

The Hashemite Kingdom of Jordan



National Center for Educational Research and Development (NCERD)

**Monitoring of Learning Achievement Project**

Project Steering Committee

Dr. Victor Billeh    Dr. Kapur Ahlawat    Dr. Tayseer Al-Nahar

Report No. (12)

**Assessment of Learning  
Achievement of Grade 4  
Students in Jordan**

**A Preliminary Report**

Dr. Kapur Ahlawat

Dr. Tayseer Al-Nahar

Dr. Victor Billeh

**Publication Series No: 27**

**July 1994**



# **A C R O N Y M S**

**\* \* \***

<b>EMIS</b>	<b>Education Management Information System.</b>
<b>ERP</b>	<b>Education Reform Plan.</b>
<b>MOD</b>	<b>Ministry of Defense.</b>
<b>MOE</b>	<b>Ministry of Education.</b>
<b>NECRD</b>	<b>National Center for Educational Research and Development.</b>
<b>UK</b>	<b>United Kingdom.</b>
<b>UNESCO</b>	<b>United Nations Educational Scientific and Cultural Organization.</b>
<b>UNICEF</b>	<b>United Nations Children's Education Fund.</b>
<b>UNRWA</b>	<b>United Nations Relief and Works Agency.</b>
<b>USA</b>	<b>United States of America.</b>

## Table of Contents

Acronyms .....	i
Table of Contents .....	ii
List of Tables .....	V
List of Figure .....	ix
Background .....	1
Objectives of the Project .....	2
A. Direct (Short-term) Objectives .....	3
B. Indirect (Long-term) Objectives .....	4

## Methodology of the Project

Sample Design .....	4
Sample Frame .....	5
Sampling Issues .....	7
Sampling Procedures .....	8
The Monitoring Study Sample .....	9
Areas of Achievement .....	10
Context of learning .....	14
Contextual Framework of the Curriculum-Referenced .....	15
Why Curriculum-Referenced? .....	16
Sequential Steps .....	17
Test Development Teams .....	19
Piloting the Tests .....	20
Analysis of the Pilot Data .....	20
Administration of the Tests .....	21
Data Collection .....	23
Data Analysis .....	24

## Results

Arabic Test .....	26
Performance at the National Level .....	26
Performance Acorss Governorates .....	27
Performance by Sex .....	28
Performance by Educational Authority .....	30
Performance by School Location .....	31
Science Test .....	33
Performance at the National Level .....	33
Performance Acorss Governorates .....	34
Gender Differences on the Science Test .....	36
Performance by Education Authority .....	37
Science Test Performance in Urban and Rural Schools .....	40

Math Test .....	42
Performance on the National level .....	42
Performance Acorss Governorates .....	42
Gender Differences on the Math Test .....	44
Performance on the Math Test in Each Education Authority .	47
Math Test Performance in Urban and Rural Schools .....	49
 Life Skills	
Performance on the National level .....	51
Performance by Governorate .....	52
Gender Differences on life skills Test .....	54
Performance on the life Skills Test in Each Education Authority .....	56
Differential Performance of Urban and Rural School Students on the Life Skills Test .....	58

## List of Tables

Table (1):	Distribution of Mainstream Schools Containing Grade 4 classes in Jordan .....	6
Table (2):	Distribution of Grade 4 Student Population in the Mainstream Schools of Jordan .....	7
Table (3):	Sampling distribution of Schools Across Governorates by Education Authority and School .....	10
Table (4):	Distribution of the Arabic Test Sample Across Governorates by Education Authority, Location, School Gender and Student Gender.....	11
Table (5):	Distribution of the Life Skills Test Sample Across Governorates by Education Authority, Location, School Gender and Student Gender.....	12
Table (6):	Distribution of the Math Test Sample Across Governorates by Education Authority, Location, School Gender, and Student Gender .....	13
Table (7):	Distribution of the Science Test Sample Across Governorates by Education Authority, Location, School Gender, and Student Gender .....	14
Table (8):	Table of Specifications for the Literacy (Arabic Language) Test .....	22
Table (9):	Table of Specifications for the Numeracy (Math) Test.	22
Table (10):	Table of Specifications for the (Science) Test .....	23
Table (11):	Table of Specifications for life Skills Test .....	23
Table (12):	Average Percent Correct on the Arabic language Test .....	26
Table (13):	Average Percent Correct Score, Standard Error, and 95% CI on the Arabic Test in Each Governorate .....	26
Table (14):	Average Percent Correct and Standard Error on Arabic Test Subscales in Each Governorate .....	28

## List of Tables

Table (15): Average Percent Correct on Arabic Test Subscales for Male and Female Students .....	29
Table (16): Average Percent Correct Score Standard Error and 95%CI for Each Education Authority .....	30
Table (17): Average Percent Correct on Arabic Test Subscales in Urban and Rural School Location .....	32
Table (18): Average Percent Correct Score, Standard, Minimum and Maximum Scores on the Science Test Subscales ..	33
Table (19): Average Percent Correct Score, Standard Error, and 95%CI of the Science Test in Each Governorate .....	34
Table (20): Average Percent Correct Score and Standard Error of Each of Science Test Subscales for Each Governorate .	35
Table (21): Average Percent Correct Score, Standard Error, and 95% CI of Male and Female Students on the Science Test Subscales .....	37
Table (22): Average Percent Correct Score, Standard Error, and 95% CI on the Science Test Subscales in Each Education Authority .....	38
Table (23): Average Percent Correct Scores on the Science Test Subscales of the 4th Grade Students in Urban and Rural Schools.....	40
Table (24): Average Percent Correct Score, Standard Error, Minimum and Maximum Score on Each Subscale of the Math Test .....	42
Table (25): Average Percent Correct Score, Standard Error, and 95%CI on the Math Test in Each Governorate .....	43
Table (26): Average Percent Correct Score and Standard Error on Each Math Test Subscale in Each Governorate .....	43
Table (27): Average Percent Correct Score, Standard Error, and 95%CI of the 4th Grade Male and Female Students on Each Math Test subscale .....	46

## **List of Tables**

Table (28): Average Percent Correct Score, Standard Error, and 95%CI of the Grade 4 Students on Each Math Test Subscale .....	48
Table (29): Average Percent Correct Score on Each Math Test subscale in Urban and Rural School .....	49
Table (30): Interval Frequency Distribution of the Percent Correct Score on the life Skills Test .....	51
Table (31): Average Percent Correct Score on Each Subscale of life skills Test .....	52
Table (32): Average Percent Correct Score and Standard Error on Each Subscale at the Governorate level .....	53
Table (33): Performance on the Life Skills Subscales in Each Governorate .....	54
Table (34): Average Percent Correct Score, Standard Error, and 95% CI of the Male and Female Students on Each subscale and Total Score of the Life Skills Test .....	55
Table (35): Average Percent Correct Score, Standard Error, and 95% CI on the Subscales of the Life Skills Test in Each Education Authority.....	57
Table (36): Average Percent Correct Score, Standard Error, and 95%CI on Each subscale and the Total Test Score by School Location.	

## List of Figures

Figure (1): Average Percent Correct Score on the Arabic Test in Each Governorate .....	27
Figure (2): Gender Difference on the Arabic Test for Grade 4 students .....	29
Figure (3): Average Percent Correct Score of the Grade 4 Students on the Arabic Test in Each Education Authority .....	31
Figure (4): Differential Performance of the Urban and Rural Schools Grade 4 Students on the Arabic language Test .....	32
Figure (5): Average Percent Correct Score of the 4th Grade Students on the Science Test in Each Governorate .....	34
Figure (6): Differential Performance of Male and Female Grade 4 Students on Various Components of the Science Test .....	36
Figure (7): Average Percent Correct Score of 4th Grade students on the Science Test in Each Education Authority .....	39
Figure (8): Location Differences on the Total Score and Subscale Score Science Test .....	41
Figure (9): Average Percent Correct Score on the Math Test in Each Governorate .....	44
Figure (10): Comparative Performance of Male and Female Grade 4 Students on Various Subscales of the Math Test .....	45
Figure (11): Average Percent Correct Score of the 4th Grade Students on the Math Test in Each Education Authority .....	47
Figure (12): Performance of Grade 4 Students on Various Math Test Subscales in Urban and Rural Schools .....	50
Figure (13): Average Percent Correct Score of the 4th Grade Students on the life skills Test in Each Governorate ....	53

## **List of Figures**

Figure (14): Gender Differences in the Performance of 4th Grade students on the Subscales and Total Score of the Life Skill Test .....	55
Figure (15): Average Percent Correct Score on the Life Skills Test in Each Education Authority .....	56
Figure (16): Performance on the Life Skills Test in Urban and Rural Schools .....	59

## Background

Since the Jomtien declaration of Education For All by the year 2000 most countries, especially the developing countries in Africa, Asia and South America have made strenuous effort to provide access to primary education for increasing number of school-age children. Primary enrollment rates have substantially increased in all the countries. Actually, in some countries education system has expanded to the extent that it reaches nearly all school-age children. Universal enrollment automatically suggests that problems of gender or regional disparity in access to schooling have been overcome by the education system. This, however, provides no assurance of gender or regional equality in terms of the quality of education, nor does it tell us about the learning achievement levels of the pupils. Cognizant of the overarching importance of monitoring the quality of Education UNESCO and UNICEF have launched upon an initiative to encourage several countries (about 30) to develop national systems for monitoring and assessment of learning achievement of the 4th grade pupils in order monitor the Learning Achievement of Education - For-All Goals.

Jordan is one of the five countries that have participated in this monitoring assessment from the very outset. This, of course, was not a happenstance. Education system in Jordan is undergoing a comprehensive reform aimed at enhancing student achievement levels by raising the quality of education. Measurement of student achievement is of critical importance to educational reform efforts. In this context, a national survey of instructional quality was focus of attention in Jordan during the same period when UNESCO/UNICEF were launching their joint project in a small set of selected countries to monitor the Education -For-All Goals. Under these circumstances the aims and objectives of the UNESCO/UNICEF sponsored international study happened to be in congruence with a subset of those of the more comprehensive longitudinal National Assessment of Instructional Quality survey planned by the National Center for Educational Research and Development (NCERD) in cooperation with the Ministry of Education (MOE) in Jordan.

This preliminary report, nevertheless, will be limited to describing the broad objectives, design, instruments and initial results of the data analysis obtained from the Grade 4 subsample of the stratified multistage random sample of the National Assessment of Instructional Quality study in Jordan.

Achieving the Goals of Education for All requires the fulfillment of two necessary basic conditions:

- (1) Provision of essential educational facilities for all the people irrespective of their age, gender, ethnicity and region of inhabitation in a country.
- (2) Effective mechanism to ensure the acquisition of required basic educational skills by the people up to acceptable standards.

Existence of these two conditions may not guarantee the achievement of the goals of basic Education - For - All in a country but it is an absolute prerequisite. While the former condition is necessary to address the fundamental issues of access and coverage, the latter implies monitoring the achievement of objectives, both quantitatively and qualitatively. Monitoring evaluation, in this sense, plays a key role in successful implementation of the action plans devised to achieve the desired objectives in the target populations.

Monitoring evaluation encompasses the assessment of the progress made in terms of both quantitative expansion of the education system and achievement levels attained by different target groups, as well as , identification of bottlenecks and weakpoints providing for modifications and improvements of different aspects of the program.

### **Objectives of the Project**

As for information related to quantitative expansion and reach of the education system, the NCERD in cooperation with the Ministry of Education maintains a comprehensive educational database and an operational Educational Management Information System (EMIS), which was used to develop indicators of progress of basic education in Jordan reported elsewhere. The main focus of this monitoring evaluation therefore, is on measurement of the achievement levels of the primary cycle students. Although Jordan has statutory free and compulsory basic education which covers (6-16) -year-age groups enrolled in grades (1-10), in order for maintaining congruity with the other countries participating in Monitoring Learning Achievement of the Education -For-All Goals, 4th grade was targeted.

The objectives of this study fall into two broad classes: (A) Direct (Immediate) and ; (B) Indirect (Distant).

## **A. *Direct (Short-term) Objectives***

Direct objectives include:

- (a) Measuring learning achievement of students at the end of the 4th- year of basic education in the following key subjects;
  - (1) Arabic Language (Literacy),
  - (2) Mathematics (Numeracy),
  - (3) Science, and
  - (4) Life Skills.
- (b) Studying students' family backgrounds, parental practices, perceptions, expectations, and attitudes related to schools and classroom activities, and their children's performance.
- (c) Studying teachers' instructional practices, attitudes, beliefs, perceptions, cognitions, expectations and background characteristics.
- (d) Studying school characteristics and principals' administrative practices, beliefs and attitudes.
- (e) Students' attitudes, beliefs perceptions and cognitions of schools, teachers, classroom practices and of themselves.
- (f) Estimating distribution parameters of salient school, family , principal, teacher and student variables across different regions, student gender, education authorities and rural / urban environments.
- (g) Studying relations between various homebackground, community, school, teacher, classroom, and student related factors and students' achievement.
- (h) Identifying strengths and weaknesses of the system, deriving policy implications and suggesting plans of action:

## **B. *Indirect (Long-term Objectives)***

The indirect objectives include:

- a) Building up national capacity for monitoring educational progress, and assessment of student achievement and instructional quality.
- b) Institutionalizing the evaluation function by establishing mechanism for monitoring learning achievement at regular intervals.
- c) Establishing a channel of communication and promoting dialogue between evaluation research and educational planning to formulate informed policy and effective plans of action.
- d) Establishing Education Management Information System (EMIS) and promoting the utilization of relevant, reliable and timely information to develop useful indicators of access, quality, and efficiency of the education system.

## **Methodology of the Project**

### **Sample Design**

Drawing an optimally efficient and representative educational survey sample within the constraints imposed by the financial resources, technical capacity and practical considerations is seldom a straightforward task. Difficult decisions involving tradeoffs between competing claims have to be made.

In countries where demographic, geographic, sociopolitical factors, and characteristics of education system or schools vary within and / or across regions; multistage stratified sampling approaches have been found more satisfactory.

The ideal **target population** for the monitoring learning achievement survey constitutes all the primary school students enrolled in the Hashemite Kingdom of Jordan. The **defined population**, however, constituted of all the grade 4 students enrolled in the mainstream schools.

## ***Sampling Frame***

While school was the sampling unit at the first stage of sampling, ideally the target population should have included all the schools, in Jordan, however, there is a substantial proportion of schools that do not have grade 4. Therefore, the schools (operating in the mainstream of education system) that do not have grade 4 classes were excluded from the defined target population of schools. The educational database at the NCERD's Education Management Information System (EMIS) maintains comprehensive census data collected annually from all the schools. This considerably facilitated the sample selection procedures and almost guaranteed the accuracy and integrity of the sampling frame.

The elements of the defined population sample frames comprising all the contained 2299 schools with 21637 Grade 4 children studying in them.

Jordan is divided administratively into 8 Governorates. These Governorates, at the time of this study, were divided educationally into a total of 23 Education Directorates.

While the MOE bears the full responsibility for providing basic education for all the school-age children (6-16 years old), there exist other nongovernment and government agencies that run a significant number of schools to provide basic education for substantial number of children in different parts of the country. The Ministry of Defense (MOD) runs a small number of schools attended mainly by the children of MOD personnel. The United Nations Relief and Works Agency (UNRWA) runs a significant number of basic schools for Palestinian refugee children. Second to MOE, a substantial number of Private schools are run by the Private Sector and are concentrated in the large cities. Although most public and UNRWA schools in Jordan are virtually unisex, there exist a large number of coed schools in which coeducation is limited to a few primary grades. Most Private schools, on the other hand, are coeducational institutions.

Table 1 presents the breakdown of schools comprising the fourth class, across the 8 Governorates by the Education Authority (MOE, MOD, UNRWA, Private), area of school Location (Urban, Rural), and the school Gender (Male, Female, Coed) based on 90-91 scholastic year data of the MOE. Table 2 shows the distribution of class 4 students enrolled in those schools.

Apparently Tables 1 and 2 display the distribution of the entire defined population of schools and 4th class students included in the sample frame of this study.

**Table 1**  
**Distribution of Mainstream Schools Containing Grade 4 Classes in Jordan**

	<b>Governorate</b>								<b>Total</b>
	<b>Amman</b>	<b>Zarqa</b>	<b>Balqa</b>	<b>Irbid</b>	<b>Mafrq</b>	<b>Karak</b>	<b>Tafileh</b>	<b>Ma'an</b>	
<b>Authority</b>									
M.O.E	493	187	134	498	221	143	49	129	1854
M.O.D		2						8	10
UNRWA	72	29	10	46	2				159
Private	190	24	13	30	3	5	1	5	271
<b>Total</b>	<b>755</b>	<b>242</b>	<b>157</b>	<b>574</b>	<b>226</b>	<b>148</b>	<b>50</b>	<b>142</b>	<b>2299</b>
<b>Location</b>									
Urban	545	107	42	132	21	22	12	36	917
Rural	212	135	115	445	205	126	38	106	1382
<b>Total</b>	<b>757</b>	<b>242</b>	<b>157</b>	<b>577</b>	<b>226</b>	<b>148</b>	<b>50</b>	<b>142</b>	<b>229</b>
<b>Gender</b>									
Male	204	78	42	215	58	32	14	47	690
Female	180	72	25	172	40	14	12	26	541
Co-Ed	373	92	90	190	128	102	24	69	1068
<b>Total</b>	<b>757</b>	<b>242</b>	<b>157</b>	<b>577</b>	<b>226</b>	<b>148</b>	<b>50</b>	<b>142</b>	<b>2299</b>

Source: NCERD Database for 1990 / 1991 scholastic year.

**Table 2****Distribution of Grade 4 Student Population in the Mainstream Schools of Jordan**

	<b>Governorate</b>								<b>Total</b>
	<b>Amman</b>	<b>Zarqa</b>	<b>Balqa</b>	<b>Irbid</b>	<b>Mafraq</b>	<b>Karak</b>	<b>Tafileh</b>	<b>Ma'an</b>	
<b>Authority</b>									
M.O.E	4545	1575	1283	4656	1867	1396	485	1099	16906
M.O.D	8	30						88	126
UNRWA	563	229	101	370	18				1281
Private	2140	370	159	386	23	58	22	83	3241
<b>Total</b>	<b>7256</b>	<b>2204</b>	<b>1543</b>	<b>5412</b>	<b>1908</b>	<b>1454</b>	<b>570</b>	<b>1270</b>	<b>21637</b>
<b>Location</b>									
Urban	5483	1039	445	1334	200	262	149	391	9303
Rural	1799	1169	1106	4116	1710	1195	358	881	12334
<b>Total</b>	<b>7282</b>	<b>2208</b>	<b>1551</b>	<b>5450</b>	<b>1910</b>	<b>1457</b>	<b>507</b>	<b>1272</b>	<b>21637</b>
<b>Gender</b>									
Male	2157	724	485	2091	572	423	172	442	7066
Female	1908	675	342	1829	405	227	136	283	5805
Co-Ed	3217	809	724	1530	933	807	199	547	8766
<b>Total</b>	<b>7282</b>	<b>2208</b>	<b>1551</b>	<b>5450</b>	<b>1910</b>	<b>1457</b>	<b>507</b>	<b>1272</b>	<b>21637</b>

Source: NCERD Database for 1990 / 1991 scholastic year.

**Sampling Issues**

The NCERD's Assessment of Instructional Quality study was designed to meet the long-term monitoring evaluation and impact assessment needs of the education system in the context of the comprehensive 10-year-long Education Reform Plan (ERP).

The issues of coverage and equitable distribution of resources and more importantly the impact of reform inputs on instructional quality measured by student achievement levels were of paramount consideration. In view of this, the selection of a nationally representative sample of all types of schools was critical to achieving both long-term and short-term goals of the study.

Besides, different education authorities, dimensions of geographical regions, rural / urban school locations, school gender and size of the school were of particular importance for the study because of

the mediatory influences of these exogenous factors on school quality and student achievement.

Moreover, the National Survey of Instructional Quality was ordained to study student achievement at three grade levels (4th, 5th, and 8th). Reasons for selecting these particular grades out of whole range of (1-12) grades will be explained elsewhere in the proper context of the broader study.

## **Sampling Procedures**

After considering the key operating factors including statistical, financial, and practicability it was decided that a stratified random sample of 245 schools would be adequate to provide for desirable accuracy of estimates of the parameters of interest, practical implementation (management and control of data collection and analyses) and would not overtax the availability of financial and manpower resources. A multistage stratified random sample of 245 schools was selected.

The 1990 - 91 Database at the NCERD's newly established Education Management Information System (EMIS) has data on each individual school in the country. A sampling frame was created of all the academic schools in the mainstream (excluding vocational schools, KGs, and special schools for the handicapped children) imposing the constraint that to belong to the sample frame set a school must have two or more students in at least one of the three grades viz., 4th, 5th, and 8th. The sample frame thus created was divided into four strata on the basis of Education Authority to which the schools belonged.

A slightly disproportional equiprobability sample of schools was selected from the sampling frame from the stratum defined by each education authority (MOE, MOD, UNRWA and Private). Disproportionality was introduced by the fact that 6 MOD schools were selected out of 18 (the total number of schools run by some other Ministries like Defense, Higher Education and Social and community Development). The 245 schools sample, thus selected, was carefully examined to ensure that it adequately represents the following critical dimensions of particular interest in Jordan:

- \* **School Size:** A previous NCERD study had classified schools into six categories on the basis of total number of students in a school and highlighted the role of school size in commanding the essential school facilities that have both direct and indirect influence on instructional quality by establishing that the basic school inputs vary with the school size.

- \* **School Location:** Whether a school is located in urban or rural area.
- \* **School Gender:** Whether a school is male, female, or coed.
- \* **Region:** North, Center, or South.
- \* **Governorate.**

A few minor adjustments, mainly with respect to school size were made to ensure the representativeness of very small schools which tended to be grossly under represented because of not meeting the grade 4 condition.

The next step was taken to ensure that the national sample represents other characteristics of schools, such as teacher qualifications, teachers' average experience, principals' qualifications and experience, availability of essential school facilities (Library, Laboratory, Multipurpose Activities Hall, and Workshop, ownership (whether the school building is owned, rented or partly owned partly rented) and shift of the school (whether a school meets under single-shift or double-shift conditions. This was done by conducting appropriate statistical tests comparing the sample with the residual sample frame and with the residual target population on every one of the aforementioned school characteristics. The statistical tests of significance dictated by the discrete or continuous nature of the variable involved were t-test and Chi Square test. The sample was not found to differ significantly from the population with respect to any of the variables at 0.05 level of significance.

### **The Monitoring Study Sample**

The sample of 205 schools meeting the requirement of the existence of the 4th grade students, is in a sense a subset of the larger national sample of 245 schools. The monitoring study sample, therefore, consists of 205 schools.

To assess the power of the sample to estimate population parameters, the differences between the weighted cases in the sample and the sampling frame in a number of school characteristics were examined.

The characteristics in which sample and population were compared were: lowest grade available in the school, highest grade available in the schools, location, gender of students, having a full-time principal, qualifications of the principal, whether the building is owned or rented, number of shifts in the school, number of teachers in the school, percentage of teachers with various levels of qualifications, number of students, sport areas, number of auditoriums, existence of library, existence of science labs, existence of workshops, total number of classrooms and number of rooms for management purposes.

In all the contrasts examined the distribution of each variable in the sample was not different from its distribution in the population.

The following section examines the properties of the 4th grade sample with the sampling frame of schools fulfilling the requirement of having 4th grade students in them. Table 3 shows the population and sampling distribution of schools having Grade 4 classes in them, in respect of Education Authority, Governorate, School Location, and School Gender. Tables 4, 5, 6 and 7 display respectively the distribution of pupils who took Arabic, Life Skills, Mathematics, and Science test, across Education Authority, Governorate, School Location and School Gender.

**Table 3**

**Sampling distribution of Schools Across Governorates by Education Authority and School**

	Governorate								Total
	Amman	Zarga	Balqa	Irbid	Mafrag	Karak	Tafileh	Ma'an	
<b>Authority</b>									
M.O.E	45	15	11	42	22	14	5	13	167
M.O.D		1						5	6
UNRWA	4	4		2					10
Private	17	3	1	1					22
<b>Total</b>	<b>66</b>	<b>23</b>	<b>12</b>	<b>45</b>	<b>22</b>	<b>14</b>	<b>5</b>	<b>18</b>	<b>205</b>
<b>Location</b>									
Urban	46	14	1	8		3		5	77
Rural	20	9	11	37	22	11	5	13	128
<b>Total</b>	<b>66</b>	<b>23</b>	<b>12</b>	<b>45</b>	<b>22</b>	<b>14</b>	<b>5</b>	<b>18</b>	<b>205</b>
<b>School Gender</b>									
Male	13	7	4	20	5	5	1	10	65
Female	17	5	0	8	5	2	2	4	43
Coed.	36	11	8	17	12	7	2	4	97
<b>Total</b>	<b>66</b>	<b>23</b>	<b>12</b>	<b>45</b>	<b>22</b>	<b>14</b>	<b>5</b>	<b>18</b>	<b>205</b>

**Areas of Achievement**

Learning Achievement is such a global concept. It embraces everything pupils acquire by learning. When narrowed down to academic achievement or achievement in the school subjects taught by means of planned instruction by teachers in the school classrooms, it still includes a variety of school subjects. Although measurement of achievement in all the school subjects would undoubtedly provide a more

comprehensive and accurate picture of learning achievement, it may come at a prohibitive cost. Like sampling of schools and classes, sampling of subjects becomes imperative. From purely statistical point of view random sampling of a few subjects from all the school subjects taught at the primary level may look attractive but since some subjects form the basic foundation for other subjects, they are considered strategically more important than others.

Expert opinion in such cases converges on the choice of a judgmental sampling strategy. In case of Jordan, Arabic Language, Mathematics, and General Science in conjunction with the Life Skills were the logical choice.

**Table 4**

**Distribution of the Arabic Test Sample Across Governorates by Education Authority, Location, School Gender, and Student Gender**

	Governorate								
	Amman	Zarqa	Balqa	Irbid	Mafraq	Karak	Tafileh	Ma'an	Total
Authority									
M.O.E	1067	463	172	1026	409	275	96	269	3777
M.O.D		41						125	166
UNRWA	169	191		68					428
Private	444	71	11	11					
Total	1680	766	183	1105	409	275	96	394	4908
Location									
Urban	1285	494	53	219		78		133	2262
Rural	395	272	130	886	409	197	96	261	2646
Total	1680	766	183	1105	409	275	96	394	4908
School Gender									
Male	356	262	62	528	94	135	16	225	1678
Female	543	172		244	120	56	54	70	1259
Coed.	781	332	121	333	195	84	26	99	1971
Total	1680	766	183	1105	409	275	96	394	4908
Student Gender									
Male	616	353	95	617	139	188	19	248	2275
Female	1064	413	88	488	270	87	77	146	2633
Total	1680	766	183	1105	409	275	96	394	4908

**Table 5**

**Distribution of the Life Skill Test Sample Across Governorates by Education Authority, Location, School Gender, and Student Gender**

	<b>Governorate</b>								<b>Total</b>
	<b>Amman</b>	<b>Zarqa</b>	<b>Balqa</b>	<b>Irbid</b>	<b>Ma'fraj</b>	<b>Karak</b>	<b>Tafileh</b>	<b>Ma'an</b>	
<b>Authority</b>									
M.O.E	1036	465	176	1006	357	274	96	264	3674
M.O.D		46						122	168
UNRWA	171	157		70					398
Private	445	69	11	11					536
<b>Total</b>	<b>1652</b>	<b>737</b>	<b>187</b>	<b>1087</b>	<b>357</b>	<b>274</b>	<b>96</b>	<b>386</b>	<b>4776</b>
<b>Location</b>									
Urban	1265	462	56	215		79		130	2207
Rural	387	275	131	872	357	195	96	256	2569
<b>Total</b>	<b>1652</b>	<b>737</b>	<b>187</b>	<b>1087</b>	<b>357</b>	<b>274</b>	<b>96</b>	<b>386</b>	<b>4776</b>
<b>School Gender</b>									
Male	354	237	63	520	85	134	16	221	1630
Female	518	163		240	119	56	53	67	1216
Coed.	780	337	124	327	153	84	27	98	1930
<b>Total</b>	<b>1652</b>	<b>737</b>	<b>187</b>	<b>1087</b>	<b>357</b>	<b>274</b>	<b>96</b>	<b>386</b>	<b>4776</b>
<b>Student Gender</b>									
Male	608	331	96	613	120	185	20	245	2218
Female	1044	406	91	474	236	88	76	141	2556
<b>Total</b>	<b>1652</b>	<b>737</b>	<b>187</b>	<b>1087</b>	<b>357</b>	<b>274</b>	<b>96</b>	<b>386</b>	<b>4776</b>

**Table 6**

**Distribution of the Math Test Sample Across Governorates by Education Authority, Location, School Gender, and Student Gender**

	Governorate								Total
	Amman	Zarqa	Balqa	Irbid	Mafrq	Karak	Tafileh	Ma'an	
<b>Authority</b>									
M.O.E	529	231	88	504	203	131	46	135	1867
M.O.D		20						61	81
UNRWA	82	94		36					212
Private	223	34	6	5					268
<b>Total</b>	<b>834</b>	<b>379</b>	<b>94</b>	<b>545</b>	<b>203</b>	<b>131</b>	<b>46</b>	<b>196</b>	<b>2428</b>
<b>Location</b>									
Urban	634	244	27	106		34		65	1110
Rural	200	135	67	439	203	97	46	131	1318
<b>Total</b>	<b>834</b>	<b>379</b>	<b>94</b>	<b>545</b>	<b>203</b>	<b>131</b>	<b>46</b>	<b>196</b>	<b>2428</b>
<b>School Gender</b>									
Male	176	130	32	263	45	63	8	111	828
Female	265	86		119	60	25	26	35	616
Coed.	393	163	62	163	98	43	12	50	984
<b>Total</b>	<b>834</b>	<b>379</b>	<b>94</b>	<b>545</b>	<b>203</b>	<b>131</b>	<b>46</b>	<b>196</b>	<b>2428</b>
<b>Student Gender</b>									
Male	314	169	52	307	64	92	9	123	1130
Female	520	210	42	238	139	39	37	73	1298
<b>Total</b>	<b>834</b>	<b>379</b>	<b>94</b>	<b>545</b>	<b>203</b>	<b>131</b>	<b>46</b>	<b>196</b>	<b>2428</b>

Table 7

**Distribution of the Science Test Sample Across Governorates by Education Authority, Location, School Gender, and Student Gender**

	Governorate								
	Amman	Zarqa	Balqa	Irbid	Mafrag	Karak	Tafileh	Ma'an	Total
Authority									
M.O.E	523	227	84	502	197	134	49	131	1847
M.O.D		21						64	85
UNRWA	87	95	5	34					216
Private	218	36		5					264
Total	828	379	89	541	197	134	49	195	2412
Location									
Urban	643	243	26	111		38		66	1127
Rural	185	136	63	430	197	96	49	129	1285
Total	828	379	89	541	197	134	49	195	2412
School Gender									
Male	169	132	30	262	48	66	8	113	828
Female	275	80		113	58	27	27	34	614
Coed.	384	167	59	166	91	41	14	48	970
Total	828	379	89	541	197	134	49	195	2412
Student Gender									
Male	281	185	45	306	72	91	10	123	1113
Female	547	194	44	235	125	43	39	72	1299
Total	828	379	89	541	197	134	49	195	2412

## Context of Learning

Learning of any sorts, however, does not occur in vacuum. Usually learning takes place in contextual environments.

Both community and school play powerful interactive roles in pupil learning achievement. In order for studying the structure of school and community factors that affect student learning achievement, it was decided to collect information on a number of contextual variables including, students' home environment, family background, parental practices and expectations; principals' background, and administrative behavior and attitudes; teachers' background and instructional practices, attitudes and expectations, and students' attitudes, perceptions and

cognitions. All these broad dimensions were tapped by means of four carefully designed, developed, and field-tested self-report instruments, namely: (1) Parent Questionnaire, (2) Principal Questionnaire, (3) Teacher Questionnaire, and (4) Student Questionnaire.

**Contextual Framework of the Curriculum-Referenced Achievement Tests:** The past decade (1980s) has been marked by education reforms in many countries, both developing and developed. A common feature shared by all current education reform programs is their emphasis on quality rather than quantity. Improving student achievement levels, enhancing critical thinking and problem solving skills, and developing higher level cognitive skills and abilities seems to be the central theme running through the stated objectives of most educational reform programs. This is certainly the major objective of the ten-year Education Reform Plan (ERP) in Jordan. Since the quantitative expansion had been attained at the cost of quality, the ERP in Jordan focuses on the qualitative aspects of the school product, i.e., improving the quality of basic and secondary school graduates without relenting the pursuit for achieving the goals of universal basic education for all in Jordan by the year 2000.

The learning achievement in Jordan had to be measured in the context of the ongoing education reform primarily aimed at enhancing student achievement levels by raising the quality of basic education.

Under the auspices of Education Reform Plan (ERP) the goals of basic education had been discussed and clarified, the curricula were developed and strategies for the development, design and production of textbooks, instructional technology, learning materials and instructional approaches were laid out.

In view of the ERP context the achievement tests were designed to serve several purposes. One was to establish the present achievement levels of students to serve as baseline data or benchmarks against which future progress could be compared and inferences could be made about the impact of education reform. Other purposes included comparing achievement of different groups of students and studying regional, locational and gender differences in student achievement; and obtaining formative information through identification of the areas of relative strength and weakness across content and skill domains of each subject.

Curriculum - referenced test development approach was adopted so that present status of students' knowledge, skills, and conceptual understanding underpinning broad curricular objectives in the core subjects could be established to inform present policy decisions and to serve as a frame of reference for the future plans. The tests were

designed such that test scores could be interpreted in both norm-referenced and criterion-referenced contexts for crosssectional and longitudinal comparisons on the one hand, and for formative information on the other.

**Why Curriculum - Referenced? :** The assumption here is that even though new curricula for all the basic school subjects have been developed under the ERP, the broad curricular objectives of teaching Mathematics, Science, and Arabic in the basic grades have not changed much. After all, any math curriculum, new or old, must aim at developing basic computation and problem solving skills and any language curriculum must include vocabulary, reading comprehension and writing skills among its objectives.

The point is, to the extent the objectives to be achieved are essentially similar, that is, the skills, abilities, concepts and attitudes to be acquired and developed are the same, tests can be developed to measure them such that the differences in student achievement levels obtained under different sets of conditions (curriculum, textbooks, teaching practices, etc., before and under reform) can be attributed to the changes in teaching / learning conditions introduced by the reform program. So, for a given subject, there exists a broad spectrum of concepts and skills that is common to different curriculums that may vary in content and approach from one state or even one board of education to another, as in the USA, and from one school system to another as in the UK. It is this common core of elements shared by different sets of curricula of the same subject across states or countries or across temporal dimension, (when the old curriculum is replaced by the new one, as in the case of Jordanian ERP) that provides justification for the standardized tests, national and international assessment of academic achievement and compatibility of norms across different populations.

This however, should not be taken as a plea to deny the existence of differences among different curriculums of a particular subject.

Alternative approaches; differential emphases, sequential arrangements, and modes of presentation of the subject matter; and inclusion or exclusion of specific topics or sets of objectives are too obvious to be ignored. It is expected that certain differences between the old curricula and the new ones reconstructed under the auspices of the ERP in Jordan are found. The realization of differences makes it necessary upon us to devise such approaches to test construction as would take into account the authentic differences and thus allow for making legitimate comparisons between different sets of test scores.

Having taken into account the classroom interactions and other mediating process-variables, the acid test of the effectiveness of reform inputs lies in student achievements. Thus the improvement in the quality of school product or lack thereof reflected by the magnitude of gains in student achievement levels will be the ultimate measure of the success of education reform program in Jordan. Now, to measure the gain or improvement in such abstract constructs as achievement in school subjects, one needs to establish certain baseline parameters against which subsequently estimated gain parameters can be compared.

## Sequential Steps

Test construction is not an easy task, nor is it a one-time activity. To construct valid and reliable achievement tests needs careful advance planning, perseverance, endurance and long-term commitment to sustained hard work on the part of a team of specialists trained and experienced in testing and evaluation. What follows is a brief description of various steps that were taken to develop achievement tests for the purpose of tracing the impact of reform interventions and establishing the current status of student achievement levels in Jordan.

### I Mapping the Old and the New Curriculums

This involved listing the objectives defined by the old and the new curriculums for a particular subject and grade side by side in parallel columns and establishing one to one correspondence: This resulted in three subsets of broad curricular objectives.

1. Objectives common to both the new and the old curriculums.
2. Objectives unique to the new curriculum.
3. Objectives unique to the old curriculum.

The following figure is a diagrammatic illustration of a mapping outcome.

## Grade 4 Math Curriculum

	Old Objectives		New Objective	
		Common set		
	a	_____	a	
	b	_____	b	
	.	_____	.	
	.	_____	.	
	.	_____	.	
	k	_____	k	
Unique to Old	x		o	
	y		o	
	z		o	
	o		m	
	o		n	Unique to New
	o		q	

- II      Determining Relative Emphasis Given to Each Objective in Both Old and New Curricula.
- III     Matching the Weights Given to Various Objectives in Both Curricula.
- IV     Listing the Content Substance (Topics, Concepts, Skills, Facts, etc.) Implied and / or Explicitly Covered by Each Curriculum (Old and New).
- V      Establishing Correspondence and Identifying Common and Unique Areas.
- VI     Determining Relative Emphasis Given to Each Area in Both the Curricula.
- VII    Preparing a Table of Specifications for the Common Objectives and Content Areas. (This constituted the major portion of the achievement test).
- VIII   Preparing Separate Tables of Specifications for the Unique Parts.

- IX Determining Relative Importance of the Common and Unique Parts in the Test.
- X Constructing Test Items Using Most Appropriate Item Format for Testing Each Objective and Content. (At least 4-times the Required Number of Items was the target).
- XI Revising the Test Items and Subjecting them to Independent Expert Reviews.
- XII Compiling the Test in a Suitable Format and Writing Appropriate Instructions.
- XIII Trying the Test on a Representative Sample Under Normal Conditions.
- XIV Analyzing the Test Data. Conducting Item Analyses.
- XV Revising the Test and Modifying the Test Items in Light of the Item Analyses Information.
- XVI Trying the Revised Test Once Again on a Different Sample. Conducting Item Analyses and Establishing Item Statistics (Psychometric Properties).
- XVII Selecting the Desired Number of Sound Items for Each Cell of the Table of Specifications and Compiling the Test Giving it a Final Shape.
- XVIII Administering the Test to The Target Samples and Determining its Psychometric Properties.

**Test Development Teams:** Four specialized teams of experts (one each for Arabic, Math, Science and Life Skills) were formed. Each team was charged with the responsibility of developing the achievement test following the general guidelines and underlying criteria given to them and thoroughly discussed with each team.

Each team consisted of university professors, subject supervisors, and teachers of the subject and met several times to discuss the procedures to be followed by each person at each step with an NCERD coordinator. To start with, each member of each team underwent a through orientation and worked at each step in close cooperation with other members of the team. Output was reviewed, discussed, and revised at every step of the test development process under close supervision of and in full cooperation with on NCERD coordinator.

When the agreement was reached on the table of specifications determining the relative emphasis to be given to various content and skill domains, items were constructed to measure the specific objectives to be tested. A variety of item formats were used. Each item format was selected on the basis of its effectiveness in measuring specific an objective and content. Each item of the item pool created for a test was independently reviewed by each member of the team and then discussed by the whole team with the NCERD coordinator.

The pool of items thus created was compiled and about 30 copies were produced. An item rating scale was designed to rate each item on multiple criteria. A Panel of 8 to 12 expert judges including testing supervisors, school teachers from public and private schools, principals and directors was invited at the NCERD. Each item of each pool was first independently rated by each judge on the multiple criteria rating scale and later on discussed with the test construction team in a combined session organized at the NCERD. The test construction team revised the items and modified the item pool in the light of the comments and suggestions of these independent reviews.

**Piloting the Tests:** The pool of items for each test was randomly divided into several forms which were administered to a pilot sample following a matrix sampling design. Each item was administered to at least 200 students scattered over 12 different schools representing different education authorities, area of location, and student gender.

**Analysis of the Pilot Data:** Each form of a test was subjected to classical test theory based item analysis procedures. Every item was scrutinized by a team of two or three experts including an experienced psychometrician and all the information surrounding the item and the test form was assessed and reassessed. The difficulty indices, the item remainder correlations, value of Alpha if item were deleted, and distracter analysis data were considered simultaneously, and the language, and wording , and format of the item along with its response alternatives and response formats were reexamined.

Lists of items were prepared indicating complete information on all the idices and criterion ratings on each item. Criteria were developed for the final selection of test items which were used to select test items for the final form of the test. The final form of each test although retained a much smaller number of the original items, maintained the relative weights of the original tables of specifications and thus conserved the content and curricular validity of each test. The final form was

compiled jointly by the original test construction team and the NCERD coordinator.

### **Administration of the Tests**

By the time the 4th grade achievement tests were ready to be used the academic year 1992-93 had already set in. The students who were promoted to Grade 5 from Grade 4 had been about a month and a half in the 5th grade. In November 1992, the 4th grade Arabic, Math, and Science achievement tests were administered by well-trained teams of test administrators in 205 schools of the national sample to 5th grade students at the beginning of the 5th grade as entry level tests.

The tests were scored, data were entered into the computers and cleaned and analyzed. Complete data from a national sample of about 5000 students on these achievement test became available. These tests were subjected to item analysis procedures again.

A few items (one or two) in each test were found to behave aberrantly. The errors were identified and duly rectified for the final administration of tests to 4th grade students in May 1993.

What follows is a brief outline of each instrument presented in the form of its Table of Specifications.

Based on the results of item analysis and test statistics, and taking into considerations different weights given to different content/ skill levels, 36 items were selected to compose the final version of Literacy Test (Arabic Language), 50 items were selected to compose the final version of the Numeracy Test, and 60 items were selected for the final version of the Science test. Tables 8, 9, and 10 respectively present the Table of Specifications for Arabic, Math and Science Achievement tests.

**Table 8****Table of Specifications for the Literacy (Arabic Language) Test**

Content \ Skill	Comprehension	Application	Synthesis	Evaluation	Total
Reading	5	2	2	1	10
Composition	0	1	4		5
Writing & Dictation	0	2	1	1	4
Grammar	10	6	1	0	17
<b>Total</b>	<b>15</b>	<b>11</b>	<b>8</b>	<b>2</b>	<b>36</b>

**Table 9****Table of Specifications for the Numeracy (Math) Test**

Content \ Skill	Conceptual Understanding	Application	Problem Solving	Total
Number	7	6	1	14
Operation	4	4	5	13
Fractions	2	4	2	8
Decimals	2	1	0	3
Geometry	4	0	0	4
Measurement	3	4	1	8
<b>Total</b>	<b>22</b>	<b>19</b>	<b>9</b>	<b>50</b>

**Table 10****Table of Specifications for the (Science) Test**

Content \ Skill	Knowledge	Application	Intergration	Total
Life Sciences	12	4	0	16
Physical Sciences	7	11	7	25
Earth and Space Sciences	6	3	10	19
<b>Total</b>	<b>25</b>	<b>18</b>	<b>17</b>	<b>60</b>

The same procedure as used in developing achievement tests (Literacy and Numeracy and Science) was adopted in developing the Life Skills Test. The final version of the Test comprised of 37 items measuring the domains agreed upon by the workshop's participants (Paris, Feb.22 - 26, 1993). The items measured basic knowledge, understanding and application of principles related to health and nutrition, environment, daily life and safety, and civic education. Table 11 presents the table of specifications for the final version of the Life Skills instrument.

**Table 11****Table of Specifications for Life Skills Test**

Content \ Skill	Knowledge	Application	Total
Health & Nutritiors	6	7	13
Environment	5	7	12
Daily Life	2	2	4
Civic Education	3	5	8
<b>Total</b>	<b>16</b>	<b>21</b>	<b>37</b>

### **Data Collection**

Standardized procedures for instructing students and establishing conditions for testing were developed by NCERD and pilot tested. To ensure reasonably accurate and generalizable information capable of generating sufficiently valid conclusions, a team of 13 area coordinators and a team of 42 test administrators was trained and charged with the responsibility of the test administration and data collection. Field work was supervised by the Jordanian National Task Force to ensure uniform standards of test administration. Data collection accomplished in two phases, each lasting for one week.

- 1) Phase 1 (May 10th - May 17th, 1993).  
During this phase, Students questionnaire and Life Skills Test were administered to the students in the sampled schools. Teacher's, Principal's and Family Background Questionnaires distributed were also during this phase.
- 2) Phase II (May 18th - May 25th, 1993).  
During this phase, achievement tests of Literacy, Numeracy, and Science were administered to the students in the sampled schools.

## **Data Analysis**

**Stock-taking and Checking:** As the completed tests started coming back from the field they were checked against the a priori prepared lists of students in the targeted grade 4 class in each school. Each student and school was given a specific code number. Any delay in the arrival of test packages from any test administrator was tracked and 100% recovery was ensured.

**Preparing Codebooks :** A Detailed code-book defining each variable, variable labels, range of values, and value labels was prepared for each questionnaire and achievement test.

**Data Entry Form:** Data entry forms were prepared for each test and questionnaire using SPSS/PC's Data Entry Procedure, and data were entered directly from the tests and questionnaires under the supervision of NCERD staff. Data entry errors were controlled by selecting randomly 5% of the protocols entered by each individual daily and checking them by independent double entry. If any person's work produced more than 2% errors the whole batch was entered against the aforesaid validation procedure. When the data entry was completed different data files were merged together and preliminary statistical analyses were performed on each variable in each instrument with the purpose of detecting any type of abnormalities in the data. If a particular error could not be tracked down and corrected directly in the raw data files, it was referred back to the test or questionnaire and finally resolved.

**Analysing the Data:** Due to diversified nature of variables and target populations a three-stage data analysis and interpretation strategy was developed. The first stage concerned with verificatory, validity and descriptive analysis of each instrument. The statistical procedures applied at this stage included, frequency analysis, decomposition of ramified variables and scales, interitem correlations, scale and subscale

reliability indices, distributional properties of various types of scale and subscale scores, descriptive statistics and variance analyses.

The purpose of the first set of analyses was to establish the reliability and other properties of interest of each instrument, while that the descriptive and variation analyses was to establish the current standards of achievement and current norms of certain practices at the national, regional (governorate) education authority, student gender, and location (rural / urban) levels, as well as, to study the differences in the performance of students in different groups of schools defined by the preceding set of exogenous factors.

The second-stage analyses involved deriving indicator indices and parsimonious sets of latent concepts in the major domains of variables and studying relationships between student achievement and important family background, parental, community, school and classroom factors.

The third-stage of data analysis involved searching for explanatory statistical models using Multilevel Regression and Linear Structural Covariance Analysis models.

At present, the first stage has been accomplished, i.e., the achievement test scores have been analyzed and students' achievement in different subjects has been described at variables levels of aggregation (national, governorate, education authority, student gender and school location).

Second-stage analyses (correctional analyses conducted to study which variables which variables in each domain influence student achievement and to what extent) are under way. Home background and parental domain has been studied and the variables that significantly correlate with student achievement have been identified.

## Results

### *Arabic Test*

#### Performance at the National Level

The Arabic language test was administrated to 4908 fourth grade students in 205 schools representing all educational authorities and Governorates in Jordan. Table 12 presents the performance of students on the whole test and its subscales.

**Table 12**

#### **Average Percent Correct on the Arabic Language Test**

Subscale	Average Percent Correct	Standard Error	Lower 10 %		Upper 10 %	
			Min	Max	Min	Max
<b>Total Score</b>	<b>54.19</b>	<b>0.33</b>	<b>0</b>	<b>21.85</b>	<b>83.8</b>	<b>100</b>
<b>Content</b>						
Reading	62.01	0.34	0	26.67	91.00	100
Comprehension	50.74	0.43	0	7.50	90.00	100
Dictation	55.06	0.37	0	14.58	85.42	100
Grammar	50.4	0.35	0	17.65	82.35	100
<b>Skill</b>						
Understanding	57.94	0.34	0	25.00	89.58	100
Application	53.11	0.34	0	20.10	84.31	100
Synthesis	50.39	0.39	0	11.79	87.14	100

It is clear from Table 12 that the average percent correct on the test is 54.19. Students' performance on the content subscales is in the following order: Reading, Dictation, Comprehension, and Grammar. On skills, performance was the highest on item related to understanding, then application, and finally on syntheses. The performance of the upper 10% of students, ranges on the total score from 84-100, from 91 to 100 on Reading subscale, from 85.4 to 100 on Dictation subscale, and from 82.3 to 100 on Grammar subscale. On the other hand, the performance of the lower 10% ranges from 0 to 22 on the total score, from 0 to 27 on Reading subscale, from 0 to 7.5 on Comprehension subscale, from 0 to 14.6 on Dictation subscale, and from 0 to 17.7 on Grammar subscale.

## Performance Across Governorates

Tables 13 gives the average percent correct score on the test as a whole for each Governorate while Figure 1 presents their baragraph.

**Table 13**

**Average Percent Correct Score, Standard Error, and 95% CI  
on the Arabic Test in Each Governorate**

Governorate	Average % correct	Standard Error	95% confidence interval
Amman	60.77	0.55	59.67 - 61.87
Zarqa	55.34	0.85	53.64 - 57.04
Balqa	56.87	1.43	54.01 - 59.73
Irbid	49.4	0.67	48.06 - 50.74
Ma'raq	50.93	0.99	48.95 - 52.91
Karak	39.35	1.28	36.79 - 41.91
Tafleh	48.37	2.01	44.35 - 52.39
Ma'an	51.29	1.18	48.93 - 53.65

**Figure 1**

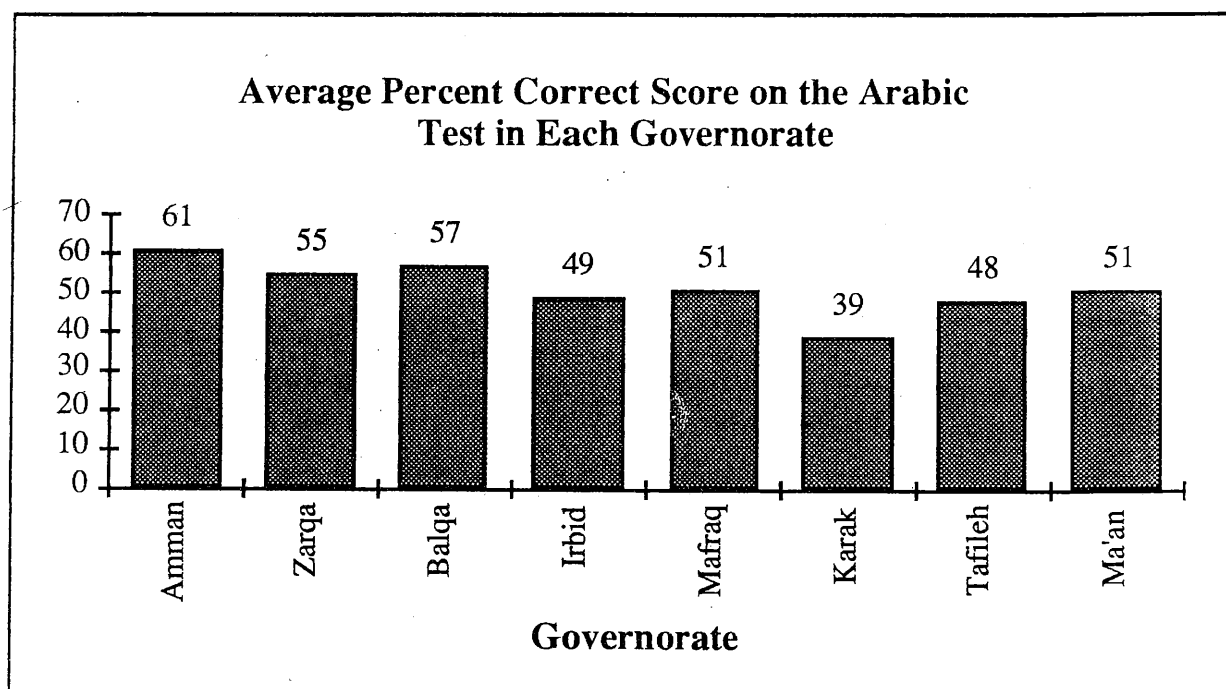


Figure 1 clearly shows that students in Amman Governorate outperformed the students in other governorates ( $\alpha .05$ ) except for Balqa. The lowest performance was for students of Karak Governorate.

On each content subscale also (Table 14), the performance of capital Amman students was the highest while Karak Governorate students was the lowest.

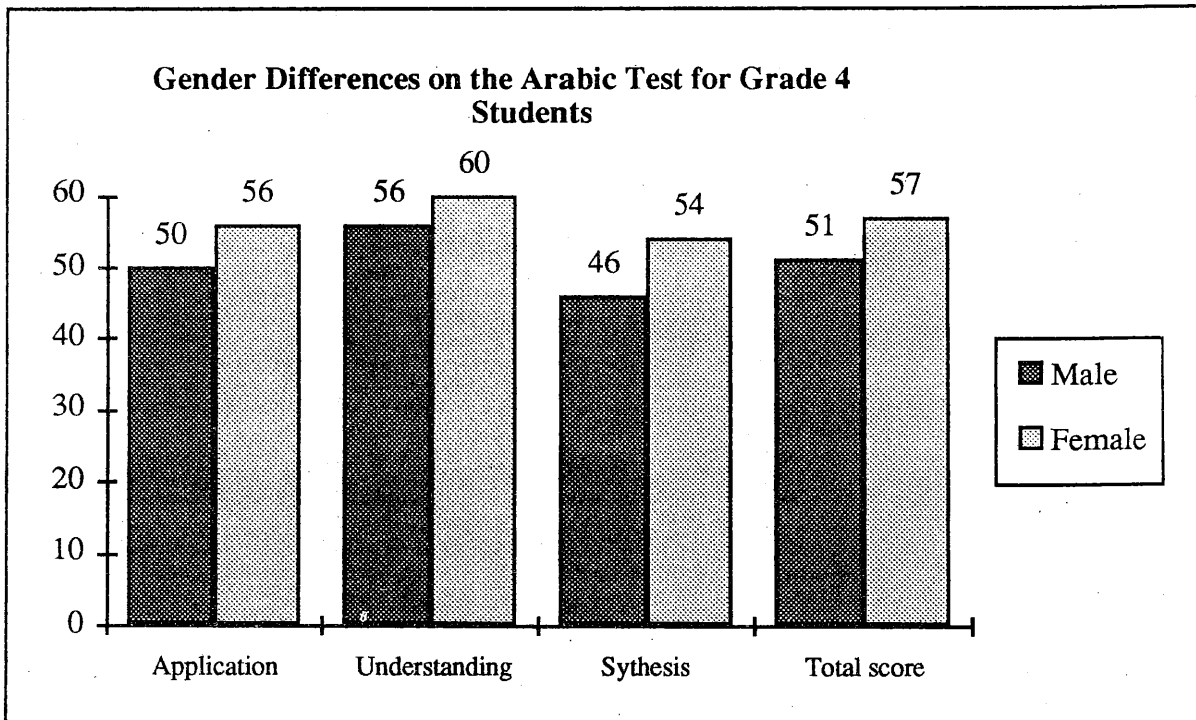
**Table 14**  
**Average Percent Correct and Standard Error on Arabic Test**  
**Subscales in Each Governorate**

	Amman	Zarqa	Balqa	Irbid	Mafraq	Karak	Tafileh	Ma'an
<b>Reading</b>								
Average % Correct	67.40	65.02	62.85	58.31	58.33	48.09	51.76	59.22
Standard Error	.57	.88	1.57	.71	1.10	1.40	2.18	1.22
<b>Comprehension</b>								
Average % Correct	58.38	55.21	55.04	43.05	49.67	31.86	51.73	43.05
Standard Error	.70	1.07	1.97	.92	1.41	1.70	3.00	1.51
<b>Dictation</b>								
Average % Correct	59.71	55.89	59.39	52.09	54.94	39.31	48.74	52.61
Standard Error	.61	.96	1.55	.78	1.19	1.62	2.35	1.24
<b>Grammar</b>								
Average % Correct	57.82	49.55	53.29	45.39	46.00	36.41	45.30	48.73
Standard Error	.59	.90	1.58	.71	1.05	1.37	2.05	1.33

### Performance by Sex

Results (Table 15) indicated that the performance of female students was statistically significantly higher on the total score and on all content subscales than that of the male students. The difference between the means of male and female students was the highest on comprehension subscale (9%). A graphic display of the differential performance of male and female students on each subscale and total score of the Arabic Test is presented by Figure 2.

**Figure 2**



**Table 15**

**Average Percent Correct on Arabic Test Subscales for Male and Female Students**

	Average % Correct	Standard Error	95% confidence Interval
<b>Total Score</b>			
Males	51.23	.49	50.25 - 52.21
Females	56.75	.43	55.89 - 57.61
<b>Reading</b>			
Males	59.49	.51	58.47 - 60.51
Females	64.19	.46	63.27 - 65.11
<b>Comprehension</b>			
Males	45.89	.64	44.61 - 47.17
Females	54.93	.57	53.79 - 56.07
<b>Dictation</b>			
Males	51.28	.56	50.16 - 52.4
Females	58.33	.48	57.37 - 59.29
<b>Grammar</b>			
Males	47.93	.52	46.89 - 48.97
Females	52.54	.47	51.6 - 53.48

The order of performance from highest to lowest was almost the same for both girls and boys. For boys, the order is as follows: Reading,

Dictation, Grammar, and Comprehension. For girls, the order is as follows: Reading, Dictation, Comprehension, and Grammar.

### Performance by Educational Authority

Table 16 shows the performance of the 4th grade students on the Arabic Language Test by educational authority, while Figure 3 gives its graphic display.

**Table 16**

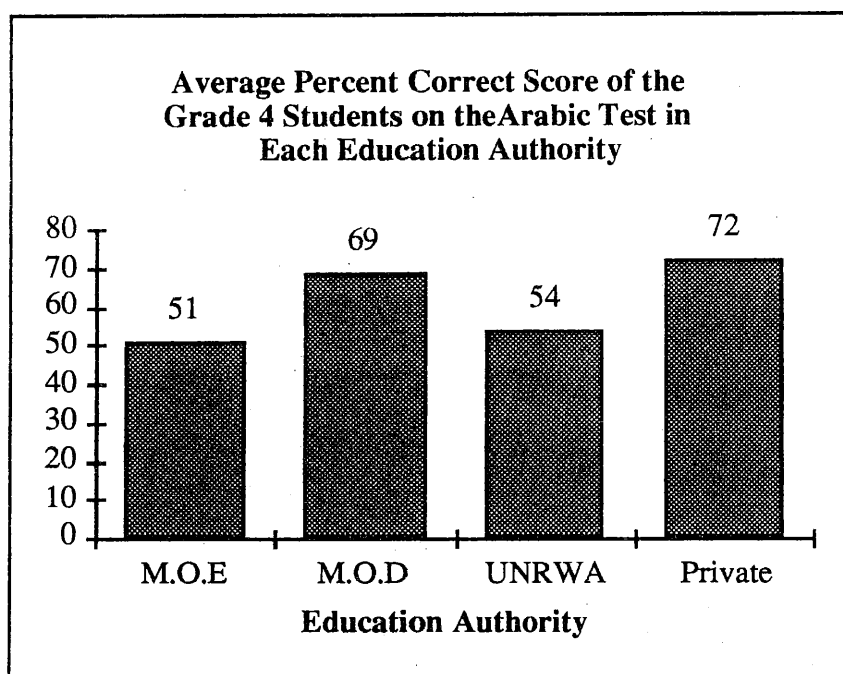
#### **Average Percent Correct Score Standard Error and 95% CI in Each Education Authority**

Authority	MOE	MOD	UNRWA Schools	Private Schools
<b>Total Score</b>				
Average % Correct	50.9	69.14	54.5	72.49
Standard Error	0.36	1.45	1.12	0.81
95% CI	50.18 - 51.62	66.24 - 72.04	52.26 - 56.74	70.87 - 74.11
<b>Reading</b>				
Average % Correct	58.78	75.96	65.24	77.87
Standard Error	0.39	1.37	1.18	0.84
95% CI	58 - 59.56	73.22 - 78.7	62.88 - 67.6	76.19 - 79.55
<b>Comprehension</b>				
Average % Correct	47.11	64.15	51.61	71.42
Standard Error	0.49	1.88	1.45	1.03
95% CI	46.13 - 48.09	60.39 - 67.91	48.71 - 54.51	69.36 - 73.48
<b>Dictation</b>				
Average % Correct	52.81	63.49	54.9	68.41
Standard Error	0.42	1.73	1.32	0.95
95% CI	51.97 - 53.65	60.03 - 66.95	52.26 - 57.54	66.51 - 70.31
<b>Grammar</b>				
Average % Correct	46.92	67.91	48.94	70.61
Standard Error	0.38	1.77	1.2	0.91
95% CI	46.16 - 47.68	64.37 - 71.45	46.54 - 51.34	68.79 - 72.43

A glance at Figure 3, reveals that the Private schools' students and the Ministry of Defense (MOD) schools' students outperformed both UNRWA & Ministry of Education (MOE) students on the test as a whole. Moreover, inspection of subscale means in Table 15 reveals that the performance of UNRWA students was statistically significantly higher than that of the Ministry of Education students. On all content subscales, private schools students and Ministry of Defense schools' students outperformed other authorities. In all authorities, students

performance was the highest on Reading subscale and the lowest on Grammar subscale for both Ministry of Education and UNRWA students, whereas the lowest performance for Private school students and Ministry of Defense students was on Dictation subscale.

**Figure 3**



### **Performance by School Location**

Table 17 and Figure 4 show the comparative performance of the 4th grade students of the urban and rural schools on various subscales of the Arabic Test.

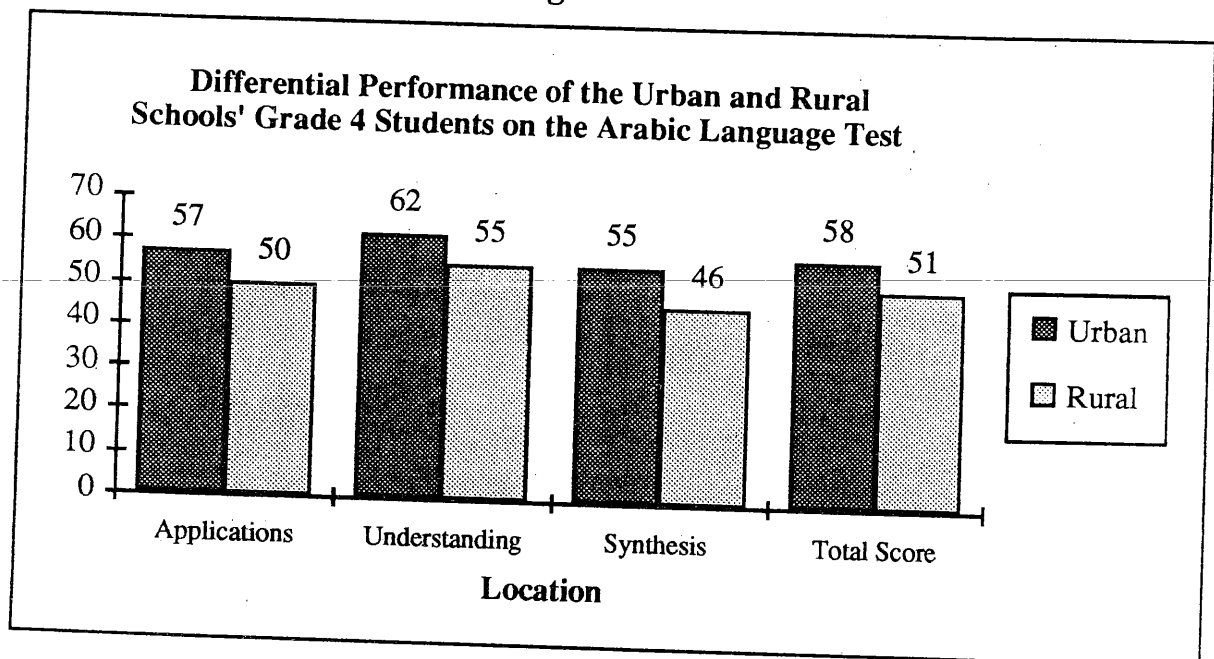
**Table 17**

**Average Percent Correct on Arabic Test Subscales in Urban and Rural School Location**

Location	Average Percent Correct	Standard Error	95% CI
<b>Total Score</b>			
Urban	58.19	.49	57.21 - 59.17
Rural	50.77	.43	49.91 - 51.63
<b>Reading</b>			
Urban	65.46	.51	64.42 - 66.5
Rural	59.07	.46	58.15 - 59.99
<b>Comprehension</b>			
Urban	55.68	.63	54.42 - 56.94
Rural	46.51	.58	45.35 - 47.67
<b>Dictation</b>			
Urban	57.70	.54	56.62 - 58.78
Rural	52.81	.50	51.81 - 53.81
<b>Grammar</b>			
Urban	54.77	.53	53.71 - 55.83
Rural	46.66	.45	45.76 - 47.56

It can be seen from Table (17) and Figure 4 that urban school students outperformed the rural school students on the total score as well as on all content subscales with a difference of 8%. The order of performance on the subscales was almost the same for urban and rural students. That is, their performance was the highest on Reading subscale and the lowest on Grammar subscale.

**Figure 4**



# Science Test

## Performance at the National Level

The Science Test was administrated to 2412 fourth grade students in 205 schools representing all governorates and educational authorities in Jordan.

Table (18) presents the average percent correct score on the test and its subscales.

**Table 18**

### **Average Percent Correct Score, Standard Error, Minimum, and Maximum Scores on the Science Test Subscales**

Subscale	Average percent correct	Standard Error	Lower 10 %		Upper 10 %	
			Min	Max	Min	Max
<b>Total Score</b>	<b>41.58</b>	<b>.37</b>	<b>.00</b>	<b>21.67</b>	<b>68.33</b>	<b>95.00</b>
<b>Content</b>						
Human	41.05	.46	.00	11.11	77.78	100.00
Water	34.83	.50	.00	.00	60.00	100.00
Physical	42.53	.42	.00	17.86	71.43	100.00
Environment	47.31	.55	.00	14.29	85.71	100.00
Universe	42.02	.60	.00	.00	75.00	100.00
Earth	41.46	.65	.00	.00	100.00	100.00
Plants	34.21	.53	.00	.00	75.00	100.00
<b>Skill</b>						
Knowledge	41.02	.39	.00	16.00	68.00	100.00
Understanding	43.82	.42	.00	17.39	73.91	100.00
Application	38.45	.44	.00	16.67	66.67	100.00

It can be seen from Table (18) that the average percent correct score on the test is approximately 42%. That is, students on the average were able to respond correctly to 42% of the test items. On test subscales, the average percent correct ranges from 34.2% (plants subscale) to 47.3 (Environment subscale). When we examine the performance of the lower and upper 10% of students, we can see that on four of the content subscales all students in the lower 10% failed to answer any of the test items correctly. The performance of upper 10% of students ranges from 60 to 100%.

## Performance Across Governorates

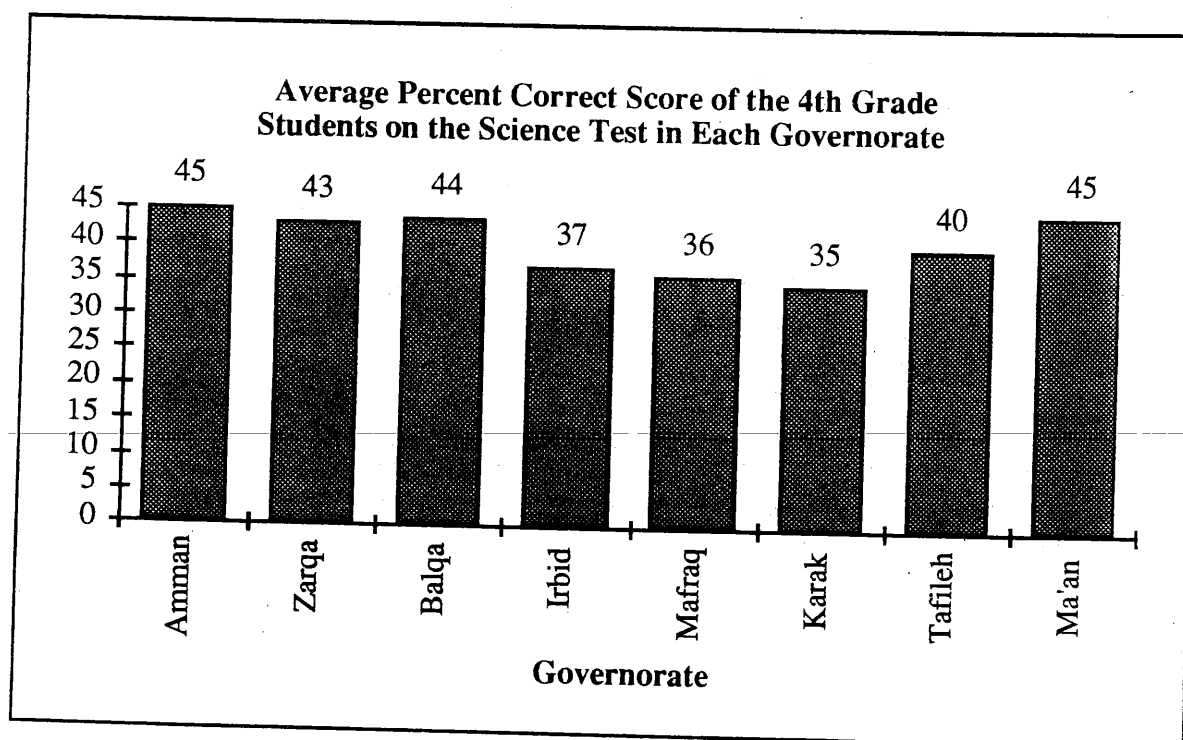
Table 19 shows the performance of 4th grade students on the Science Test in each Governorate, and Figure 5 gives a graphic display of intergovernorate comparisons.

**Table 19**

**Average Percent Correct Score, Standard Error, and 95% CI on the Science Test in Each Governorate**

Governorate	Average % Correct	Standard Deviation	Standard Error	95% Confidence Interval
Amman	45.17	22.37	0.78	43.61 - 46.73
Zarqa	44.03	23.36	1.2	41.63 - 46.43
Balqa	38.83	23.81	2.52	33.79 - 43.87
Irbid	36.64	22.23	0.96	34.72 - 38.56
Mafraq	32.54	19.16	1.37	29.80 - 35.28
Karak	36.73	22.15	1.91	32.91 - 40.55
Tafleh	44.44	18.98	2.71	39.02 - 49.86
Ma'an	41.71	22.17	1.59	38.53 - 44.89

**Figure 5**



**Table 20****Average Percent Correct Score and Standard Error of Each Science Test Subscales in Each Governorate**

Governorate	Amman	Zarqa	Balqa	Irbid	Mafraq	Karak	Tafileh	Ma'an
<b>Water</b>								
Average % Correct	35.77	38.79	37.30	31.53	35.23	28.51	30.20	36.31
Standard Deviation	25.48	25.48	24.34	22.74	23.55	22.79	19.63	22.86
Standard Error	.89	1.31	2.58	.98	1.68	1.97	2.80	1.64
<b>Physical</b>								
Average % Correct	46.40	43.80	44.06	37.47	37.33	34.97	38.92	48.28
Standard Deviation	21.18	21.65	20.52	18.89	17.92	16.03	12.03	24.01
Standard Error	.74	1.11	2.17	.81	1.28	1.38	1.72	1.72
<b>Environment</b>								
Average % Correct	51.88	47.61	51.69	40.88	41.84	40.72	48.98	52.82
Standard Deviation	26.65	27.00	28.12	26.33	23.47	26.65	24.57	27.27
Standard Error	.93	1.39	2.98	1.13	1.67	2.30	3.51	1.95
<b>Universal</b>								
Average % Correct	46.47	41.16	52.81	39.37	34.39	31.34	42.86	42.05
Standard Deviation	29.92	28.05	33.17	28.91	27.33	27.71	25.52	28.40
Standard Error	1.04	1.44	3.52	1.24	1.95	2.39	3.65	2.03
<b>Earth</b>								
Average % Correct	46.62	44.68	41.95	35.49	37.06	32.84	36.73	41.20
Standard Deviation	33.16	33.03	33.15	30.52	30.45	28.90	31.36	29.23
Standard Error	1.15	1.70	3.51	1.31	2.17	2.50	4.48	2.09
<b>Plants</b>								
Average % Correct	36.35	32.52	39.04	31.93	32.23	31.16	30.10	37.69
Standard Deviation	26.95	25.58	26.10	26.05	26.28	21.77	21.03	27.62
Standard Error	.94	1.31	2.77	1.12	1.87	1.88	3.00	1.98

Examination of Figure 5 reveals that the performance of students in Amman, Zarqa, and Tafileh governorates was better than that of the students in Irbid and Mafraq governorates.

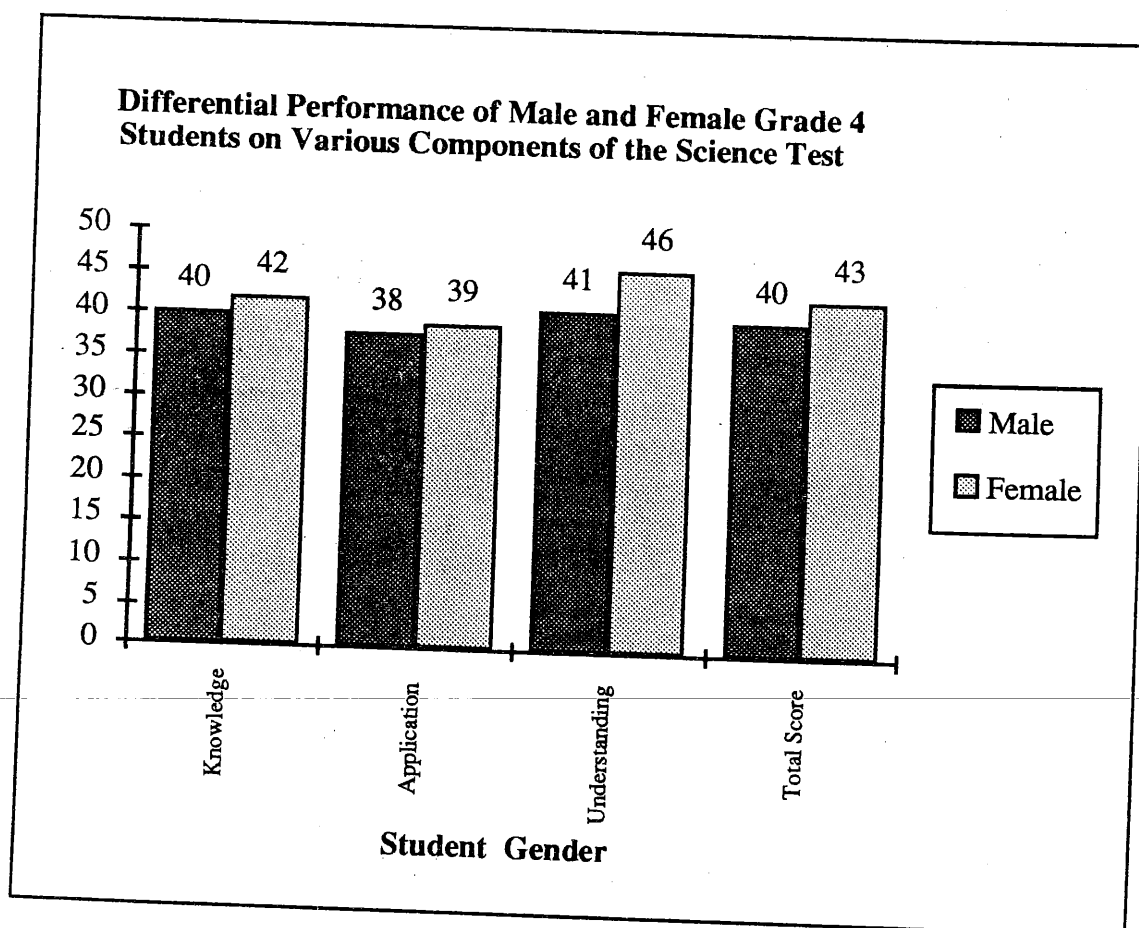
When we take content subscales of the test into consideration (Table 20), we notice that for all governorates the highest performance was on the Environment subscale and the lowest performance was on Water and Plants subscales. On Water subscale, students in Zarqa and Balqa Governorates scored higher than those in other Governorates. On Physical Sciences subscale, students in Ma'an and Zarqa governorates scored higher than students in other governorates. On both Environment and Plants subscales, performance of students in Amman, Balqa, and

Ma'an Governorates was higher than that of students in other Governorates. Finally, on Universe and Earth subscales students in Amman Governorate scored higher than did students in other governorates.

### Gender Differences on the Science Test

Results displayed by Figure 6 (detailed in Table 20) indicate that female students scored statistically significantly higher on the total test, and on Water, Environment, and Universe subscales than the male students. Differences on other subscales were not statistically different. For both groups the performance was the highest on Environment subscale and the lowest on Water & Plants subscales.

Figure 6



**Table 21****Average Percent Correct Score, Standard Error, and 95% CI of Male and Female Students on the Science Test Subscales**

Location	Average % Correct	Standard Deviation	Standard Error	95% Confidence Interval
<b>Total Score</b>				
Males	40.18	19.27	.58	39.02 - 41.34
Females	42.87	17.13	.48	41.91 - 43.83
<b>Content</b>				
<b>Human</b>				
Males	40.55	23.81	.72	39.11 - 41.99
Females	41.53	21.50	.60	40.33 - 42.73
<b>Water</b>				
Males	33.00	24.58	.74	31.52 - 34.48
Females	36.47	24.02	.67	35.13 - 37.81
<b>Physical</b>				
Males	41.57	22.29	.67	40.23 - 42.91
Females	43.43	19.32	.54	42.35 - 44.51
<b>Environment</b>				
Males	43.91	27.63	.83	42.25 - 45.57
Females	50.33	25.84	.72	48.89 - 51.77
<b>Universe</b>				
Males	38.52	29.25	.88	36.76 - 40.28
Females	45.17	29.11	.81	43.55 - 46.79
<b>Earth</b>				
Males	41.13	31.69	.95	39.23 - 43.03
Females	41.84	32.44	.90	40.04 - 43.64
<b>Plants</b>				
Males	33.03	26.05	.78	31.47 - 34.59
Females	35.34	26.32	.73	33.88 - 36.80

**Performance by Education Authority**

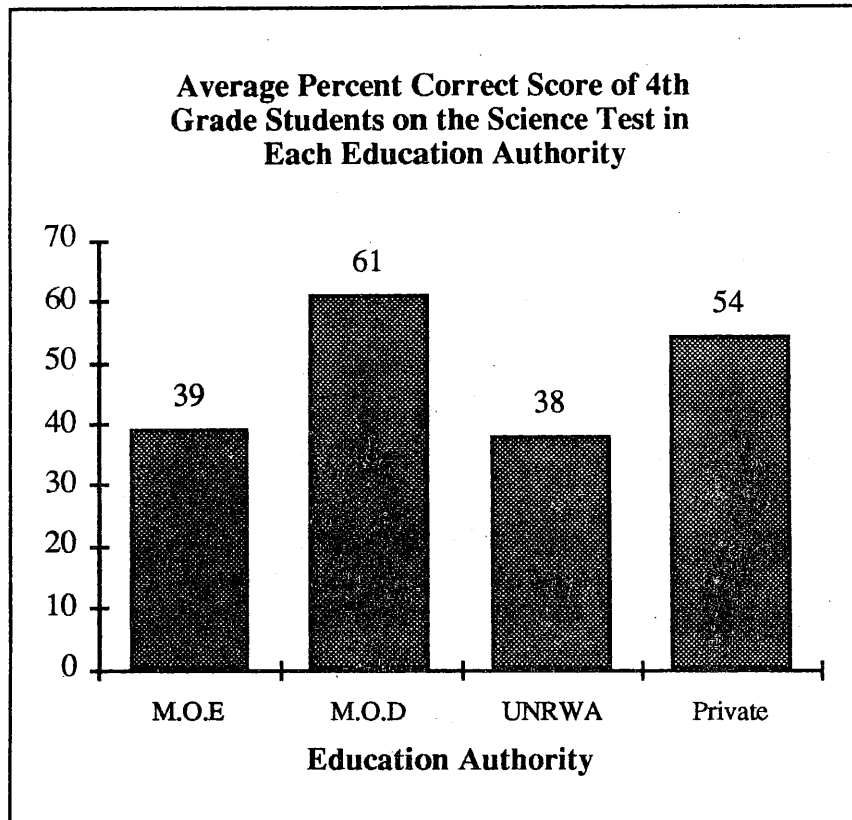
Table 22 shows related statistics and Figure 7 gives a graphic display of the performance of 4th grade students on the Test and its subscales by education authority.

Table 22

**Average Percent Correct Score, Standard Error, and 95% CI  
on the Science Test Subscales in Each Education Authority**

	MOE	MOD	UNRWA Schools	Private Schools
<b>Total Score</b>				
Average % Correct	39.34	61.39	38.32	53.52
Standard Deviation	17.14	18.45	16.64	18.21
Standard Error	0.4	2	1.13	1.12
95% CI	38.54 - 41.14	57.39-65.39	36.06-40.58	51.28-55.76
<b>content</b>				
<b>Human</b>				
Average % Correct	38.6	58.69	37.65	55.26
Standard Deviation	21.63	21.04	20.32	23.86
Standard Error	0.5	2.28	1.38	1.47
95% CI	37.6-39.6	54.13-63.25	34.89-40.41	52.32-58.2
<b>Water</b>				
Average % Correct	33.86	44.71	35.28	38.11
Standard Deviation	24.34	24.62	24.59	23.35
Standard Error	0.57	2.67	1.67	1.44
95% CI	32.72-35	39.37-50.05	31.94-38.62	35.23-40.99
<b>Physical</b>				
Average % Correct	40.03	67.73	38.41	55.22
Standard Deviation	19.48	22.11	19.69	19.93
Standard Error	0.45	2.4	1.34	1.23
95% CI	39.13-40.93	62.93-72.53	35.73-41.09	52.76-57.68
<b>Environment</b>				
Average % Correct	44.64	68.07	43.72	62.28
Standard Deviation	26.32	23.08	24.5	25.96
Standard Error	0.61	2.5	1.67	1.6
95% CI	43.42-45.86	63.07-73.07	40.38-47.06	59.08-65.48
<b>Universe</b>				
Average % Correct	40.04	52.35	38.77	55.21
Standard Deviation	28.84	29.28	29.8	28.66
Standard Error	0.67	3.18	2.03	1.76
95% CI	38.7-41.38	45.99-58.71	34.71-42.83	51.69-58.73
<b>Earth</b>				
Average % Correct	39.58	53.73	39.66	52.15
Standard Deviation	31.64	29.59	30.94	34.33
Standard Error	0.74	3.21	2.11	2.11
95% CI	38.1-41.06	47.31-60.15	35.44-43.88	47.93-56.37
<b>Plants</b>				
Average % Correct	32.9	47.06	32.06	41
Standard Deviation	25.64	25.12	23.73	30.16
Standard Error	0.6	2.72	1.61	1.86
95% CI	31.7-34.1	41.62-52.5	28.84-35.28	37.28-44.72

**Figure 7**



Results presented in Table (22) and graphed in Figure 7 show that Ministry of Defense (MOD) students scored the highest on the Science Test than students in all. Private school students scored higher than students in Ministry of Education (MOE) schools and UNRWA schools. Finally, the performance of MOE & UNRWA students was not statistically different.

On all the subscales of the Science Test, MOD and Private school students scored higher than the students of both MOE and UNRWA schools. No significant differences were found between MOE and UNRWA students on any of the subscales. Although MOD students scored higher than Private school students on almost all the subscales, the differences, however, were not statistically different except on physical sciences subscale. Finally, performance was the highest on Environment subscale for students in different authorities and the lowest on water and plants subscales.

### Science Test Performances in Urban and Rural Schools

Table 23 and Figure 8 present the levels of performance of the 4th grade students by school location.

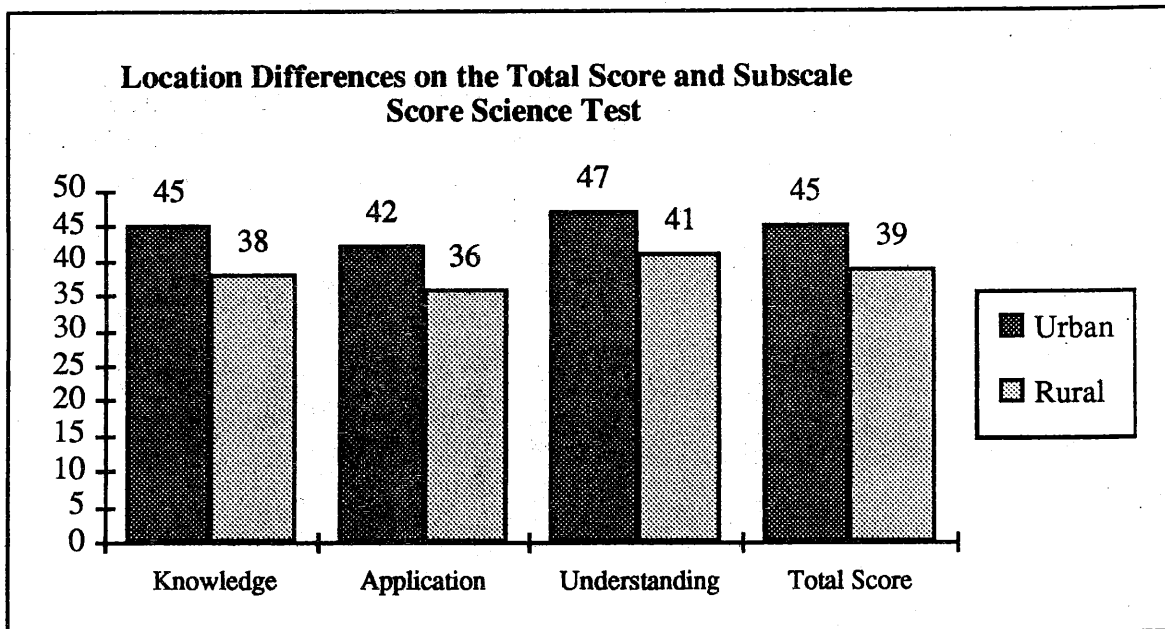
**Table 23**

**Average Percent Correct Scores on the Science Test Subscales of the 4th Grade Students in Urban and Rural Schools**

Location	Average % Correct	Standard Deviation	Standard Error	95% CI
<b>Total Score</b>				
Urban	44.91	18.76	.56	43.79 - 46.03
Rural	38.66	17.22	.48	37.70 - 39.62
<b>Content</b>				
<b>Human</b>				
Urban	45.25	22.82	.68	43.89 - 46.61
Rural	37.36	21.77	.61	36.14 - 38.58
<b>Water</b>				
Urban	37.28	24.92	.74	35.80 - 38.76
Rural	32.68	23.67	.66	31.36 - 34.00
<b>Physical</b>				
Urban	45.64	21.22	.63	44.38 - 46.90
Rural	39.76	20.00	.56	38.67 - 40.91
<b>Environment</b>				
Urban	51.72	27.34	.81	50.10 - 53.34
Rural	43.45	25.87	.72	42.01 - 44.89
<b>Universe</b>				
Urban	45.03	29.48	.88	43.27 - 46.79
Rural	39.38	29.02	.81	37.76 - 41.00
<b>Earth</b>				
Urban	44.96	33.42	1.00	42.96 - 46.96
Rural	38.39	30.62	.85	36.69 - 40.09
<b>Plants</b>				
Urban	36.42	27.51	.82	34.78 - 38.06
Rural	32.28	24.87	.69	30.90 - 33.66

It is clear from Table 23 and Figure 8 that the performance of urban school students on the total score and on all content subscales was statistically significantly higher than that of the rural school students. For both urban and rural students, the highest performance was on Environment subscale and the lowest on Water and Plants subscales.

**Figure 8**



# Math Test

## Performance on the National Level

The Mathematics Test was administered to 2428 fourth grade students in 205 schools representing all governorates and educational authorities in Jordan. Results indicate that the national average on the test is 30 which reflects a very low performance. Table (24) presents the average percent correct score on the test and its subscales.

**Table 24**

**Average Percent Correct Score, standard Error, Minimum and Maximum Score on Each Subscale of the Math Test**

Subscale	Average Percent Correct	Standard Error	Lower 10%		Upper 10%	
			Min	Max	Min	Max
<b>Total Score</b>	30.01	.37	0	8.00	56.00	90.00
<b>Content</b>						
Numbers	33.11	.51	0	0	75.00	100
Operations	24.40	.42	0	0	53.85	92.3
Theory	29.41	.51	0	0	66.67	100
Fractions	30.1	.43	0	0	62.5	100
Decimals	40.99	.65	0	0	100	100
Geometry	31.54	.44	0	8.3	58.3	100
<b>Skills</b>						
Conceptual Understanding	31.55	.39	0	9.1	59.1	95.45
Procedural knowledge	33.55	.44	0	5.26	63.16	94.74
Problem Solving	18.8	.39	0	0	44.4	88.9

Table (24) shows that the performance was the highest on Decimals subscale (41.7%) and the lowest was on Operations subscale (24.4%). In terms of cognitive skills, results indicated that the lowest performance was on problem-solving (19% approximately). Also noticeable was the fact that the lower (100%) of students failed to answer correctly any of the questions on five out of the six content subscales.

## Performance Across Governorates

Table 25 and Figure 9 present the levels of performance on the Math Test by governorate, while, Table 25 gives the performance on each subscale at the governorate level.

**Tables 25**

**Average Percent Correct Score, Standard Error, and 95% CI on the Math Test in Each Governorate**

Governorate	Average Percent Correct	Standard Deviation	Standard Error	95% CI
Amman	31.86	18.84	0.65	30.56-33.16
Zarqa	33.03	20.16	1.04	30.95-35.11
Balqa	31.06	14.67	1.51	28.04-34.08
Irbid	28.76	18.31	0.78	27.20-30.32
Mafaq	29.42	17.42	1.22	26.98-31.86
Karak	18.90	14.04	1.23	16.44-21.36
Tafileh	25.78	12.62	1.86	22.06-29.50
Maan	28.36	16.43	1.17	26.02-30.70

**Table 26**

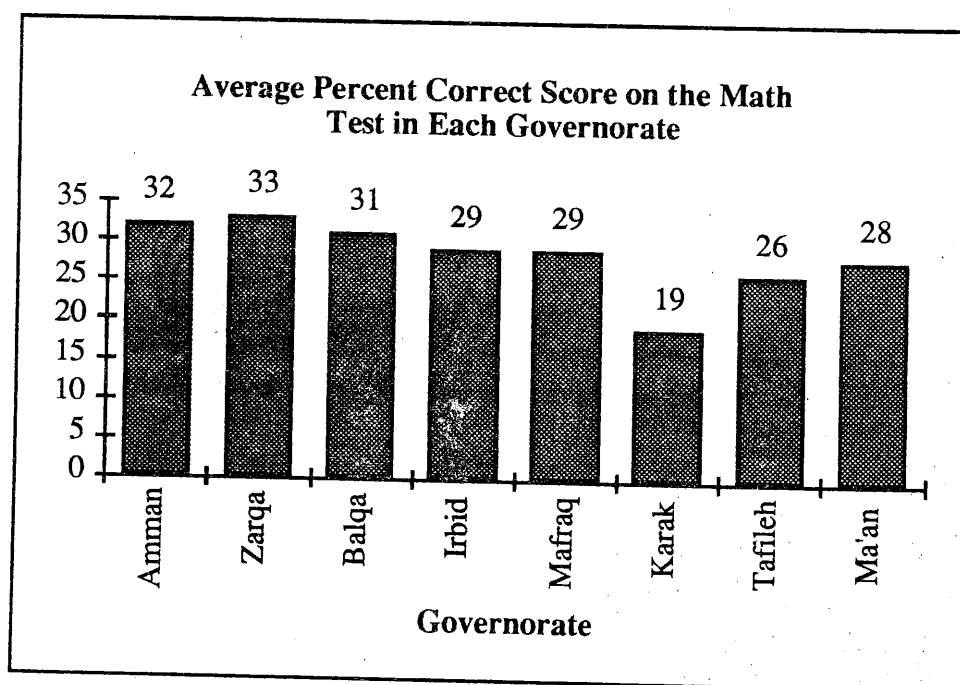
**Average Percent Correct Score and Standard Error on Each Math Test Subscale in Each Governorate**

	Amman	Zarqa	Balqa	Irbid	Mafraq	Karak	Tafileh	Ma'an
<b>Numbers</b>								
APC	34.46	39.35	29.52	32.59	31.03	22.52	19.84	30.87
Standard Deviation	26.12	27.06	22.81	24.36	22.30	20.36	17.59	20.07
Standard Error	.90	1.39	2.35	1.04	1.56	1.78	2.59	1.43
<b>Operations</b>								
APC	26.48	28.96	24.47	23.12	20.80	13.68	18.06	22.61
Standard Deviation	20.78	23.59	18.22	20.54	19.27	14.82	12.69	19.44
Standard Error	.72	1.21	1.88	.88	1.35	1.29	1.87	1.39
<b>Theory</b>								
APC	30.26	31.22	33.87	28.75	32.92	16.54	24.64	28.15
Standard Deviation	26.24	25.8	23.75	24.82	24.73	17.84	20.10	24.05
Standard Error	.91	1.33	2.45	1.06	1.74	1.65	2.96	1.72
<b>Fractions</b>								
APC	32.58	31.93	29.26	28.51	28.76	18.42	32.07	29.21
Standard Deviation	21.72	22.84	17.54	20.77	21.05	16.79	19.48	21.25
Standard Error	.75	1.17	1.81	.89	1.48	1.47	2.87	1.52
<b>Decimals</b>								
APC	44.76	44.59	41.49	36.64	44.66	25.70	29.71	38.95
Standard Deviation	32.91	33.63	27.95	29.93	29.81	27.6	25.56	31.26
Standard Error	1.14	1.73	2.88	1.28	2.09	2.41	3.77	2.33
<b>Geometry</b>								
APC	33.03	31.97	36.44	30.5	32.55	21.95	33.51	29.8
Standard Deviation	21.72	21.13	18.97	22.05	23.32	17.67	18.6	19.8
Standard Error	.75	1.09	1.96	.94	1.64	1.54	2.75	1.41

APC: Average Percent Correct.

Examination of Table 25 and Figure 9 reveals that the highest performance on the test was shown by the students of Zarqa governorate and the lowest by the students of Karak governorate. Taken as a whole, students in Amman, Zarqa, Balqa, Irbid, Mafraq, Tafleleh and Ma'an Governorates scored significantly higher than did students in Karak.

**Figure 9**

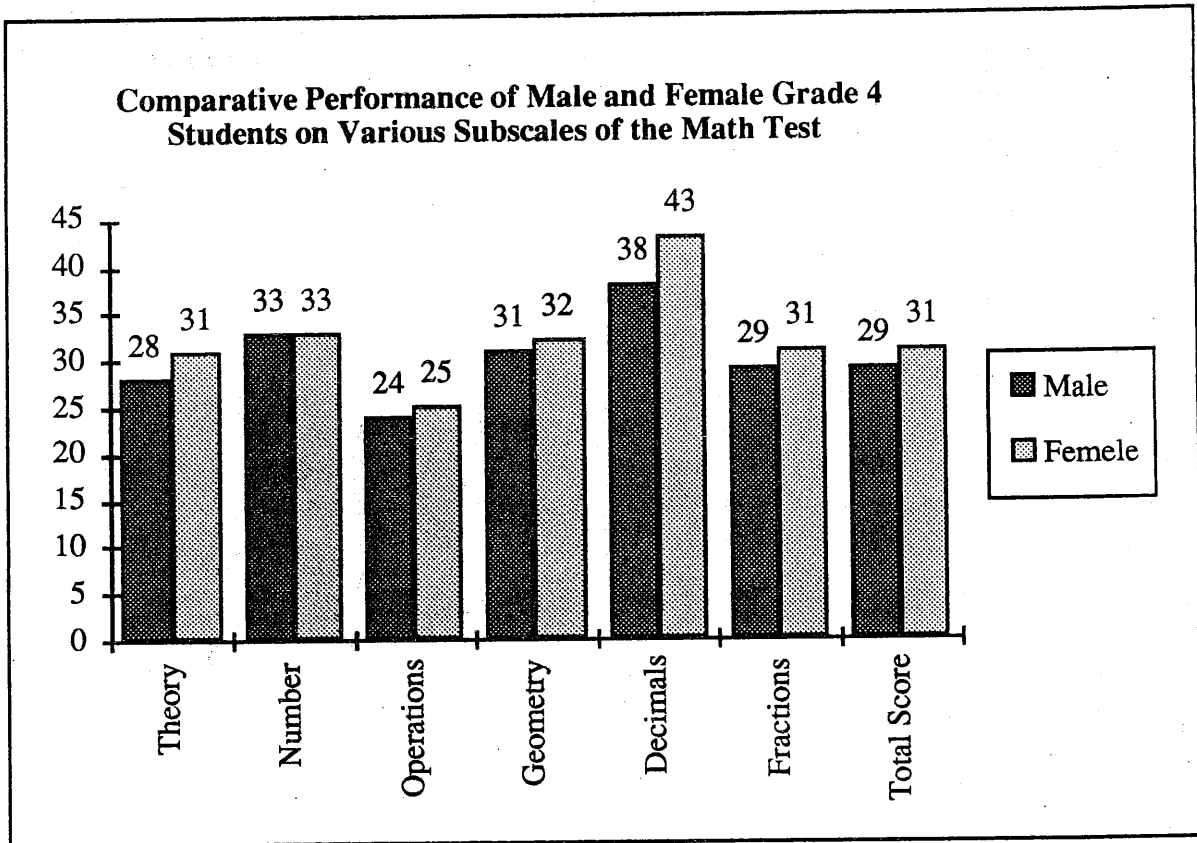


Inspection of Table 26 leads us to conclude that in all Governorates the highest performance was on the Decimals subscale, and the lowest was on Operations subscale.

### **Gender Differences on the Math Test**

Results presented in Table 27 and displayed by figure 10 indicate that differences between the performance of male and female students were not statistically significant except on the Fractions' and the Decimals subscales whereon female students, contrary to popular belief, outperformed the male students. For both groups the highest performance on content subscales was on Decimals subscale and the lowest on Operations subscale. In terms of skills, students of both sexes performed higher on procedural knowledge items than on conceptual understanding and problem solving items.

**Figure 10**



**Table 27**

**Average Percent Correct Score, Standard Error, and 95% CI of the  
4th Grade Male and Female Students on Each Math Test Subscale.**

	Average Percent Correct	Standard Deviation	Standard Error	95%CI
<b>Total score</b>				
Males	29.21	18.66	.56	28.09-30.33
Females	30.71	18.18	.50	29.71-31.71
<b>Numbers</b>				
Males	33.25	25.31	.75	31.75-34.75
Females	32.99	24.56	.68	31.63-34.35
<b>Operations</b>				
Males	23.82	20.96	.62	22.58-25.06
Females	24.90	20.62	.57	23.76-26.04
<b>Theory</b>				
Males	27.96	24.47	.73	26.50-29.42
Females	30.68	25.73	.71	29.26-32.10
<b>Fractions</b>				
Males	28.53	21.54	.64	27.25-29.81
Females	31.41	21.19	.59	30.23-32.59
<b>Decimals</b>				
Males	38.35	32.03	.95	36.45-40.35
Females	43.40	31.44	.87	41.56-45.04
<b>Geometry</b>				
Males	31.15	21.96	.65	29.85-32.54
Females	31.88	21.06	.58	30.72-33.04
<b>Conceptual understanding</b>				
Males	31.26	18.86	.56	30.14-32.38
Females	31.80	19.47	.54	30.72-32.88
<b>Procedural knowledge</b>				
Males	32.39	22.45	.67	31.05-33.73
Females	34.55	20.79	.58	33.39-35.71
<b>Problem solving</b>				
Males	17.50	18.67	.56	16.38-18.62
Females	19.94	19.18	.53	18.88-21.00

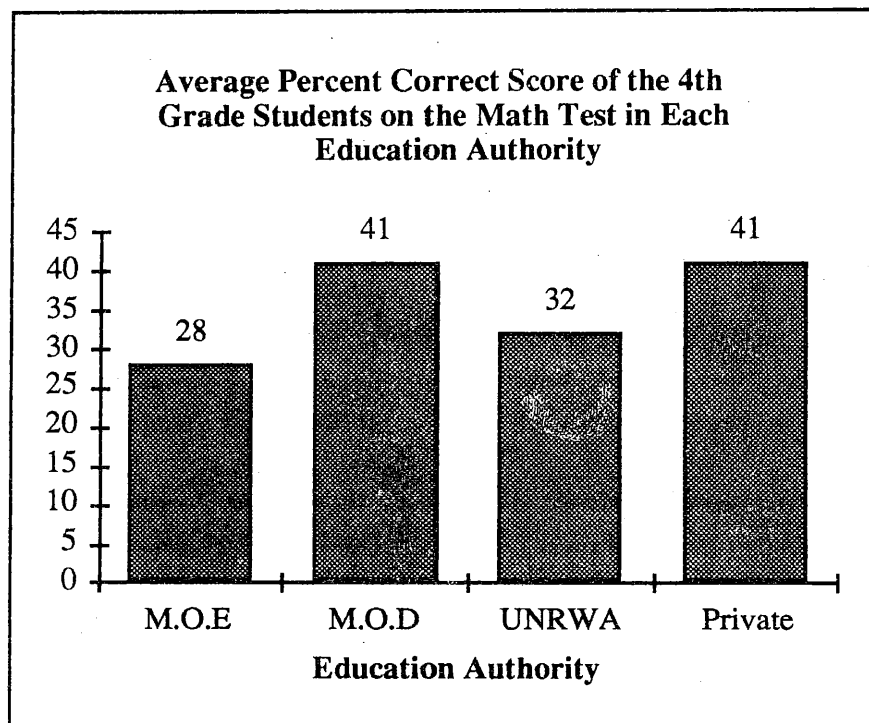
### Performance on the Math Test in Each Education Authority

Results presented in Table 28 and graphically depicted in Figure 11 show that students in both Private and Ministry of Defense schools scored significantly higher than those in the Ministry of Education and UNRWA students on the Math test. Moreover, UNRWA schools also performed better on the test than did the Ministry of Education students.

On the Math Test subscales (both content and skills) both Private school and Ministry of Defense students scored higher than Ministry of Education and UNRWA school students on all subscales except on three (Decimals, Theory and Procedural Knowledge) subscales.

In all the education authorities, performance (among content subscales) was the highest on the Decimals subscale and the lowest on Operations subscale. Among skills subscales, student performance was the highest on Procedural Knowledge items and the lowest on Problem-solving items.

**Figure 11**



**Table 28**  
**Avarage Percent Correct Score, Standard Error, and 95% CI of the Grade 4 Students on Each Math Test Subscale**

Authority	M.O.E	M.O.D	UNRWA	Private
<b>Total Score</b>				
Average % Correct	27.76	41.36	31.52	41.10
Standard Deviation	17.50	18.04	18.97	19.02
Standard Error	.40	2.00	1.30	1.16
95%CI	26.96-28.56	37.36-45.36	28.92-33.52	38.78-43.42
<b>Numbers</b>				
Average % Correct	30.36	39.35	38.50	46.18
Standard Deviation	23.82	19.58	26.10	27.57
Standard Error	.55	2.18	1.79	1.68
95%CI	29.26-31.46	34.99-43.71	34.92-42.08	42.82-49.54
<b>Operations</b>				
Average % Correct	22.08	37.99	25.44	35.65
Standard Deviation	19.77	22.24	20.20	22.38
Standard Error	.46	2.46	1.39	1.37
95%CI	21.16-23.00	33.07-42.91	22.66-28.22	32.91-38.93
<b>Theory</b>				
Average % Correct	27.63	38.48	31.05	37.81
Standard Deviation	24.34	23.96	26.49	27.87
Standard Error	.56	2.66	1.82	1.70
95%CI	26.51-28.75	33.16-43.80	27.41-43.69	34.41--41.21
<b>Fractions</b>				
Average % Correct	27.60	42.44	33.55	40.81
Standard Deviation	20.49	24.18	22.47	20.92
Standard Error	.47	2.69	1.54	1.28
95%CI	26.66-28.54	37.06-47.82	30.47-36.63	38.25-43.37
<b>Decimals</b>				
Average % Correct	38.67	47.74	41.82	54.48
Standard Deviation	31.30	37.62	29.94	31.35
Standard Error	.72	4.18	2.06	1.91
95CI	37.23-40.11	39.38-56.1	37.7-45.94	50.66-58.30
<b>Geometry</b>				
Average % Correct	29.62	45.47	29.76	42.13
Standard Deviation	20.78	21.69	23.03	20.40
Standard Error	.48	2.41	1.58	1.25
95%CI	28.66-30.58	40.65-50.29	26.6-32.92	39.63-44.63
<b>Conceptual understanding</b>				
Average % Correct	29.49	45.40	31.99	41.33
Standard Deviation	18.39	16.80	20.16	19.86
Standard Error	.43	1.87	1.38	1.21
95%CI	28.63-30.35	41.66-49.14	29.23-34.75	38.91-43.75
<b>Procedural knowledge</b>				
Average % Correct	30.77	41.98	37.24	47.39
Standard Deviation	20.51	20.84	22.50	22.18
Standard Error	.47	2.32	1.55	1.36
95%CI	29.83-31.71	37.34-46.62	34.14-40.34	44.67-50.11
<b>Problem solving</b>				
Average % Correct	17.15	30.18	18.29	27.28
Standard Deviation	17.72	25.53	19.33	21.42
Standard Error	.41	2.84	1.33	1.31
95%CI	16.33-17.97	24.50-35.86	15.63-20.95	24.66-29.90

## **Math Test Performance in Urban and Rural Schools**

Table 29 presents levels of performance on the Math Test and its subscales by school location.

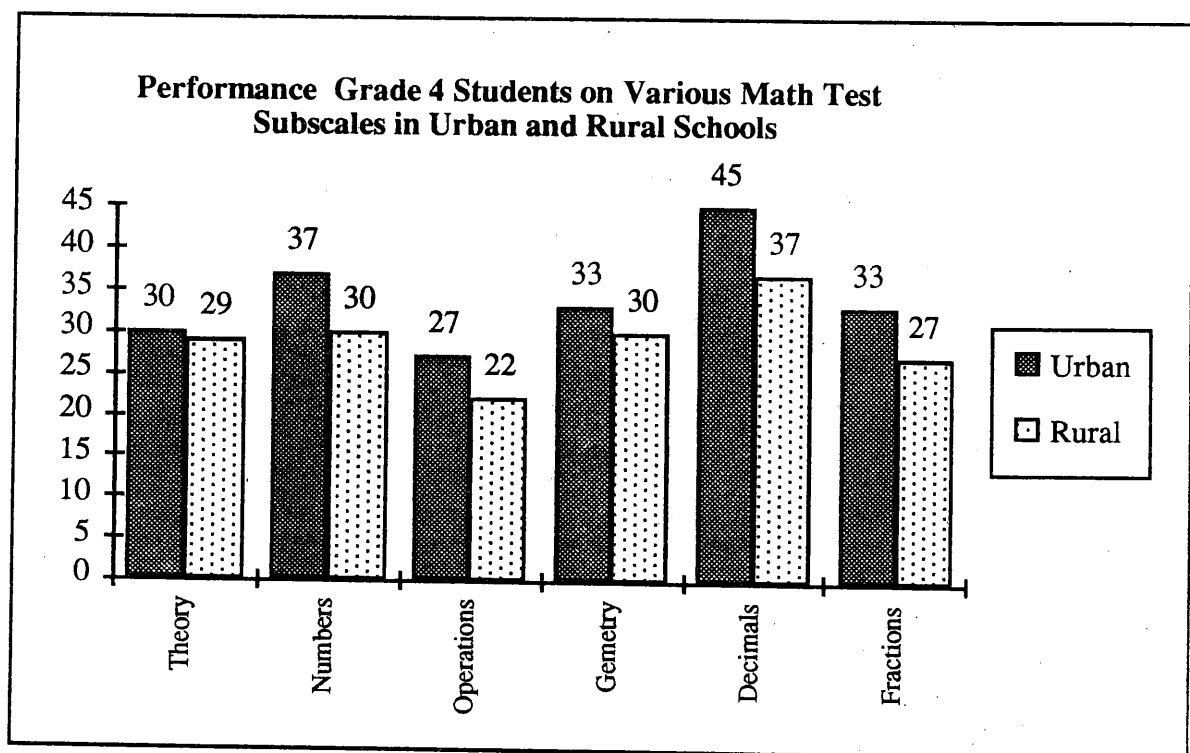
**Table 29**

### **Average Percent Correct Score on Each Math Test Subscale in Urban and Rural Schools**

	Average Percent Correct	Standard Deviation	Standard Error	95%CI
<b>Total score</b>				
Urban	32.52	18.89	.57	31.38-33.66
Rural	27.90	17.74	.49	26.92-28.88
<b>Numbers</b>				
Urban	36.85	26.29	.79	35.27-38.43
Rural	29.97	23.22	.64	28.69-31.25
<b>Operations</b>				
Urban	26.72	21.61	.65	25.42-28.02
Rural	22.45	19.85	.55	21.35-23.55
<b>Theory</b>				
Urban	30.41	25.20	.76	28.89-31.93
Rural	28.58	25.14	.69	27.20-29.96
<b>Fractions</b>				
Urban	33.23	21.49	.65	31.93-34.53
Rural	27.41	20.96	.58	26.25-28.57
<b>Decimals</b>				
Urban	45.32	32.64	.98	43.36-47.28
Rural	37.35	30.63	.84	35.67-39.03
<b>Geometry</b>				
Urban	33.30	21.58	.65	32.00-34.60
Rural	30.06	21.30	.59	28.88-31.24
<b>Conceptual Understanding</b>				
Urban	33.62	19.64	.59	32.44-34.80
Rural	29.80	18.63	.51	28.78-30.82
<b>Procedural Knowledge</b>				
Urban	36.78	22.19	.67	35.44-38.12
Rural	30.87	20.72	.57	29.68-32.01
<b>Problem Solving</b>				
Urban	20.85	19.87	.60	19.65-22.05
Rural	17.08	18.02	.50	16.08-18.08

It is clear from Table 29 and Figure 12 that urban students outperformed the rural students on the total score as well as on all the subscales except for the Theory subscale where the difference between the performance of the two groups was not statistically significant. For both locations, the highest performance (among content subscales) was on Decimals and the lowest on Operations. Among skills, the highest performance for students of both urban and rural locations was on Procedural Knowledge items and the lowest on Problem-solving items.

**Figure 12**



## Life Skills

### Performance on the National Level

The life skills test was administered to 4776 fourth grade students in 205 schools representing all governorates and educational authorities in Jordan.

Results indicated that the national average percent correct on the test was 61.2. Table (30) presents the frequencies of the average percent correct categorized in deciles.

**Table 30**

**Interval Frequency Distribution of the Percent  
Correct Score on the Life Skills Test**

Interval	Freq%	Cum%
Less than 10%	0.30	30
10 - < 20 %	0.60	0.9
20 - < 30 %	1.80	2.7
30 - < 40 %	4.80	7.5
40 - < 50 %	13.00	20.5
50 - < 60 %	24.40	44.9
60 - < 70 %	23.60	68.5
70 - < 80 %	24.00	92.5
80 - < 90 %	7.00	99.5
90 - 100 %	0.30	99.8
	<b>100</b>	

It can be seen from Table (30) that 79.5% of the fourth grade students who took the test had an average percent correct greater than 50. However, only 7.5% of those who took the test, scored higher than 80% (the acceptable level of performance).

Table 31 presents averages of percent correct by content and levels of skills measured by the test.

**Table 31**  
**Average Percent Correct Score on Each Subscale**  
**of the Life Skills Test**

Subscale	Average Percent Correct	Standard Error	Lower 10 %		Upper 10 %	
			Min	Max	Min	Max
<b>Total Score</b>	<b>61.2</b>	<b>0.21</b>	<b>0</b>	<b>42</b>	<b>79</b>	<b>96</b>
<b>Content</b>						
Health & Nutrition	68.4	0.24	0	47	89	100
Environment	51.8	0.32	0	25	75	100
Daily Behavior	48.7	0.25	0	27	73	100
Civic Education	69.3	0.28	0	44	89	100
<b>Skill</b>						
Knowledge	64.9	0.22	0	44	81	100
Application	55.9	0.24	0	35	75	100

It can be seen from Table (31) that the performance of fourth grade students was the best on the Civic Education subscale followed by Health & Nutrition, then Environment, and finally on the Daily Behavior Subscale. In terms of skills, performance was better on the Knowledge subscale than on the Application subscale with a difference of 9%.

### **Performance by Governorate**

Table 32 shows the average percent correct score in each Governorate on the test as a whole, while Figure 13 gives its graphic presentation.

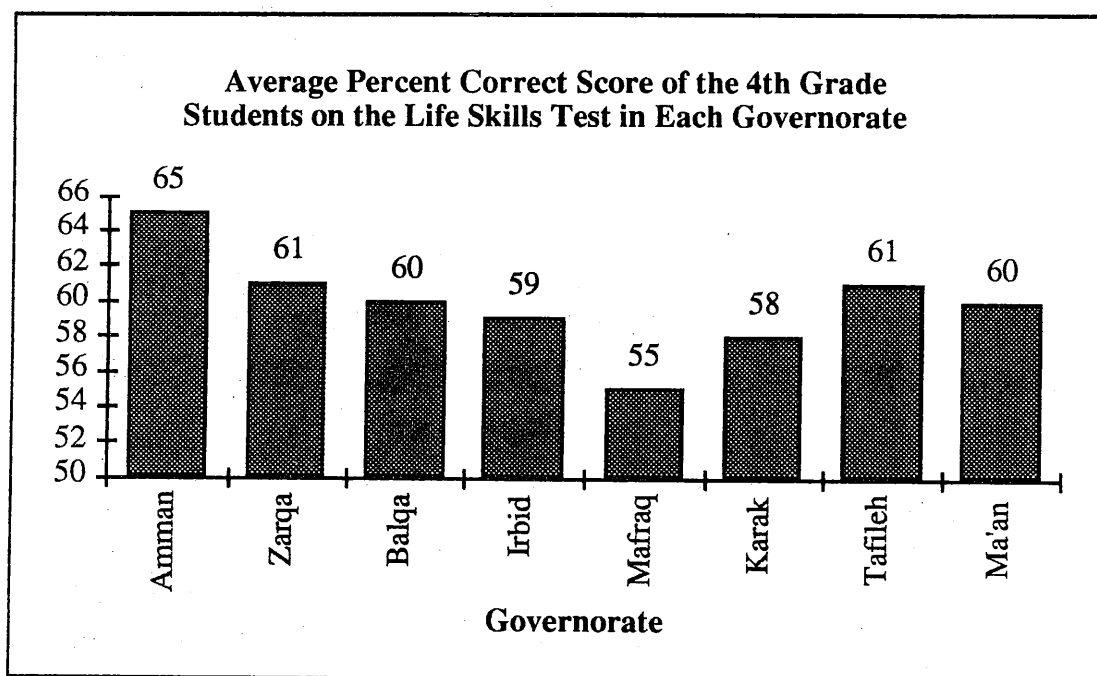
The performance levels of students in each Governorate on each subscale of the Life Skills Test are given in Table

**Table 32**

**Average Percent Correct Score and Standard Error on Each Subscale at the Governorate Level**

Governorate	Average Percentage Correct	Standard Deviation	Standard Error	95% Confidence Interval
<b>Amman</b>	64.8	13.7	.34	64.12 - 65.48
Zarqa	60.7	13.9	.51	59.68 - 61.72
Balqa	60.1	14.4	1.05	58.00 - 62.2
Irbid	59.1	15.4	.46	58.2 - 60.00
Ma'fraj	55.1	13.4	.71	53.7 - 56.5
Karak	57.7	13.1	.79	56.1 - 59.3
Tafleh	60.9	11.2	1.10	58.7 - 63.1
Ma'an	60.4	12.2	.62	59.2 - 61.6

**Figure 13**



**Table 33**

**Performance on the Life Skills Subscales in Each Governorate**

Governorate	Amman	Zarqa	Balqa	Irbid	Ma'raq	Karak	Tafileh	Ma'an
<b>Health &amp; Nutrition</b>								
Average % Correct	72.5	68.3	67.5	65.9	62.4	64.9	64.9	66.8
Standard Devition	16	16.5	16.1	17.8	16.3	15.4	12.2	14.9
Standard Error	0.39	0.61	1.2	0.54	0.86	0.93	1.4	0.76
<b>Enviroment</b>								
Average % Correct	55.3	51.7	52.6	50.3	42.8	46.3	55.4	52.1
Standard Devition	21	21.3	22.4	22.6	21.2	21.2	20.1	20.6
Standard Error	0.52	0.78	1.6	0.68	1.1	1.2	2.1	1.1
<b>Daily life</b>								
Average % Correct	51.8	45.8	47.1	47.2	44.6	48.9	53.9	48.7
Standard Devition	17.4	16.7	16.9	17.7	15.9	15.3	17.2	15.6
Standard Error	0.42	0.62	1.2	0.53	0.84	0.92	1.7	0.8
<b>Civic Education</b>								
Average % Correct	72.9	70.5	66.9	66.8	63.5	63.5	66.1	68.8
Standard Devition	18.2	18.1	20.3	19.9	19.4	18.7	14.7	18.1
Standard Error	0.45	0.67	1.4	0.6	1	1.1	1.5	0.92

Examination of Table 33 reveals that the students in Amman Governorate outperformed the students in other Governorates ( $\alpha = .05$ ). Moreover, performance of students in the Governorates of Zarqa, Balqa, Irbid, Tafileh, and Ma'an was statistically higher than performance of Ma'raq students.

when we take content subscales of the test into consideration (Table 33), we notice that on Health & Nutrition and Civic Education subscales the highest performance was of Amman students while the lowest performance was of Ma'raq students. On the Environment and Daily Life subscales, the highest performance was for Tafileh students and the lowest was of Ma'raq students.

**Gender Differences on Life Skills Test**

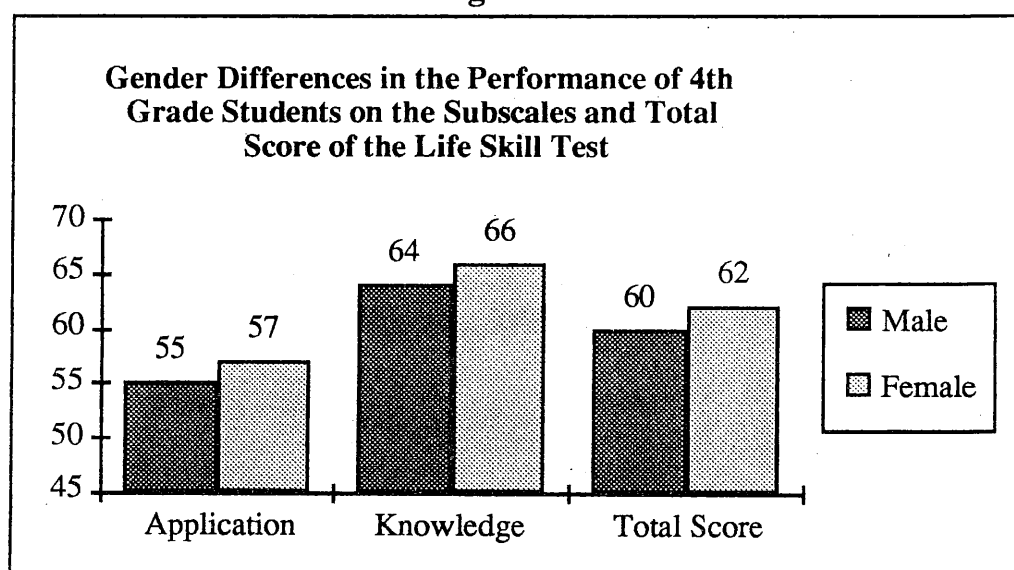
Results (Table 34) displayed by Figure 14 indicated that performance of female students was significantly higher on the total score and on the Environment, Civic Education, and skill level subscales than that of the male students. The pattern of performance was the same for the male and female students. That is, for both girls and boys, the highest performance was on Civic Education subscale, followed by Health & Nutrition, then Environment, and finally on the Daily life subscale. Performance of both sexes was also higher on Knowledge items than on Application items.

Table 34

**Average Percent Correct Score, Standard Error, and 95% CI  
of the Male and Female Students on Each Subscale and  
Total Score of the Life Skills Test.**

	Average Percent Correct	Standard Deviation	Standard Error	95% Confidence Interval
<b>Total Score</b>				
Males	60.1	14.7	0.31	59.5 - 60.7
Females	62.1	13.8	0.27	61.6 - 62.6
<b>Health &amp; Nutrition</b>				
Male	67.5	16.8	0.36	66.8 - 68.2
Female	69.2	16.6	0.33	68.5 - 69.9
<b>Environment</b>				
Male	50.1	22.1	0.47	49.2 - 51
Female	53.2	21.5	0.43	52.3 - 54.1
<b>Daily life</b>				
Males	48.2	17.9	0.38	47.4 - 49.00
Females	49.2	16.4	0.33	48.5 - 49.9
<b>Civic Education</b>				
Males	67.9	19.6	0.42	67.1 - 68.7
Females	70.4	18.3	0.36	69.7 - 71.1
<b>Knowledge</b>				
Males	64.1	15.6	0.33	63.4 - 64.8
Females	65.7	14.4	0.28	65.1 - 66.3
<b>Application</b>				
Males	54.7	16.6	0.35	59.5 - 60.70
Females	56.9	16.1	0.32	61.6 - 62.60

Figure 14

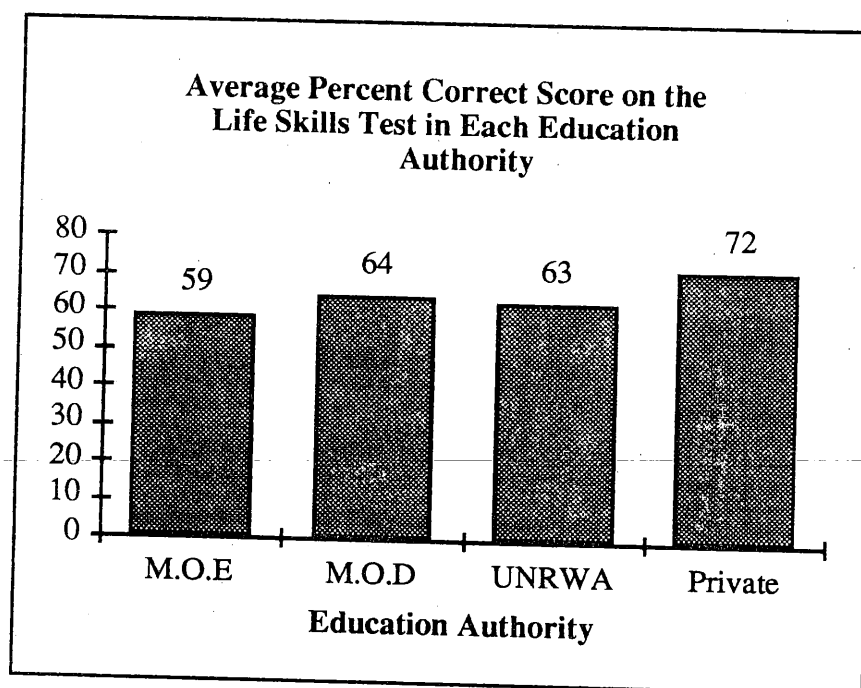


### Performance on the Life Skills Test in Each Education Authority

Results given in Table 35 and illustrated by Figure 15 indicate that Private school students outperformed the students in all other educational authorities on the Test as a whole. The order is as follows: Private schools, Ministry of Defense schools, UNRWA schools, and Ministry of Education school. The 95% Confidence Interval (CI) for the total score tells as that the average percent correct score of the private school students is statistically significantly higher than that of the students in other education authorities. On all the subscales also, the order of performance was the same as for the total score in a sense that Private schools students outperformed the students in all other education authorities, on the one hand, and the performance of Ministry of Defense and UNRWA school students was higher than that of the students of the Ministry of Education schools.

Finally, in all education authorities, students' performance was higher on Civic Education and Health and Nutrition subscales than their performance on the Environment and Daily Life subscales.

**Figure 15**



**Table 35**

**Average Percent Correct Score, Standard Error, and 95% CI on the Subscales of the Life Skills Test in Each Education Authority**

Authority	M.O.E	M.O.D	UNRWA	Private
<b>Total Score</b>				
Average % Correct	59.1	63.9	63.4	71.7
Standard Deviation	14.1	10.6	14.3	10.2
Standard Error	0.23	0.82	0.71	0.44
95% CI	58.6 - 59.7	62.3 - 65.5	62 - 64.8	70.8 - 72.6
<b>Health &amp; Nutrition</b>				
Average % Correct	66.4	70.1	70.3	79.5
Standard Deviation	16.7	13.5	17.1	12.1
Standard Error	0.3	1.1	0.86	0.52
95% CI	49 - 50.4	67.9 - 72.3	68.6 - 72.00	78.5 - 80.5
<b>Environment</b>				
Average % Correct	49.7	55.7	53.1	64.1
Standard Deviation	21.9	19.1	20.4	17.9
Standard Error	0.36	1.4	1.1	0.78
95% CI	49.00 - 50.4	52.9 - 58.5	55.3 - 50.9	62.5 - 65.7
<b>Daily Life</b>				
Average % Correct	47.1	50.6	50.9	57.4
Standard Deviation	17.1	14.4	17.2	15.7
Standard Error	0.28	1.1	0.87	0.68
95% CI	46.5 - 47.70	48.4 - 52.8	49.2 - 52.6	56 - 58.8
<b>Civic Education</b>				
Average % Correct	66.9	74.7	73.2	79.9
Standard Deviation	19.1	15.5	19.4	13.7
Standard Error	0.32	1.2	0.97	0.59
95% CI	66.3 - 67.5	72.3 - 77.1	71.3 - 75.1	78.7 - 81.1

**Differential Performance of Urban and Rural School Students on the Life Skills Test**

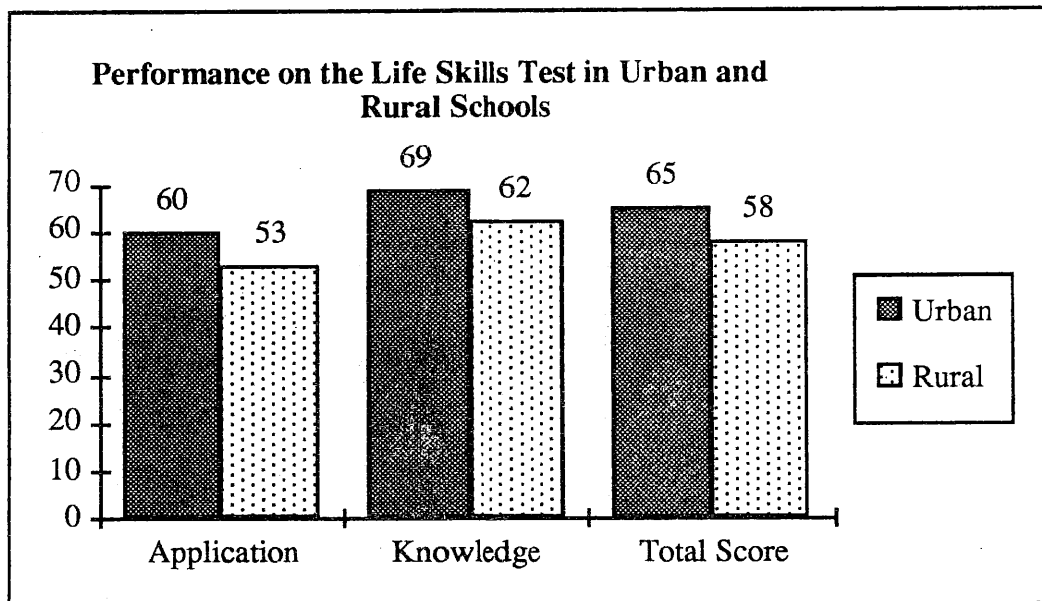
Table (36) shows the relevant statistics about the performance of 4th grade students by school location, whereas, Figure 16 presents graphic illustration of that their performance.

**Table 36**

**Average Percent Correct Score, Standard Error, and 95% CI on Each Subscale and the Total Test Score by School Location**

	Average Percent Correct	Standard Deviation	Standard Error	95% CI
<b>Total Score</b>				
Urban	64.4	13.3	.28	63.8 - 65.00
Rural	57.9	14.2	.28	57.3 - 58.5
<b>Health &amp; Nutrition</b>				
Urban	72.4	15.5	.33	71.7 - 73.1
Rural	64.9	16.9	.33	64.2 - 65.6
<b>Environment</b>				
Urban	55.9	20.9	.43	55 - 56.8
Rural	48.2	21.9	.45	47.3 - 49.1
<b>Daily Life</b>				
Urban	50.7	17.2	.37	50 - 51.4
Rural	47.1	16.9	.33	46.4 - 47.8
<b>Civic Education</b>				
Urban	73.5	17.8	.38	72.7 - 74.3
Rural	65.5	19.2	.38	64.7 - 66.3
<b>Knowledge</b>				
Urban	68.5	13.8	.29	67.9 - 69.1
Rural	61.9	15.3	.30	61.3 - 62.5
<b>Application</b>				
Urban	59.6	16.1	.34	58.9 - 60.3
Rural	52.7	16.1	.31	52.1 - 53.3

**Figure 16**



It can be seen from Table 36 and Figure 16 that urban school students outperformed rural school cohorts on the total score as well as on test subscales (content & skills) with an average difference of 8%. Moreover, the performance of both urban and rural students was higher on Civic Education and Health subscales than their performance on Environment and Daily Life subscales. Students of , both locations scored higher on Knowledge items compared to their performance on Application items.

