (Portugal) it was positive and in 9 of them there was no significant relationship.

#### Homework and science achievement

Question 19 cited in Table (S7) and Question 6 as shown in Table (S8) refer to the amount of homework. Q19 refers to time spent doing sience homework each week while Q6 asked about time spent on homework for all school subjects each day.

Table (S8)

Frequency and Percentage of Responses in Each Category

# Q6: How much time do you usually spend each day on homework for all school subjects?

Label	Value	Freq.	%
No home work	1	23	4.3
1/2 hour or less	2	47	8.9
One hour	3	119	22.5
Two hours	4	151	28.5
> Two hours	5	183	34.6
Missing	0	6	1.1
Total		529	100

As can be verified from Table (S8) about 63% of 8th grade students in the total Science Test sample spent 2 hours or more each day on homework for all school subjects. The relationship between time devoted to homework for all school subjects every day and performance on the Science Achievement Test, as is evident from Row 6 of Table 32, was positive and significant (p<.01). In the West Bank also, this relationship was positive and highly significant (p<0.005) as reported by (Ahlawat et al., 1992).

In the *IAEPII* international study this correlation was positive in 9 out of 20 countries, negative in one and Zero in 10 of them.

As to the time spent on science homework each week, as reported in Table (S7), about 11.7% of respondents reported spending 4 hours or more each week doing Science homework. The largest percentage of the students around (44%) spent one hour or less each week on science homework. In West Bank, 14% of the students reported spending 4 hours or more each week doing science homework. The linear correlation between time spent doing science homework each week and science achievement is not significant in Table 32 Row 19. In the West Bank also this relation was not significant (Ahlawat et al., p.135).

The relationship between amount of time spent on science homework and achievement in science in the international study, was positive in 7 countries, negative in 3 and Zero in 10 of them.

As to how often did students watch their teacher doing science experiments and how often do they experiment with science on their own or with others, 4% of the students never saw the teacher do science experiment and about 31% never participated in conducting a science experiment or conducted one by themselves.

As regards to how often did students watch their teacher doing science experiments and its relationship with science achievement, Row 15 of Table 32 shows a nonsignificant correlation with all components of science achievement test while row 16 of Table 32 shows a significant negative correlation between students' participation in doing science experiments and their achievement on the Nature of Science component of the Science Test. In the West Bank the relationship of both variables with science achievement was negative with all components of the Science Test.

Item 17 of the questionnaire deals with how often do students watch a film, video or TV program about science in school. About 49% of the students watch such programs less than once a week including about 30% of them never watched such programs.

It is apparent from Row 17 of Table 32 that this activity, contrary to expectations, shows the tendency to be negatively correlated with all the Science Test components. The correlation, however, was not significant in any case.

# Amount of testing and science achievement

Frequency of testing is shown in Q18 Table (S6). Strangely enough, about 3% of the respondents from all the three fields reported that they have never taken a test or quiz in school, while on the other hand, about 3% reported taking it every day. Around 32% of the students reported taking a test in science once a week and the highest percentage of responses 35.5% reported taking a test 2 or 3 times a week. The IAEPII international study results indicate that schools in most countries do not use tests frequently to evaluate student performance in science. Tests are mostly used in Taiwan and the Soviet Union, both high performing countries, and the United States and Jordan, lower scoring groups. From 67% to almost 90% of students in these 4 groups reported being tested at least once a week. Less than one half of the students from most other 20 participating countries reported weekly testing (Lapointe et al., 1992).

The relationship between the amount of testing and science achievement in the three fields is negative but nonsignificant while in the West Bank sample it was negative and highly significant (Ahlawat et al., 1992, p.136).

Descriptive results as reported in *IAEPII* international study suggest that testing is relatively infrequent among most *IAEPII* participants, and even among those countries that use tests more frequently - Taiwan, the Soviet Union, The U.S. and Jordan - the relationship between amount of testing and performance on the Science Test is not consistent. The amount of testing correlated with science achievement negatively in 5 countries, Zero in 12 countries and positively in 3 countries out of 20.

# Attitude toward science and science achievement

Students bring to school certain attitudes toward education in general and toward specific school subjects. These attitudes, to some extent, contribute to, and are a product of, academic success. Students who approach a school subject enthusiastically are more likely to do well in that subject and, conversely, students who succeed in a content area are more likely to develop positive attitudes toward it. The students questionnaire enquired about the extent the students agreed with the following statements:

- Q 7: Much of what is learned in Science is useful in everyday life.
- Q 8: It is important to know some science in order to get a good job.
- Q 10: Scientists discover new facts and ideas about science by making observations and doing experiments.
- Q 12: I am good at science.

As shown in Table (S4) by students' responses on Q7, Q8, Q10 and Q12, the majority of 8th grade students in the entire sample expressed positive attitudes toward Science.

With reference to these different attitude items 84% to 90% of the students in the total sample have definitely positive attitudes towards Science. In the *IAEPII* study the highest percentage of students with positive attitudes towards science was in Jordan (82%) while the lowest was in Korea (23%) - a notable exception - which was the highest achiever in science (Ahlawat et al., 1992, p. 137).

Everyone of the four attitude items presented in Rows 7,8,10 and 12 of Table 32 correlated positively with every component of the Science Test.

More than 70% of the students have agreed with the statement that facts and ideas abut science cannot be questioned or changed (Q11). This means, more intelligent students disagreed with the statement, as a consequence, the correlation with science achievement components is negative but statistically nonsignificant.

# Leisure reading and science achievement

While reading for fun is not directly related to performance on the Science Test, consistent readers tend to be high achievers in many academic areas.

Q 5 in Table (S9) shows responses as to the frequency of reading for fun outside school. Around 16% of the respondents reported that they never or hardly ever read for fun; while the largest percentage (44%) reported that they read once or twice a week.

Table (S9)
Frequency and Percentage of Responses in Each Category

Q5: How often do you read on your own for fun outside of school?

Label	Value	Freq.	%
Every day	4	127	24.2
Once or twice a week	3	232	43.9
Once or twice a month	2	84	15.9
Never or hardly ever	1	82	15.5
Missing	0	4	0.8
Total		529	100

Around one quarter (24.2%) of the students claimed reading on their own for fun almost daily.

*IAEPII* international study results show that there is a positive relationship between leisure reading and science achievement in 16 out of 20 populations. In UNRWA students' population too, leisure reading has positive correlation with everyone of the Science Test components but its magnitude did not attain statistical significance at the  $\alpha = 0.05$  level.

# Science for girls or for boys or for both girls and boys

Q 13 in the Students' Questionnaire seeks the opinion of 8th grade students on whether science was more for boys or more for girls or about equally for both. Table (S10) presents the response frequency and percentage thereof in each response category; 16% reported that science is more for boys than for girls, 13% reported that science is more for girls than for boys and the majority 71% reported that it suits boys and girls equally/

Table (S10)

Frequency and Percentage of Responses in Each Category

# Q13. With which of the following statements about science do you agree?

Label	Value	Freq.	%
- Science is more for boys than for	1	84	15.9
girls Science is more for girls than for	2	68	12.9
boys Science is for both boys and girls	3	372	70.3
equally Missing	0	5	0.9
Total		529	100

In Jordan and the West Bank, groups of students stating science is more for girls than for boys scored significantly lower on science achievement test than did either of the other two opinion groups (Ahlawat et al., 1992).

Oneway ANOVA conducted in this study, however, did not show any significant differences among the Science Test means of the three groups.

# Chapter VII

# MAJOR RECOMMENDATIONS

- 1. In-depth study of the findings of the report should be conducted at UBH (A) and at the field level to identify specific points of weakness of UNRWA students in Math and Science.
- 2. A conference / workshop should be conducted at UBH (A), to be attended by two school supervisors from each Field (Math and Science) and GESs (Math & Science) to work on the following:
  - a) To prepare a list of points of weakness of students achievement in Math and Science.
  - b) To identify the content areas in Math and Science textbooks which match the areas of weakness in students achievement.
  - c) To identify possible reasons / factors behind each point of weakness and prepare guide-lines for preparation of a teacherguide to be distributed to Math and Science teachers to assist them to improve their methods of teaching.
  - d) To prepare two prototype enrichment material packages one in Math and one in Science of specific points of weakness in Math & Science.
- 3. Math & Science committees at Field level should conduct an in-depth study of the students achievements in Math and Science and propose a plan of action to improve student achievement in Math and Science in their Fields.
- 4. GESs (Math & Science) should identify the international consensus of Math & Science concepts and skills and work on preparation of instructional material to cover these concepts and skills for each grade to be distributed to teachers for teaching / learning purposes.
- 5. GESs (Math & Science) together with the Math and Science committees at Field level should prepare an Item-Bank for each grade (4-10) to be used by Math and Science teachers for diagnostic purposes.

- 6. In-service Institute of Education short courses should be conducted for Math and Science teachers at all Fields to train them on methods of teaching the main content areas in Math (Numbers, Measurements, Geometry, Algebra, and Data Analysis) and the main cognitive skills (Computation, Procedural Knowledge and Problem Solving), and the main content areas in science (Life Sciences, Physical Sciences, Nature of Science) and main skills (Knowledge, Application and Integration).
- 7. The report of the study should be distributed to all concerned Palestinians who should be encouraged to utilize the findings of the report when they prepare their own Math and Science curricula, textbooks and teacher training.
- 8. A detailed plan of action should be prepared by GESs (Math & Science) for the implementation of all the above-mentioned recommendations together with a plan of follow-up of the implementation of the plan of action in the Fields.
- 9. The results of this report should not be used by any means for warning, punishment or reward of any staff member in *UNRWA*. On the contrary these results should encourage the educational administrators to facilitate the work of the technical staff.
- 10. The study should be repeated after two years (in 1996) for comparative purposes with the results of this report and to evaluate the effectiveness of the plans which will be prepared to improve the students achievements in Math & Science.
- 11. An Arabic version of this report should be prepared and distributed to all *UNRWA* Math and Science teachers and to be discussed with them through short meetings.
- 12. The report should be utilized in the in-service courses, especially in the Headteachers and school supervisors courses, as for example, when dealing with the topic "Action Research".

# References

- Ahlawat et al. (1993). Student Achievement in Jordan And the West Bank: A comparative Perspective. National Centre for Educational Research and Development Amman Jordan.
- Lapointe, A.E. Askew, J.M., and Mead, N.A. (1992). Learning Mathematics. Princeton N.J.: NAEP Educational Testing Service.
- Lapointe, A.E., Askew, J.M., and Mead, N.A. (1992) Learning Science: Princeton N.J.: NAEP Education Testing Service.