THE HASHEMITE KINGDOM OF JORDAN
EDUCATION REFORM FOR KNOWLEDGE ECONOMY II (ERfKE II)

Classroom Observation Baseline Study Report

Monitoring & Evaluation Partnership (MEP) Project*

March, 2012

* This report is a product of collaboration between National Center for Human Resources Development (NCHRD) and World Education, Inc. (WEI) researchers under the Monitoring & Evaluation Partnership (MEP) project. MEP is a four-year (2010-2015) USAID-funded project implemented by World Education with the aim to strengthen the technical capacity of NCHRD and to provide financial support for a series of program quality evaluations for the Government of Jordan's Education Reform for Knowledge Economy (ERfKE II) program. The Classroom Observation study is the *first* of a series of evaluation studies supported by the MEP project.

Table of Contents

Acronyms List

CO	Classroom Observation
ECD	Early Childhood Development
ECE	Early Childhood Education
EMIS	Education Management Information System
ERfKE I and II	Education for the Knowledge Economy I and II
E-TVET Council	Employment–Technical and Vocational Education and Training Council
GOJ	Government of Jordan
HKJ	Hashemite Kingdom of Jordan
ICT	Information and Communications Technology
ICDL	International Computer Driving License
JEI	Jordan Education Initiative
KE	Knowledge Economy
KG	Kindergarten
M&E	Monitoring and Evaluation
MEP	Monitoring and Evaluation Partnership
MOE	Ministry of Education
MOD	Ministry of Defense
NAfKE	National Assessment for Knowledge Economy
NCHRD	National Center for Human Resources Development
NES	National Education Strategy
PCA	Principle Component Analysis
TOR	Terms of Reference
S.D.	Standard Deviation
SPSS	Statistical Package for the Social Sciences
SR	School Rationalization
WEI	World Education, Inc.
UNRWA	United Nations Relief and Works Agency

Executive Summary

The primary objective of this classroom observation baseline study is to fully understand a current status of teaching and learning practices in classrooms of the basic and secondary schools in Jordan. There will be a follow up study in two years for assessing the change and difference in classroom learning and teaching. The baseline study is carried out under the Monitoring and Evaluation Partnership (MEP) project supported by USAID and implemented by NCHRD and World Education, Inc. For this study, we observed randomly selected 852 classroom teachers who are teaching Arabic, Math and Science subjects for grades 5, 9, and 11 in 43 directorates and under various authorities and project interventions (Ministry of Education (MOE) schools, Ministry of Defense (MoD) schools, private and UNRWA schools, as well as Discovery, Madrasti and ERSP schools[1]. This is a comprehensive baseline study.

This baseline report mainly features the analysis results of several major domains of teachers' teaching practices in classrooms such as classroom management, student-centered active learning and teaching methodologies, (also known as progressive learning and teaching methodologies), methods to motivate students, and application of student assessments by various groups such as types of project intervention in the schools (Discovery, ERSP, and MADRASTI), school authority, locality, teacher gender, teacher qualification and training, subjects, and so on. It also intends to stimulate a policy-level dialogue and debate on what factors may explain the variation in teachers' utilization of student-centered active learning and teaching methodologies and how these methodologies correlate with the student achievement (measured in NAfKE assessment results). We hope that the findings from this classroom observation study will serve as a solid baseline and the analysis results may assist the MOE to further monitor and evaluate changes in teaching and learning practices and to reorient interventions and inform teacher training programs in the future in the future to improve teaching and learning practices within MOE schools, ultimately improving student learning achievement outcomes.

A random selection of a large sample size, use of scientific methods, development of a reliable classroom observation instrument, training of 65 experienced MOE observers, empirical data analysis strategies, and collaborative team work under the MEP project has made this classroom observation baseline study unique and credible and useful.

Specifically, this baseline study has provided useful information in the following areas:

1) The current profile of teaching practices in MOE schools in Jordan;

2) The extent to which teachers performed well or poorly on the use of student-centered active learning and teaching and other teaching methodologies in the classroom;

3) The large variations in the key measures of teaching and learning behavior in class between and within directorates, and other comparison groups such as schools under various authorities and project interventions;

[1] ERSP schools are intervention schools under the USAID funded project, Education Reform Support Project in Jordan

4) Which factors (limiting to teacher characteristics) significantly explain the variation in the measure of student-centered active learning and teaching.

In addition, we integrated the classroom observation data with NAfKE student assessment data to answer the question if the use of student-centered active learning and teaching methodologies correlates with student achievement.[2]

Major Findings:

1) Teachers in Jordan are generally good "classroom managers" in time management, checking student attendance, presenting lesson objectives and plan and organizing orderly class activities, etc. These however are considered as very basic requirement for being a classroom teacher regardless which pedagogical approach or methodology teacher would apply. Female teachers are significantly better than male teachers in this examined domain.

2) Teachers in Jordan are not yet all "SCALT" teachers using student-centered active learning and teaching methodology. They have much room for improvement. Only 16.2% of teachers are ranked as excellent teachers in applying student-centered active learning and teaching methodology (SCALT). Almost 20% are considered very poor performance in applying the methodology.

3) Female teachers are significantly better than male teachers in applying the student-centered active learning and teaching methodologies in classroom even after considering all things equal.

4) ERSP school teachers are significantly better in applying the student-centered active learning and teaching methodologies in classroom than all other school teachers (private schools, MoE schools (non-project), MoD schools, UNRWA schools, Discovery Schools, Madrasti schools).

5) There is a significant variation in applying the student-centered active learning and teaching methodologies between directorates and within directorates. In other words, good experience in applying SCALT methodologies exists in all places locally and fostering learning and exchange of experience between and within directorates could be really beneficial for educators in Jordan.

6) A large proportion (90%) of the variance in the student-centered active learning and teaching has not been explained. There is much room for identifying "policy manipulatable variables" such as specific training programs, student-teacher ratio, class size, etc. These types of relational analysis could be further explored by integrating data from multiple sources, levels and years.

7) Teachers in Jordan generally speaking are not doing much in classrooms to stimulate students' motivation to learn. Based on our analysis, only 2% teachers are considered as "excellent" teachers stimulating students' motivation to learn in classroom. Comparing averages of the composite score of the practice to stimulate student motivation, female teachers again are better than male teachers.

[2] We note that 2011 NAfKE assessment data is available under a separated study but integrated for a suggested correlation analysis to identify initial relationship between the student-centered active learning and teaching methodology (SCALT) and student learning achievement (NAfKE assessment).

8) Teachers in Jordan use limited student assessment techniques in classroom teaching. The composite indicator in this is very low. The average score is only 2.4 out of the maximum value 7 indicating that teachers rarely used student assessment methods to assess students. However, with this low score, female teachers yet again perform significantly better than male teachers in this aspect.

9) Teachers in Jordan rarely use computers or other modern technology in classroom teaching. For example, only 7% of teachers used computers in Arabic and Math classes and 4% of teachers used computers in Science classes. However, it is known if this is the fact that computers are readily available for teachers to use in classroom but they don't use them during the observed classes or computers are not yet assigned to these teachers for use in classrooms. This is a snapshot of the situation of computers in use in class teaching time.

10) Comparing the observation results with the opinions by the 65 observers, we found there is a significant discrepancies in assessing the measured domains. Mostly observers overrated teachers performance in the measured domains such as classroom management, applying methods to motivate student learning, and to assess student performance in classes. However, the observers' opinions about the student-centered active learning and teaching is mostly in line with or slightly underrated than the observed data.

11) Correlation results show that in grade 5, there is no relationship between SCALT and NAfKE in any of the 3 subjects. But in grade 11, there is a significant relationship between SCALT and NAfKE in all 3 subjects. A similar situation is observed in 9th grade, with the exception of Science. The higher the teachers' utilization of SCALT methodologies in the classroom, the higher their students' NAfKE performance in Science. In 11th grade, all correlation coefficients were positive and significant, which demonstrate that SCALT methodologies might be an important predictor of students' scores in higher grades across subjects. However, the magnitude of that relationship varied by subject.

The main results of this baseline study suggest that Jordanian teachers have much room for improvement utilizing student-centered active learning and teaching methodologies, integrating technology into classroom teaching and learning, motivating students in learning, and providing in-classroom student assessment feedback to students. This study also shows that a significant number of excellent teachers in the measured domain areas can be found between and within many directorates or urban and rural areas. Learning from the high performance teachers is critical to replicate this success through the Jordanian education system.

As we have completed the baseline study, we are fully aware of the remaining challenges and unanswered research inquiries facing the Jordan education system. For example, how to validly and reliably assess students' critical thinking and problem solving skills in Jordan? Can these 21st century skills be effectively taught by teachers in Jordan? Why do female teachers perform significantly better in almost all aspects of the outcome measures than their male counterparts even considering all other things equal? Recognizing these challenges, we are continuing our research and evaluation journey

and strengthening our analytical capacity as education analysts to support ERfKE II program and improve the quality of Jordan education system.

1. Introduction

1.1 Background and Context

In early 2003, the Government of Jordan (GOJ) launched the comprehensive Education Reform for Knowledge Economy (ERfKE) program with the aim to empower general education graduates with the knowledge, skills and competencies that enable them to effectively participate in the Knowledge Economy (KE) and thus to improve their futures. The ERfKE program is implemented in two stages. ERfKE I (2003-2009) aimed to produce high-quality graduates who could effectively compete and participate in a competitive global knowledge based economy. The ERfKE II program (2010-2015) is designed to build on the objectives of ERfKE I but with a renewed focus on the system changes necessary for a shift to education for a knowledge economy. It reflects a deliberate focus on ensuring that the system changes (curriculum and assessment reform, teacher development, policy and strategy capacity) are manifested in changes in learning outcomes in schools and classrooms. ERfKE II components are as follows:

- **Component I**: Establishment of a National School-based Development System
- **Component 2**: Policy, Planning, M&E and Organizational Change
- **Component 3**: Teaching and Learning Resources
- **Component 4**: Special Focus Program Development (Early Childhood Development, Vocational Education, and Special Education)
- **Component 5**: Quality Physical Learning Environments

Established in 1989, the National Center for Human Resources Development (NCHRD) is a semi-autonomous research body in Jordan. The primary goal of NCHRD is to create a balance between the outputs of training and education programs on the one hand and the knowledge, skills, abilities and attitudinal requirements of the labor market on the other hand. NCHRD is also in charge of executing the external monitoring and evaluation activities for the various educational interventions of the Ministry of Education (MOE). This role is considered an extension of what has been done by NCHRD during the first phase of ERFKE (2003-2009), namely the implementation of 35 evaluation studies in the areas of curriculum, student assessment, teacher training, information and communication technology and early childhood development. Under ERfKE II, NCHRD is charged with implementing 30 studies, of which five baseline studies, including this baseline evaluation of teaching and learning practices in MOE schools.

The Classroom Observation Baseline Study is aligned with Component 3 of the ERfKE II program, namely to improve the teaching and learning through various program interventions aiming at: teacher employment, teacher utilization, and professional development policies and practices. These core areas are managed under two sub-components: 1) Teacher policy, training and professional development; 2) Curricula improvements and the development of assessment and learning resources – including the use of ICT and e-learning.

1.2 Previous Relevant Studies by NCHRD

In 1995, NCHRD conducted a study aimed to identify teaching and learning practices and activities carried out by the teachers in Jordanian school classrooms, and to correlate the variation in teaching practices with teachers' educational qualifications and experience.[3] The results of this study showed that teachers often used traditional instructional methods in their classrooms, for example, explaining the lesson without clarifying the objectives of the lesson, or rarely linking the learning subjects with what was learned previously. On average, 60% of the time was spent on explaining the lesson. While teacher may have asked some questions, most of these questions were not shown to stimulate thinking, but rather requests for factual information.[4] In addition, the earlier study showed that there is no statistical difference in the educational practices of teachers by grade and topic or based on the qualifications or educational experience of the teachers.

In 2008, NCHRD conducted a formative evaluation study[5] aimed to evaluate the developed curricula training programs as well as the assessment strategies implemented through the staff of the MOE. The results indicated that teacher training programs on the developed curricula and/or assessment strategies were successful at several different levels, but also suffered from various deficiencies that limited the effectiveness of those programs.[6]Study results showed that 33% of the teachers encourage students to debate and dialogue in the classroom, 60% link learning to life, 65% take into account individual student differences, and about 53% provide reasoning questions that develop student's analytical thinking. More particularly, discussion methods were used by 70% of the teachers and working group methods were used by 53%, whereas the percentages of teachers who used other methods such as: project based learning method, analyze and criticism of learning topics, self-learning and workshops method were modest. Based on the results, the study recommended to design trainings to enhance teacher pedagogical skills to have a more explicit impact on the development of student skills in dealing with real life contexts.

In 2009 NCHRD conducted another study to analyze[7] the quality of classroom resources, particularly student and teacher textbooks, training materials[8], and electronic content, as well as to assess the use

[3] The sample included students in grades 4, 5, and 8 in 40 schools distributed across the country.

[4] The results showed that 75% of class time was on average allocated to the teacher's talking time – either in explanation of the lesson or classroom management – with only 25% of the time allocated to student interactive work.

[5] The formative evaluation was based on Kirkpatrick's model of the evaluation of training programs, which focus on trainee's reactions, level of the skills and knowledge acquired by trainees, as well as examining the overall results of the training programs.

[6] A variety of proposals were raised by the concerned parties to improve the training programs, including a review of the trainings' times and the establishment of training centers.

[7] Six types of instruments were used to gather data for this study: textbook analytical scheme, teacher's book analytical scheme, electronic content analytical scheme, classroom observation schedule, students' attitudes questionnaire about the new textbooks and students Interview focus groups (student's comprehension about classroom learning).

[8] The study evaluated the use of textbooks for grades 1, 4, 8 and 10 and for the following school subjects used during the 2006-2007 academic year: Arabic Language, Mathematics, Science and National and Social Studies.

of various new pedagogical skills and approaches in the classroom and students' comprehension and achievement.[9] The study results showed the following:

- Medium degree of linkages between teaching and assessment strategies and learning outcomes.
- The e-content did not differ significantly from the printed materials – basically a book in digital format – and did not employ the full potential of computer animation and media technologies.
- Limited methods used by teachers in presenting learning objectives and expectations to students prior to the lesson.
- Little use of ICT in the classroom.
- Most teachers used assessment for the purposes of examining the level of students' textbook based achievements, but did not take into account understanding.

In sum, the previous studies are relevant to the current baseline study in that they collected data evidence and presented a general profile of teachers classrooms practices and quality of technology resources. However, many of the conclusive findings (strong claims) were often not credibly supported by strong data evidence or rigorous analytical process. It is clear that large educational concepts (technically called "constructs") such as teaching of critical thinking, problem solving, cooperative learning, utilization of technology in classroom, or teacher application of student-centered active learning and teaching in class, should be scientifically measured and observed and analyzed. The current baseline study covers and emphasizes this deeper and more credible approach.

1.3 Study Objectives

The primary objective of this baseline classroom observation evaluation report is to present baseline indicators and critical analyses about the current teaching and learning practices in Ministry of Education (MOE) schools, private, UNRWA schools as well as Ministry of Defense (MoD) schools in Jordan. This baseline report describes the current profile of teaching practices by core subject teachers in Grades 5, 9 and 11 in Jordan, reports on comparative analysis results of student-centered active learning and teaching practices (also known as progressive teaching practices) among various groups such as types of project intervention in the MOE schools (Discovery, ERSP, and MADRASTI), school authority, locality, teacher gender, teacher qualification and training, subjects, and so on. It also intends to stimulate a policy-level dialogue and debate on what factors may explain the variation in teachers' utilization of student-centered active learning and teaching methodologies and how these methodologies correlate with the student achievement (measured in NAfKE assessment results). We hope that the findings from this classroom observation study will serve as a solid baseline benchmark for future follow-up studies under ERfKE II, as well as a tool to assist the MOE to effectively monitor

[9] Study sample included 117 schools, stratified with respect to sex, location (urban/rural) and authority (public/private/UNRWA), 276 teachers, selected from those teachers in the sample schools, including 75 Arabic Language teachers, 79 Mathematics teachers, 98 Science teachers, and 24 National and Social Studies teachers.

and evaluate changes in teaching and learning practices and to reorient interventions and inform necessary training programs to improve teaching and learning practices within MOE schools, ultimately improving student learning outcomes.

1.4 Research Questions

To meet the objectives, we developed a set of principle research questions which guided our development of the data collection instruments as well as the data analysis process. These questions were developed based on the interests of MOE policy stakeholders in conjunction with ERfKE II program stakeholders, partners and reform managers. The four principle research questions are as follows:

1) To what extent do teachers in Jordan apply student-centered active learning and teaching methodologies that include problem solving, critical thinking, cooperative learning, interactive methods, and alternative student assessment techniques in their classrooms?
2) If there is a significant variation in teachers' application of student-centered active learning and teaching methodologies, does the variation differ by gender, years of work experience, training status, locality, subjects taught, or other factors?
3) What factors may explain the variation in teachers' practical application of student-centered active learning and teaching methodologies controlling for the effects of other factors?
4) How does teachers' practical application of student-centered active learning and teaching methodologies correlate with student achievement in Jordan (based on NAfKE assessment results)?

2. Methodology

The methodology used for this baseline study is designed based on the objectives and principle research questions of the study.

2.1 Sample and Population

Given that the focus of the baseline study is to find out how key subject teachers in MOE schools in Jordan teach in the basic and secondary education system, our study population is all teachers who teach in grades 5, 9, and 11 who teach Arabic, Math and Science subjects in MOE schools. In order to reflect the reality of the teaching practices in Jordan, 300 schools of various types of schools for the school year 2009-2010 were randomly selected, with equal distribution of teachers in grades 5, 9 and 11. The sampling criteria is based on the sampling of NAFKE 2011 study which includes the following parameters: school sex (boys, girls, mixed), geographical region (north, center, south), the authority (MOE, private, UNRWA), school location (urban, rural). Moreover, the following special intervention schools were purposively oversampled: Discovery schools, MADRASATI schools, Education Reform Support Project (ERSP) schools. These would let us assess how teachers in these various intervention schools perform using new student-centered active learning and teaching methodologies. The total distribution of classrooms per subject, per grade is shown in the following table.

Table (1): Distribution of the sample per subject, per grade

Grade	Arabic	Math	Science	Total
5	98	99	99	296
9	101	100	96	298
11	99	99	60	258
Total	298	298	256	852

As shown in Table 1, the total number of the classrooms was 852, which is slightly below the originally selected 900 classrooms for observation. 48 schools were excluded from the study because they are not the “main stream” schools but “art and MIS” specialized schools which do not offer science subjects. Each of the selected 852 teachers was observed once for 45 minute period during a 2-week period of the observation data collection.

2.2 Measures

For this study, one of the keys is to develop a set of valid and reliable measures for the intended outcomes of the study, that is, teachers' utilization of student-centered active learning and teaching methodologies. Since the study is to observe how teachers teach in class, a classroom observation tool was specially developed by NCHRD in collaboration with the MEP project team based on policy interests of the key education stakeholders, ERfKE II program requirements, national curriculum standards and framework, previous and international experience and research, as well as the principle research questions. Following are the main sections of the classroom observation tool.

The observation items were based on the new curriculum standards developed under the ERfKE program which focus on critical thinking, problem solving and participatory learning methodologies. These items are composed of the following four parts:

- **Part I: General Profile Information** mainly covers teachers' characteristics and school information
- **Part II: Teaching and Learning Practices in the Classroom** is the essential core of the tool containing 57 carefully selected checklist items grouped under 5 major domains as follows:
  - ✓ Classroom management (10 items)
  - ✓ Utilization of student-centered active learning and teaching methodologies (instructional strategies) with 26 items distributed into five sub-domains: participatory learning (4 items), problem solving and inquiry (6 items), cooperative learning (7 items), active learning (4 items), and critical thinking (5 items)[10]
  - ✓ Use of instructional tools & media aids with 10 items distributed into two sub-domains: use of computer (7 items) and use of other educational aids (3 items).
  - ✓ Teaching practice to motivate students (6 items).
  - ✓ Utilization of student assessment (7 items).
- **Part III: Summative Evaluation by Observers** completed by the observers soon after the classroom observation and is composed of 16 items distributed in two sub-domains:
  - ✓ School management (8 items).
  - ✓ Student-centered active learning and teaching methodologies and student assessment strategies (8 items).
- **Part IV: General Notes about the Classroom** includes 6 items of both open- and close-ended questions.

Validity and Reliability of the Observation Instrument

Since the major part of the study is to examine teaching practices in the five major domains mentioned above, exploratory factor analysis and reliability tests were conducted to ensure the validity and reliability of the developed items. All items under each domain must be well tested to contribute to the overall measure of that specific domain. The factor analysis results show that the instrument was well developed and statistical results confirm valid measures of the major domains and sub-domains in the instrument. For example, for the factor analysis, Principle Component Analysis (PCA) was used to identify six domains, each of which has an Eigen value larger than 1. In fact, the sixth dimension (in this case representing the sixth domain) only contains three valid items, including 1) debating technique, 2) playing games for learning, and 3) asking students to prepare project

[10] These sub-domain groupings were slightly changed (less than 5%) when we used factor analysis to confirm our domain and sub-domain components.

materials. It was decided not to use the sixth sub-domain under the overall domain of student-centered active learning and teaching due to the following three factors: 1) lowest ranked Eigen value, 2) lopsided distribution for each item, and 3) relevance by conceptual framework. All of the five sub-domain composites (cumulative 61.8% variance contained) are almost perfectly aligned with the original design. Only one question item was regrouped (q29).

In addition, statistical reliability test was conducted for each domain or sub-domain to ensure that items grouped under each domain or sub-domain are reliable. Alpha level of .70 or above was chosen as a reliability criteria. In most cases, the domain measures are between .75 and .93.

2.3 Data Collection

Data was collected by 65 MOE directorate-level supervisors for Science, Math, and Arabic language subjects. In addition to their experience in the subjects and field observation, all observers participated in a day-long training on consistency and observation tool utilization before they went to the field. The following procedures were followed to collect the data using the classroom observation tool:

- NCHRD sent a formal letter to MOE asking to nominate supervisors for Science, Math and Arabic language based on the sample distribution.
- Selected supervisors attended one-day training by NCHRD on how to use the observation tool. NCHRD prepared an implementation manual that includes a rubric to ensure a high level of consistency during the observation process.
- An implementation schedule was given to each supervisor based on the sample distribution. The supervisor was asked to call the school in advance to coordinate the visits. The data collection spanned two weeks during the period of 11-22 May, 2011 with one or two observations per supervisor per day.
- After the implementation, each observer's set of instruments were reviewed to ensure accuracy of the work.
- The codebook was prepared.
- The data was entered and cleaned before starting the analysis.

2.4 Limitations

Classroom observation study usually takes more time, level of effort and logistic resources than other types of educational studies in the field of educational evaluation. For this study, a large number of observers were sent to schools all over Jordan for classroom observations. We recognize that it is a challenge to ensure absolute consistency in observing classes and using the observation tool effectively among all the observers even though everyone participated in a day-long training. Further, given the time and resources, we admit that we were not able to observe teachers for multiple times, nor to cover other subjects in this baseline study. Had we conducted more observations for each

teacher and in more subjects, we would have had an even greater confidence in the validity and reliability as well as representativeness of data reflecting the true reality.

3. Findings

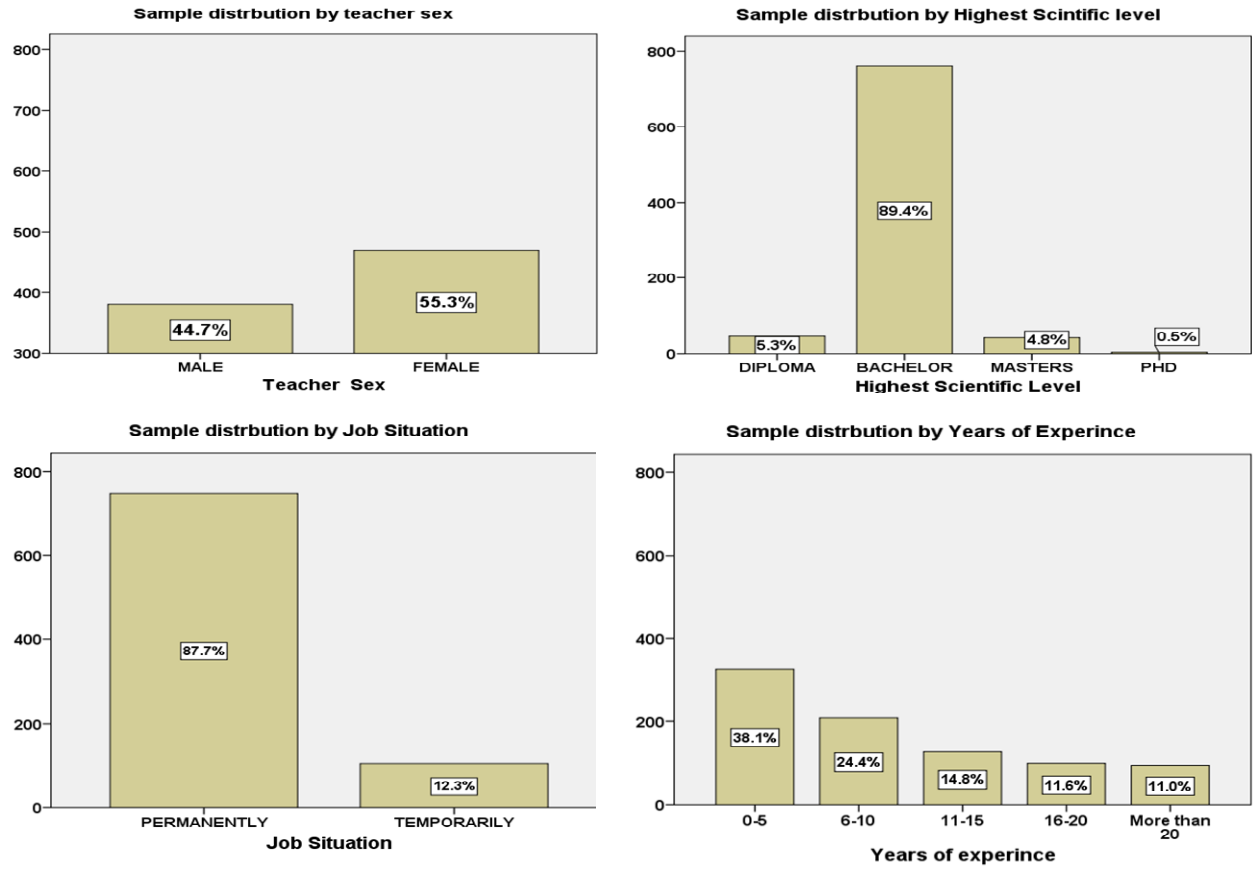
3.1 General Profile of Sampled Teachers in MOE Schools

A large number of randomly selected teachers in MOE schools were observed in their classes. These teachers are currently teaching Arabic, Math and Science subjects in Grades 5, 9 and 11, and they are located in 43 directorates. The sample intentionally covered various types of MOE schools (Discovery, Madrasti, ERSP), and other MoE schools, MoD, private and UNRWA schools. Among 852 selected teachers, 44.7% are male teachers and 55.3% are female teachers. A great majority of the teachers, 89.4% have bachelor decrees in various subjects, 4.8% have master's degrees, and 0.5% have PhD degrees. 5.3% have only reached a diploma level (below a bachelor's degree).

In terms of teachers job status, 87.7% hold permanent job status and 12.3% temporary job (who has at most a one year contract) status. Most teachers (62.5%) in the sample have less than 10 years of work experience indicating their young status, but also potential for long tenure. Only 11% of sampled teachers have worked for more than 20 years. 72.5% of the selected teachers have gone through professional job development training. 58.7% of the teachers received International Computer Driving

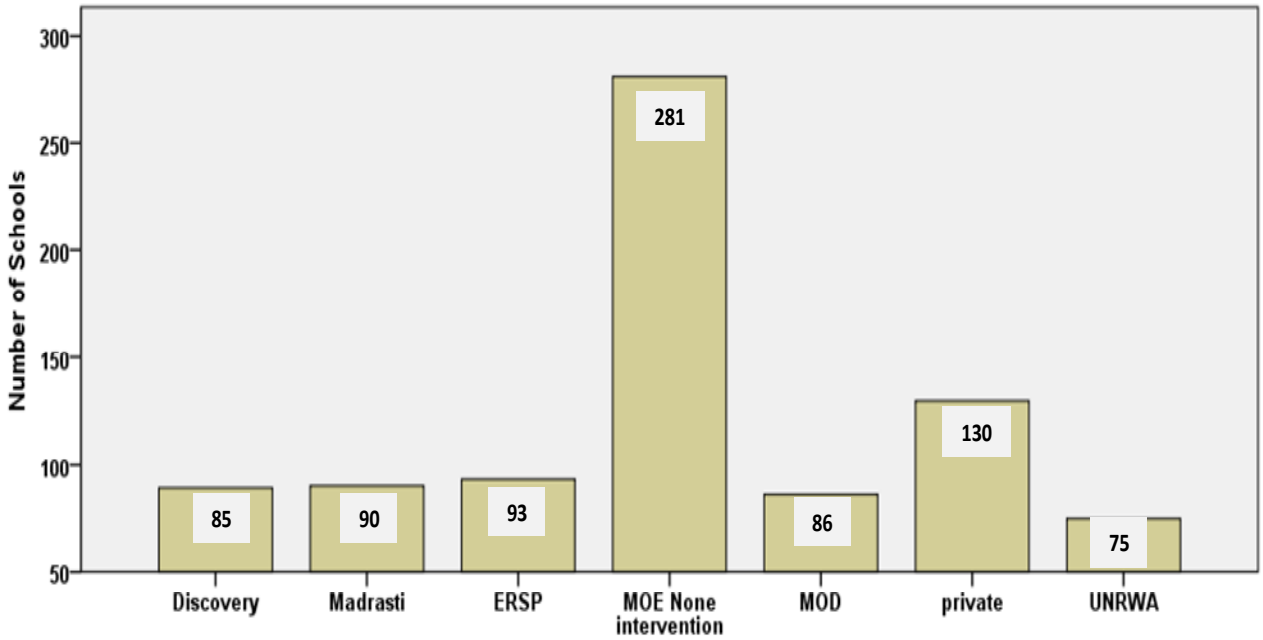
License (ICDL) training and 14.9% received INTEL training. Following Chart (01) below illustrates this.

Chart (01): General Profile of Sampled Teachers in Baseline Classroom Study



In addition, there are quite number of schools in the sample which participated in various project interventions or belong to various authorities. See Chart (02) below. The study intends to compare classroom observation results by these various authorities and types.

Chart (02): Number of Sampled Schools by Authority and Project Intervention Type



This sample profile fits well within the national profile of teachers teaching in Grades 5, 9, and 11 in Arabic, Math and Science subjects. When the profile of MoE schools only was checked, there is no significant difference from this overall profile.

3.2 Overall Classroom Observation Results

In the classroom observation instrument, a total number of 59 items were developed for observers to identify if certain specific teacher practices were observed. These practices are grouped under five major domains such as classroom management, student-centered active learning and teaching (with 5 sub-domains), utilization of IT and media aids, methods to motivate students and methods to assess student performance. In order to see how teachers in Jordan practice teaching in classrooms, the percentages of all observation check list (items) were computed to check the presence or non-presence of the specified teacher practices. These item percentages are useful to inform how many teachers (in percentages) have used or not specific teaching and learning methodologies, certain management techniques, student assessment techniques, and methods to motivate students. For example, on observation item 3, (48.1%) of teachers tracked student attendance and (51.9%) did not; on observation item 21,(32.8%) of teachers divided class into groups and (67.2%) did not. On item 38; only (7.8%) of observed teachers used projection screen to teach the observed classes and (92.2%) did not use it. To see the whole list of the percentages of all observation items, see **Appendix 1: Teaching and Learning Practices in Classroom.** However, this report goes beyond the simple examination of percentages of the observed practices. It analyzes teacher performance under pre-defined domain

areas so as to see how well teachers in Jordan perform in the baseline year. This also will set up a baseline year analytics for future data comparison and trend analysis. Next, the report describes how teachers performed in each measured domain area in the baseline year.

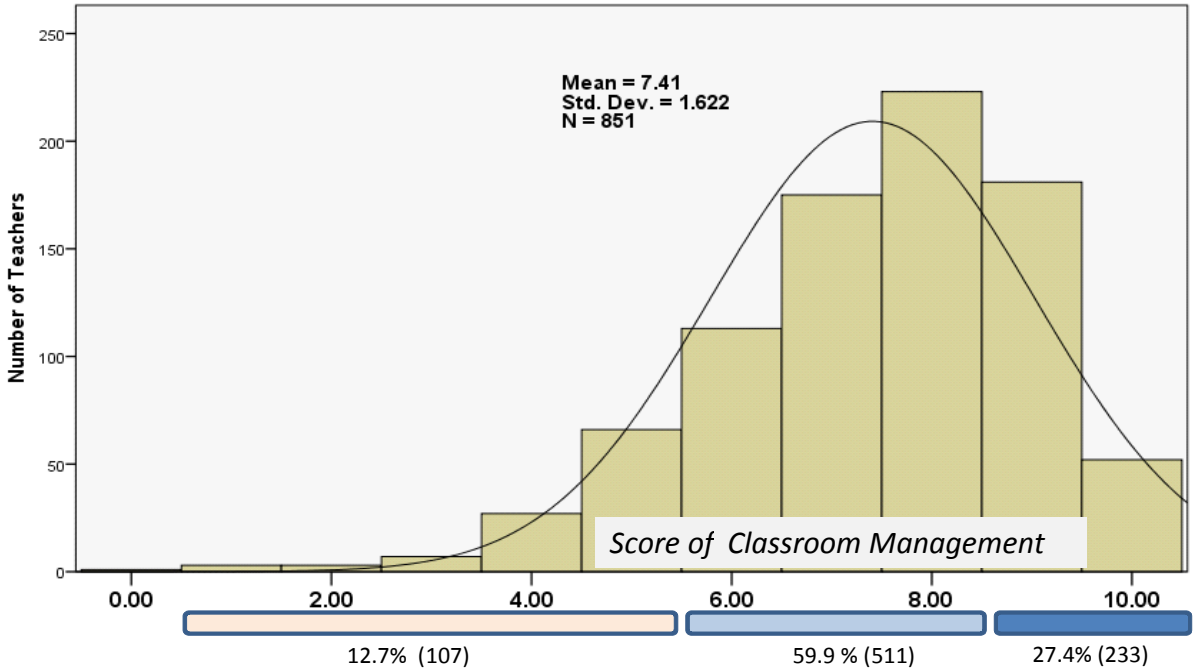
3.3 Classroom Management

Teachers who teach in schools should know how to effectively manage the teaching and learning process in classrooms regardless of which pedagogical approach he or she uses. In either a student-centered *or* a teacher-centered environment, teachers who can effectively manage the process of teaching and learning can effectively deliver the desired educational results.

To measure teachers' classroom management skills ten specific items were developed for this purpose in the observation study instrument. These items are simple but proxy measures of the domain of classroom management. With a reliability test conducted, a non-weighted statistical composite score[11] of "classroom management" was successfully developed to see how well Jordanian teachers manage their classrooms including time management, checking student attendance, presenting lesson objectives and plan and organizing orderly discussion. The score ranges from 0 (absolutely no classroom management skill at all) to 10 (a perfect classroom management score). Based on this 10-point scale, it was found that the average score for all sampled Jordanian teachers is 7.4 out of the total of 10 points with a standard deviation of 1.6 (see Chart 03 below). This indicates that teachers in Jordan are pretty good "managers" of classroom overall. 27.4% of teachers have scores of 9 or 10 as the most effective classroom managers, and only 12.7% of teachers have scores of 5 points or less as the least effective classroom managers.

[11] There are ten specific items used to measure the classroom management. We tested the item reliability of the ten "class management" items (Alpha=0.69) and computed a non-weighted summary composite score with a slightly skewed to the high score, with the mean of 7.4 and s.d. of 1.6.

Chart (03): Distribution of Composite Score of Classroom Management



Distribution of Classroom Management Score

Keeping in mind the classroom management score at a national level, the score between several dichotomous variables such as male and female, permanent and temporary were compared to see if there are any significant differences which may help us better understand potential factors that may be relevant for policy interventions or planning. Below are the composite score results by comparing separately three dichotomous variables: gender, job status, and locality:

Table (2): Composite Score of Classroom Management by Gender, Job Status, ERSP Project, and Locality

		Average Score (Composite 0-10)	N.	T, (P-value)
Teacher Gender	Female	7.65	471	
	Male	7.11	380	4.90, (.000)
Job Status	Permanent	7.45	746	
	Temporary	7.09	105	2.18, (.030)
Locality	Rural	7.43	251	
	Urban	7.39	592	0.30 (.765)

Note: If P-value for T-statistics is bigger than .05, difference in the average score between two paired categories is not statistically significant. In the above table, there is no significant difference in classroom management composite score between urban and rural teachers.

From the above table, the study finds that female teachers are significantly better than their male counterparts in terms of managing their classroom teaching. On average, female teachers scored 7.65 points while male teachers scored 7.11 points, indicating that female teachers have better classroom management skills than their male counterparts by 7.6 percentage point difference. This means that female teachers on average are better than male teachers in managing class time, tracking student attendance and their classroom activities, and conducting more orderly classes. Moreover, the study also finds that permanent teachers on average scored 7.45 while temporary teachers scored 7.09, a significant difference of 5.1%. This indicates that permanent teachers are on average better teachers than temporary teachers in terms of managing their classes. However, there is no significant difference between urban and rural teachers in their class management.

In the context of ERfKE II, the study also highlights comparative findings among various types of schools, the MoE schools, private schools, UNRWA schools, and MoD schools, as well as those schools which participated in educational project interventions (Discovery schools, Madrasati, ERSP). Table (03) shows the results of the composite scores of classroom management by the educational authority as well as the educational interventions within MOE schools. The results also showed a significant variation among schools under different authorities and projects (with ANOVA test, F-stats-2.6 and p-value=.017). The observed differences are from 7.22, the lowest for Discovery schools to 7.86, the highest for private schools. It is clear that Discovery, Madrasti and MoE non-intervention schools have lowest scores; ERSP, MoD and UNRWA schools are in the middle; and private schools have the highest although these differences are relatively small.

Table (03): Composite Score of classroom management by Authority & Educational Intervention

School Authority & Educational Intervention	Average Score	St deviation	N.
Discovery	7.22	1.61	89
Madrasti	7.29	1.58	89
ERSP	7.43	1.45	93
MOE No-intervention	7.23	1.78	281
MOD	7.49	1.40	86
Private	7.86	1.46	130
UNRWA	7.48	1.72	75

The teachers' classroom management score was examined by several other categorical variables (more than two categories) such as teacher qualification, years of experience[12], educational

[12] Years of experience, a continuous variable, is recomputed into 5 categorical variables with 1 for 0-5 years of experience, 2 for 6-10, 3 for 11-15, 4 for 16-20, and 5 for more than 20 years of experience.

qualification[13], subjects, and training in in-service training, ICDL or INTEL, but none of the analysis results show any significant differences between and among these multiple categories.

In sum, the study concludes that teachers in Jordan generally know fairly well how to manage classroom activities in terms of time management, checking student attendance, presenting lesson objectives and plan and organizing orderly discussion, etc. These however are considered as basic requirement for being a classroom teacher regardless which pedagogical approach or methodology one applies. Managing classroom well by the used criteria does not tell us much about how teachers use student-centered active learning and teaching or other traditional models of teaching.

3.3 Student-centered active learning and teaching

Student-centered active learning and teaching is regarded as progressive education methodology that teachers use in classrooms to facilitate student learning through critical thinking, problem solving, analytic inquiry, team work, and project-based learning skills in the student-centered classroom environment. The student-centered active learning and teaching model is globally recognized to equip students with the 21st century skills. Given that under the ERfKE pedagogical reform projects, most Jordanian teachers have been trained to use student-centered active learning and teaching methodologies in their classrooms, we are now at an important juncture to ask the following questions: How do the teachers teach today in Jordanian school classrooms? How differently do they teach? To what extent do they use student-centered active learning and teaching methodologies? Is the student-centered active learning and teaching methodology (SCALT) correlated with student learning achievement? In this section, the study intends to answer these questions with data evidence specifically focused on teaching practices of the student-centered active learning and teaching methodology.

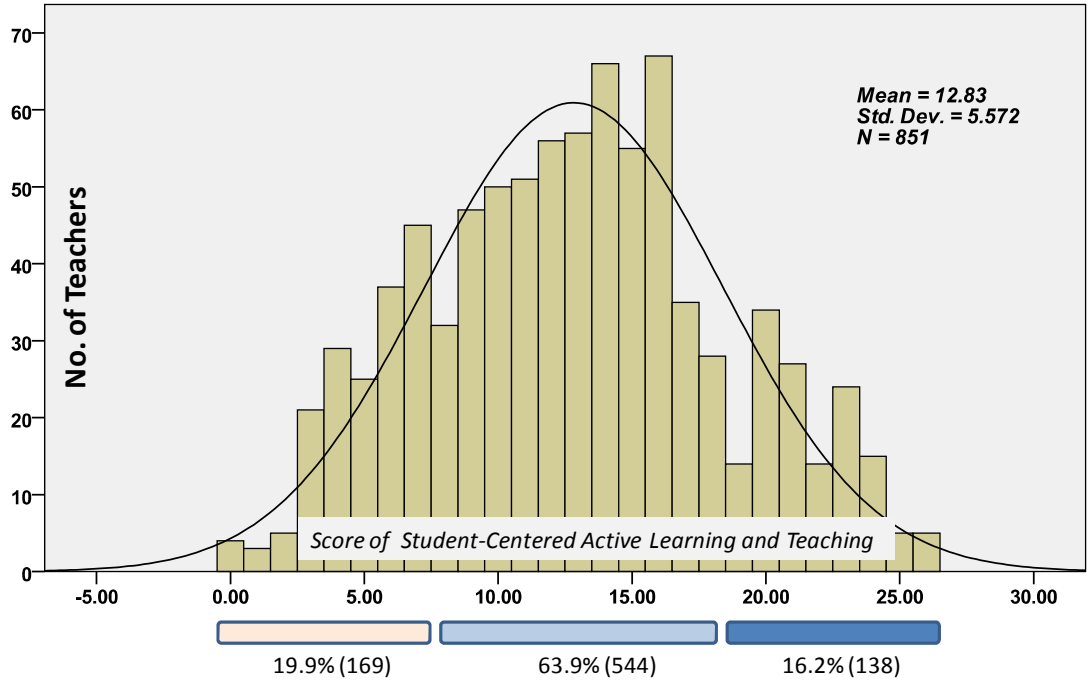
1) Composite measure of student-centered active learning and teaching

To start, the report first describes how the tool was developed and later features the analysis and findings. For the classroom observation tool, 26 items were developed specifically for measuring the overall domain of the student-centered active learning and teaching methodology. Within the domain, five sub-domains of critical thinking, problem solving, cooperative learning, participatory inquiry, and active learning were considered based on the MOE's policy interests. As discussed earlier, a two step analysis was conducted to ensure validity and reliability of the measures of the domain. First, a factor analysis was conducted within the domain to confirm if the original five sub-domains were well fitted by statistical tests (extracted eigenvalues). The results are convincingly reassuring that the larger domain of active learning and the five sub-domains are statistically confirmed as noted in section 2.2, under "Validity and Reliability of the Observation Instrument." Secondly, a reliability test was conducted to see if the 26 items were designed well enough to measure the overall domain of student-centered active learning and teaching. The reliability test shows alpha statistics of .886, which is highly reliable.

[13] There are 3 categories in educational qualification, diploma, master degree and PhD. Given only 4 teachers in our sample are PhD and 23 teachers in Master Degree, we decided that statistical data are not reliable for use.

Below is a distribution of the composite score of the domain of the student-centered active learning and teaching. The composite score identifies the variation of teachers' performance in using student-centered active learning and teaching methods. From the distribution, we may conclude that teachers on the right hand side of the chart are considered high performers of student-centered active learning and teaching while teachers on the left hand side are poor performers of student-centered active learning and teaching. Specifically, we used 1 standard deviation above and below the mean as two cut-off points to determine the excellent and poor performing teachers in this study. As indicated in the chart, 16.2% of teachers in our sample (138 teachers) are ranked as high performers in using student-centered active learning and teaching methodology, who obtained 19 points or higher. 19.9% of the teachers (169 teachers) are identified as poor performers in using student-centered active learning and teaching approach in classes. The majority of teachers in our sample (63.9%) are in the middle.

Chart (04): Distribution of Composite Score of Student-Centered Active Learning & Teaching



As we know that the current study is only a baseline study, we intend to conduct the same study in two years. By then we will be able to identify the *change* in the composite score of teachers' use of student-centered active learning and teaching methodologies. Educators would hope that there will be a significant increase in number of teachers who join the rank of high performance in using student-centered active learning and teaching methodologies in Jordanian schools. Until then, we have a good snapshot of the current teacher performance in using student-centered active learning and teaching methodologies in Jordan. We must note that those teachers who score poorly in using student-centered active learning and teaching methodologies (19.9% shown above) may prefer or feel more comfortable using traditional teaching and learning methodologies. However, we are unsure of this assertion as we believe that there may be other alternative methodologies used by teachers.

2) Variations by Gender, Job Status, Urban vs. Rural

Next, we would like to focus on a set of bi-variate analysis results to show mean differences in the composite score of student-centered active learning and teaching between groups, such as male and female teachers, temporary and permanent teachers, and urban and rural teachers. Note that a higher score means high performance in using student-centered active learning and teaching methodology while a lower score implies poor performance in using student-centered active learning and teaching methodology.

Table (4):
Composite Score of Student-centered active learning and teaching by Various Groups

		Average Score (Composite 0-26)	N.	T, (P-value)
Teacher Gender	Female	14.0	471	
	Male	11.3	380	7.3 (.000)
Job Status	Permanent	13.0	746	
	Temporary	11.4	105	2.7 (.004)
Locality	Urban	12.9	592	
	Rural	12.6	251	0.75 (.452)

Note: If P-value for T-statistics is bigger than .05, difference in the average score between two paired categories is not statistically significant. In the above table, there is no significant difference therefore in the composite score between urban and rural teachers.

Comparing male and female teachers, the study finds that female teachers on average have significantly higher composite scores in student-centered active learning and teaching than their male counterparts. The score for female teachers is 14.0 while the score for male teachers is 11.3, almost 24% point difference in the SCALT composite score in favor of female teachers. This indicates female teachers are performing significantly better than male teachers in using student-centered active learning and teaching methodology in Jordan. What causes female teachers to perform better than male teachers in student-centered active learning and teaching remains an intriguing question and deserves a more in-depth research study.

Permanent teachers perform better than temporary teachers in using student-centered active learning and teaching by a 14% point difference (actual score difference is 1.6). As we cross-examine job status against variables such as training, qualification and years of experience, the study finds that the difference in the composite scores of student-centered active learning and teaching between permanent and temporary teachers could be attributable to better training, higher qualification, and longer years of experience, associated with permanent teachers.

Between urban and rural teachers, there is no significant difference in their composite scores of student-centered active learning and teaching. This is somewhat counter-intuitive or intriguing finding for us as we hypothesized that urban teachers would do better than rural teachers, but the result shows no difference.

Examining the composite score of student-centered active learning and teaching by authority and project type, the study finds that ERSP is significantly higher than all the other schools. ERSP school is ranked the first and private schools the second. ERSP's score is 9% better than that of private school in the score of student-centered active learning and teaching and is almost 23% better than the lowest, schools under the Ministry of Defense (see Table 05 below). This indicates that teachers in ERSP schools show a remarkable lead in using student-centered active learning and teaching methodologies in Jordan.

Table (05): Composite Score of Student Centered Active Learning and Teaching by Authority & Educational intervention

School Authority & Educational intervention	Average Score	St deviation	N.
Discovery	12.60	5.56	89
Madrasti	12.70	6.09	89
ERSP	14.39	5.70	93
MOE None intervention	12.53	5.89	281
MOD	11.72	4.49	86
Private	13.21	5.06	130
UNRWA	12.93	5.29	75

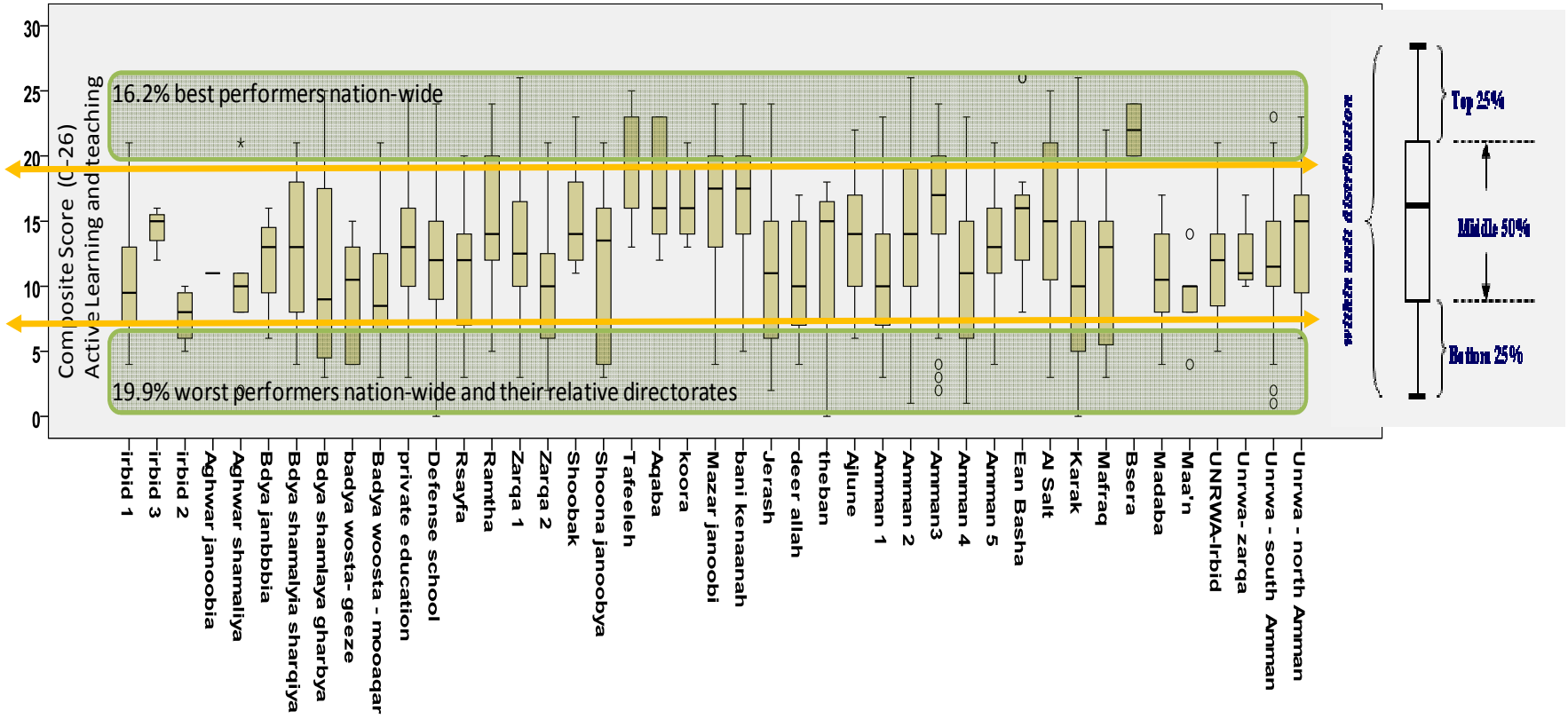
3) Variations in student-centered active learning and teaching at the directorate Level

The selected teachers represent all field directorates in Jordan. An unavoidable question is how different teachers are in using student-centered active learning and teaching approach among all of the directorates. Below is a directorate level distribution of the composite score of student-centered active learning and teaching, indicating how variant teachers are within and between directorates.

For each directorate in the box plot below, one may notice the wide range on the score bar. As the chart legend on the right hand side of the chart indicates, the box in the middle represents 50% of all teachers' performance ranges within that directorate. The line above the box represents the 25% best performers within the directorate, and line below the box represents the 25% poorest performers within the directorate. The stretched longer box and lines indicate that there is a larger variance within the directorates in teachers' use of student-centered active learning and teaching methodologies while the shorter box and lines show less variance within the directorates in teachers' use student-centered active learning and teaching methodologies

Chart (05):

Composite Score (Boxplots) of Student-Centered Active Learning and Teaching by Directorates



Boxes located high up on the chart show a range of high scores in teachers' use of student-centered active learning and teaching (e.g. Bsera and Tafeeleh). Boxes located low on the chart indicate a range of poor scores in using student-centered active learning and teaching methodologies (e.g., Irbid 2, Shoona-janoobya, Karak). In addition to that, two large horizontal lines were drawn indicating the two cut-off points on the composite score, the top line is the cut-off point where all performance score above it are considered the best performing teachers (16.2%) in teaching student-centered active learning and teaching.

The bottom line indicates the cut-off point where all teachers below the line are poor performing teachers in student-centered active learning and teaching, located in many directorates.

The overall box plot above shows there are quite large variances between directorates as well as within directorates. In other words, many directorates have their own large range of performance scores in teachers' use of student-centered active learning and teaching while overall they vary significantly between directorates. With this data evidence, the study concludes that many directorates have good performing teachers as well as poor performing teachers in using student-centered active learning and teaching methodology in Jordan. One policy implication may be that MOE encourages a wide exchange of experiences and learning within each directorate as well as between directorates in the use of student-centered teaching and learning methodologies. Poor performing teachers may learn from effective teachers within their own directorate as well as from other directorates. This may also tease out another policy inquiry, what has created such diverse performance gaps within directorates and between directorates? Do teacher characteristics explain some of the diverse variance in the composite score of student-centered active learning and teaching? This study has also reported this next.

4) Factors that explain the variance of the composite score of student-centered active learning and teaching

Earlier in the report, a large variation in the composite score of student-centered active learning and teaching was presented indicating how teachers in Jordanian schools applied the student-centered active learning and teaching methodologies. The variation suggests that teaching practices varies quite significantly. The study examined number of factors to explain why this variation. For example, will teacher characteristics be able to explain the variance in the composite score of student-centered active learning and teaching and to what extent? Will training explain some? To answer the questions, a regression analysis was conducted and the results are worth noting.

The dependent variable is the composite score of student-centered active learning and teaching (ranging 0-26)[14]. The independent variables (or explanatory variables) for teacher characteristics in Jordan that we chose to add to our final regression model are as follows:

a. teacher gender
b. teacher qualifications
c. job status (permanent or temporary)
d. years of work experience
e. academic subject (Math/Science or Arabic)
f. participation in ICDL training
g. participation in INTEL training
h. urban and rural
i. grade (5, 9, and 11)

[14] The histogram of the composite score is an almost perfect normal distribution. By a rule of regression requirement of normality in dependent variable, it is almost perfect dependent variable.

Below shows the results of the regression model:

Table (06): Regression Analysis Model: Composite Score of Student-centered Active learning and teaching as Dependent Variable

Regression Model	β [15]	β1 [16]	S.E. (β)	T-stats (P-value)
Constant (or Intercept)	**9.48**		**.703**	**13.48 (.000)**
Teacher Gender (F=1, M=0)	2.95	.26	.383	7.71 (.000)
Teacher Job (Permanent=1,Non-permenant=0)	1.26	.07	.597	2.12 (.035)
Yrs of Exp (5 yrs or less =0, more than 5 years=1)	.872	.08	.421	2.07 (.039)
Mathematics (Math=1, Science & Arabic=0)	-1.48	-.13	.444	3.34 (.001)
Science (Science=1, Math & Arabic =0)	-.543	-.05	.427	1.17 (.241)
Training ICDL (ICDL=1, Non-ICDL=0)	.491	.04	.430	1.14 (.253)
Training INTEL (INTEL=1, Non-INTEL=0)	.652	.04	.556	1.17 (.241)
School Location (Rural=1, Urban=0)	.609	.05	.421	1.45 (.148)
Grade 11 (G11=1, G9=0, G5=0)	.434	.04	.474	.916 (.360)
Grade 9 (G9=1, G11=0, G5=0)	.136	.01	.443	.307(.759)
R^2= .098				
F-statistics =9.05				
P-value= .000				

From the regression analysis results presented in Table 6, the study finds that four significant factors that help explain the variation in the composite score of student-centered active learning and teaching (SCALT) and six insignificant factors. However, R-square is small (r^2= .098), indicating there is about 10% of variation in teachers' SCALT composite score that has been explained by the 10 factors in the model. That suggests there might be other factors that can be identified in future research.

Gender is the most significant factor in explaining variation in teachers' SCALT scores[17]. The ? coefficient associated with gender is 2.95 indicating the magnitude of the difference in composite scores (dependent variable) between male and female teachers (T-stats=7.71, P-value=.000). Since female is coded 1 and male is coded 0, female teachers obtained significantly higher SCALT composite scores than their male colleagues on average, even after controlling for all other 9 factors in the model.

[15] Unstandardized coefficients are presented in this column for the purpose of presentation of the relationship between each factor and the composite score.

[16] Standardized coefficients are presented in this column for the purpose of ranking the size of factors' influence on the composite score.

[17] For measuring the size of each factor's contribution, we used the unstandardized coefficient for each factor.

An examination of teacher job status (permanent, non-permanent) shows that being a permanent teacher is associated with an unstandardized 1.26 [?] coefficient controlling for the other factors (T-stats=2.12, P-value= .035). In addition, the coefficient associated with years of experience (0.872) suggests that more experienced teachers tend to obtain slightly higher, but significant, scores in student-centered active learning and teaching than teachers who are less experienced.

Subjects also matter. It is largely expected and confirmed that teachers who teach Math are less likely to apply student-centered active learning and teaching methodologies than teachers who teach Science and Arabic. Given the methodologies associated with active learning and teaching, such as cooperative learning, critical thinking, and group discussion, literature and language arts, and many topics under Science subjects could be more naturally conducive to progressive teaching methodologies than math. The significant negative coefficient in math (as compared to Science and Arabic) indicates that Math teachers scored significantly lower than Science and Arabic teachers. Those results suggest that math teachers' scores on the SCALT composite were, on average, 1.48 points lower than their Science and Arabic counterparts, after considering all the other factors equal.

The remaining factors, training ICDL, INTEL, location (urban vs. rural), and grades have no significant effects on the composite score after controlling for all other variables in the model (p-values= .253, .241, .148, .360, and .759, respectively).

In sum, the results of the regression analysis inform us that, after taking into consideration the effects of all factors included in the model above, the following is true:

1) Gender is the most significant factor in explaining variance in teachers' SCALT scores. Female teachers are significantly more likely to apply student-centered active learning and teaching methodology than their male counterparts in Jordan.
2) Teachers' job status (permanent/temporary) have a positive effect on teachers' SCALT score. Years of experience also affects teachers' SCALT score, although the magnitude of the effect is smaller than job status. Permanent teachers and teachers with longer than 5 years of service are more likely to apply SCALT methodology.
3) Among 3 subjects, math teachers are least likely to apply SCALT methodology in their classroom even controlling for other factors.
4) Factors such as ICDL and INTEL training, school location (urban vs. rural), and grades taught by teachers (5th, 9th, or 11th) have no significant effect on teachers' SCALT composite score after we take into consideration the effects of the other factors.

From the results of the regression analysis, we understand that we continue to face the following policy inquiries that may require more research and analyses[18]:

[18] Currently, we are trying to integrate multiple databases to aim to address these inquiries.

- Why are female teachers more likely to apply SCALT methodologies than male teachers even after we consider other factors such as qualifications, years of work experience, or subjects taught?
- How does the application of SCALT in classrooms correlate with student achievement in NAfKE? and
- Why don't variations in teacher training explain the variation in classroom application of SCALT after considering all the other factors?

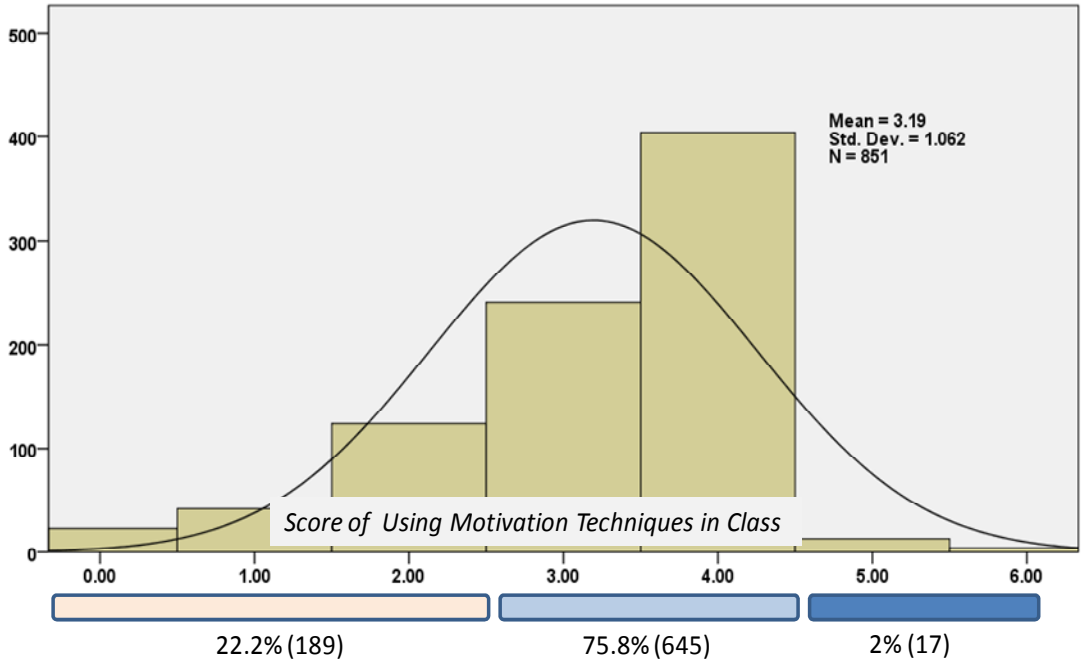
We must point out that these findings from the regression analysis may appear contradictory to our findings from T-tests or ANOVA analysis presented earlier in the report. However, they are not contradictory. In the findings presented earlier from the results of T-test and ANOVA, the aim was to compare mean differences between two or more groups to see if the observed mean differences could be significantly large enough to infer to the population (all public school teachers in grades 5, 9 and 11) in Jordan. In the regression analysis, we aim to see if a factor significantly explains the variation in SCALT composite score considering all other factors in the model equal.

3.4 Motivation

To stimulate student motivation in class is a teacher's essential responsibility in teaching. Encouraging student creativity as well as performance, stimulating participation, and showing satisfaction with student efforts independent of quality all contribute to increasing student motivation. This is the area of our investigation for this section.

To measure teachers' motivation six specific items were developed for this purpose in the observation study instrument. The summed scaled score was computed. The score ranges from 0 (absolutely no motivation at all) to 6 (perfect motivation). The distribution of the score is truncated at point 5 indicating very few teachers reaching 5 or 6 or using the methods to stimulate student learning motivation in class. Based on this 6-point scale, it was found that the average score for all sampled Jordanian teachers is 3.2 out of the total of 6 points (with standard deviation of 1.1). See the table below.

Chart (6): Distribution of Composite Score of Practicing "Motivation" Techniques



In order to distinctively identify "most effective" and "least effective" teachers in stimulating student motivation, all sampled teachers were divided into three categories. By taking 1 standard deviation plus and minus the mean, we classified teachers who scored five and six as "most effective" and teachers who scored one and two as "least effective" in stimulating student motivation. In the analysis, study finds that in Jordan, only very small 2% of all sampled teachers are considered as "effective" at in stimulating student motivation. 22.2% of the sampled teachers are not doing much in class to encourage students' motivation. 75.8% are considered as "average motivators". Also between male and female teachers, there is no significant difference in "most effective" teachers in terms of boosting student motivation. Given the results, the study concludes that teachers in Jordan must significantly improve their techniques in motivating student learning in class as we deem it one of the most important aspects of teacher responsibility in class. This may require specialized teacher training on how to motivate student learning in class and become enthusiastic in teaching by themselves as role model.

For this, we compared the score between several dichotomous variables such as male and female, permanent and temporary, and school location to see if there are any significant differences that would help us better understand potential factors that may be relevant for policy interventions or planning. Below are the composite score results by comparing four dichotomous variables, gender, job status, and school location respectively:

Table (7):
Composite Score of Motivation by Gender, Job Status, and School Location

		Average Score (Composite 0-6)	N.	T, (P-value)
Teacher Gender	Female	3.26	471	
	Male	3.11	380	1.99, (.04)
Job Status	Permanent	3.25	746	
	Temporary	2.83	105	3.61, (.00)
School location	Urban	3.17	592	
	Rural	3.24	251	0.51 (0.61)

Note: If P-value for T-statistics is bigger than .05, difference in the average score between two paired categories is not statistically significant. In the above table, there is no significant difference in motivation composite score between urban and rural.

From the above table, we find that female teachers are significantly better than their male counterparts in terms of their efforts to motivate students. On average, female teachers scored 3.23 points while male teachers scored 3.11 points, indicating that female teachers make slightly higher efforts to motivate their students than their male counterparts, by about 5% difference. This means that female teachers on average are better than male teachers in praising students for their work, less likely to mock students after they make mistakes, and regularly encouraging student participation and opinions in class.

In addition, the study examined how permanent and temporary teachers scored differently in this composite measure. Permanent teachers on average scored 3.25 while temporary teachers scored 2.83, a significant difference of 12.9%. This indicates that permanent teachers are on average better teachers than temporary teachers in terms of motiving students to learn.

We also examined teachers' motivation score by several other categorical variables (more than two categories) such as years of experience[19], highest educational qualification, highest scientific degree[20], subjects, and training in ICDL or INTEL. Following are the composite score results:

[19] Years of experience, a continuous variable, is recomputed into 5 categorical variables with 1 for 0-5 years of experience, 2 for 6-10, 3 for 11-15, 4 for 16-20, and 5 for more than 20 years of experience.

[20] There are 3 categories in educational qualification: bachelor's degree, master's degree and PhD. Given only 4 teachers in our sample have PhDs ,and 23 teachers master's degrees, we decided that statistical out are not reliable for use.

Table (8): Composite Score of Motivation by Teacher Qualification, Years of Experience, Subject, Project Intervention and Training in ICDL or INTEL

		Average Score (Composite 0-6)	N.	F, (P-value)
Highest educational qualification	Bachelor's degree	3.36	123	
	Master's degree	3.14	23	2.60, (.020)
	PHD	2.75	4	
Years of experience	0-5 years	3.05	324	
	6-10 years	3.24	208	2.39, (.02)
	11-15 years	3.26	126	
	16-20 years	3.14	99	
	More than 20 years	3.09	94	
Subject	Arabic	3.12	297	
	Math	3.05	298	1.03, (.37)
	Science	3.11	256	
Training in ICDL	Yes	3.26	500	
	No	3.10	351	3.72,(.000)
Training in INTEL	Yes	3.23	127	2.39,(.017)
	No	3.12	724	

Note: If P-value for F-statistics is smaller than .05, difference in the average score between two paired or more categories is statistically significant. In the above table, there is no significant difference in motivation composite score between teachers of various subjects.

From the above table, we find that teachers holding bachelor's degrees are significantly better at stimulating student motivation in the classroom than teachers holding higher level degrees. On average, teachers holding bachelor's degrees scored 3.36 points while teachers holding PhDs scored only 2.75 points. We must note that there are only 4 teachers in our sample with a PhD degree, making the comparative statistics difficult to determine at this stage. However, we believe that this sharp difference may still hint a real truth in reality that teachers with a PhD may not necessarily know how to stimulate student motivation for learning.

Examining years of work experience, we found that teachers with medium years of experience (6-15) have the best performance in stimulating student motivation in the classroom. Both groups of teachers with five years or less and those with 16 or more years of experience are significantly poorer

performers in stimulating students to learn. This leaves us with an intriguing question, what factors may explain the fact that teachers with 6-15 years of experience are the best at stimulating student motivation in Jordan? This deserves a further and more in-depth study.

Based on the analysis, there is a clear link between training and improvement in a teachers' ability to stimulate student motivation. All teachers who went through the ICDL and INTEL training courses are significantly more likely to stimulate student motivation in classroom teaching.

We have found no significant differences in teachers' classroom practice to motivate students among schools under different authorities or project interventions (see table below).

Table (9): Composite Score of Motivation by Authority & Educational intervention

School Authority & Educational intervention	Average Score	St deviation	N.
Discovery	3.14	1.06	89
Madrasti	3.07	1.29	89
ERSP	3.25	0.99	93
MOE None intervention	3.16	1.07	281
MOD	3.30	0.99	86
Private	3.29	1.01	130
UNRWA	3.21	1.01	75

3.5 In-Classroom Student Assessments

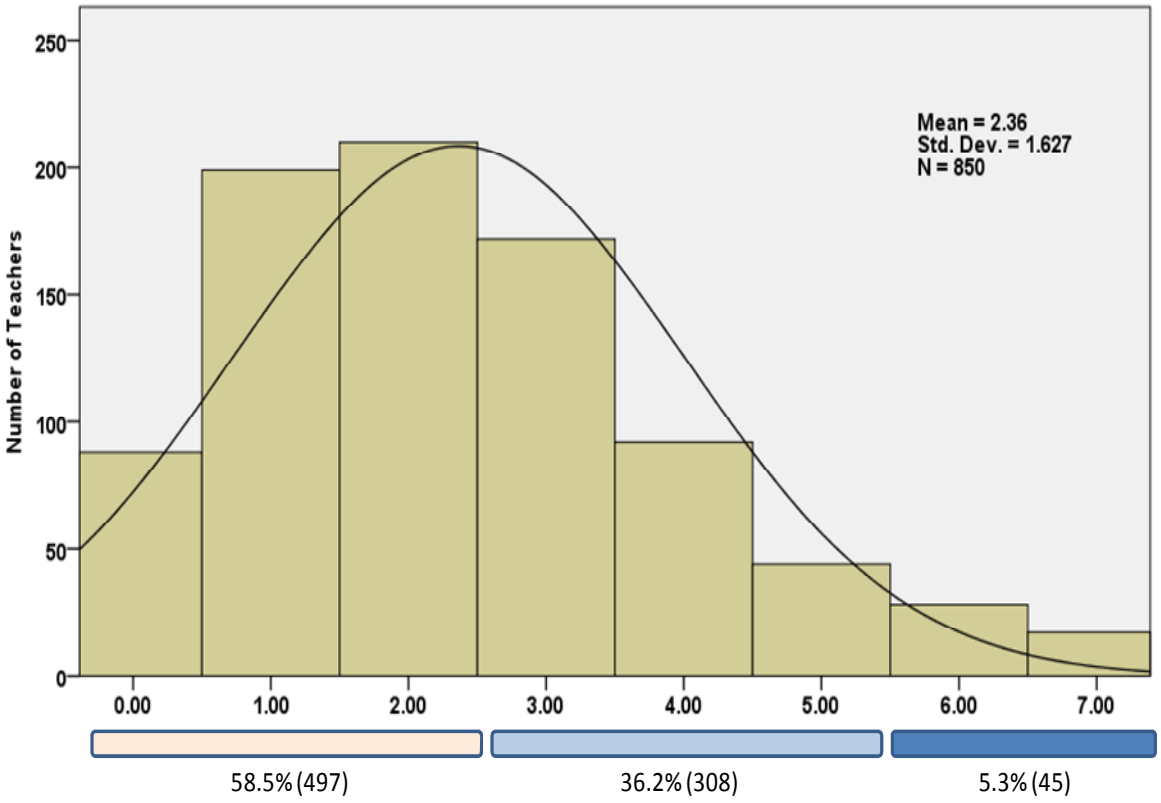
The use of in-classroom student assessment techniques is critical to ensure quality learning outcomes. Teachers can use in-classroom student assessments not only for regular assessment of student progress and understanding of the material, but also to provide feedback to students to encourage student learning, performance and improvement. Teachers can also use student self-assessment and peer-to-peer assessment techniques to encourage individual and peer-to-peer learning and motivation.

To measure teachers' use of in-classroom student assessment techniques seven specific assessment items were developed in the observation study instrument. Together they were designed to measure a sub-domain of teacher practice of using student assessment. With a reliability test conducted, we successfully developed a statistical composite score[21] of "student assessment" to see how well Jordanian teachers used assessment techniques to work with students in class. The score ranges from 0 (never used any assessment) to 7 (constantly assess and feedback). Based on this 7-point scale, we found that the average score for all sampled Jordanian teachers is 2.4 out of the total of 7 points (with

[21] There are seven specific items used to measure the students assessment. We tested the item reliability of the seven " students assessment " items (Alpha=0.69) and computed a non-weighted summary composite score with a histogram skewed to the low end, with the mean of 2.36 and s.d. of 1.63. This composite score is not normally distributed.

standard deviation of 1.6). Low average scores and large standard deviation indicate that most teachers rarely use classroom assessment tools, nor do they encourage peer-to-peer assessments. Although there is not yet a specific national expectation for this domain, we hope that this analysis establishes a baseline measure for future use of in-classroom student assessment techniques in Jordanian schools. If we rate teachers with composite scores of six and seven as the most effective teachers in this aspect, then only 5.3% of all teachers practice in-classroom student assessment. In fact, 58.5% (scored between 0-2 points) of all teachers have never in student assessment techniques in the classroom. This is quite low performance in this regard.

Chart (7): Distribution of Composite Score of Using Student Assessment Techniques



Keeping in mind the low scores in student assessment at the national level, we ran additional comparisons between multiple categorical variables: male and female teachers, permanent and temporary teachers, and rural and urban localities, training in ICDL/INTEL (yes/no), subject (Arabic, Math, Science), and project intervention schools. Below are the resulting composite score results by variable:

Table (10): Composite Score of Student Assessment by Multiple Categorical Variables

		Average Score (Composite 0-10)	N.	T or F, (P-value)
Teacher Gender	Female	2.60	471	4.78, (.000)
	Male	2.07	379	
Job Status	Permanent	2.39	746	1.43, (.153)
	Temporary	2.15	105	
Location	Rural	2.41	251	.527,(.598)
	Urban	2.34	591	
Training in ICDL	Yes	2.47	500	2.30,(.021)
	No	2.21	350	
Training in Intel	Yes	2.68	127	2.35,(.019)
	No	2.31	723	
Subject	Arabic	2.33	297	4.47,(.012)
	Math	2.57	297	
	Science	2.16	256	

Note: If P-value for T-statistics is bigger than .05, difference in the average score between two paired categories is not statistically significant. In the above table, there is no significant difference in student assessment composite score between temporary and permanent, and rural and urban.

From the above table, we find that once again, female teachers are significantly better than their male counterparts in terms of using in-classroom student assessment techniques. On average, female teachers scored 2.60 points while male teachers scored 2.07 points, indicating that female teachers are significantly more likely to use student assessment techniques than their male counterparts. Although low actual score, the difference in the score between male and female is 25%. This indicates that female teachers on average are much more likely than male teachers to, use checklist and rating scales for assessing student performance, encourage peer-to-peer assessment, discuss homework in class, and provide students feedback on their class performance.

In addition, the study finds how teachers of Arabic, Math, and Science subjects scored differently in this composite. Math teachers on average scored 2.57, Arabic on average scored 2.33, and Science on average scored 2.16, a significant descending variation among the three subjects respectively. This indicates that teachers on average use more in-classroom student assessment techniques in Math than in Arabic, and teachers of Science scored the worst.

As with other domains, the analysis shows that teachers with training are significantly more likely to use in-classroom student assessment techniques than teachers without training. Teachers with and without ICDL scored very differently in this composite measure. ICDL trained teachers on average scored 2.47 while non-ICDL teachers scored 2.21, a significant difference of 12%. This indicates that training in ICDL teachers are on average better in using in-classroom student assessment techniques than non-ICDL teachers. With training in INTEL, teachers scored 2.68 while non-INTEL teachers scored 2.31, a significant difference of 16%. The evidence above tell us that teachers with the training (either ICDL and/or INTEL) on average are significantly more likely to use in-classroom student assessment techniques to improve student learning than teachers without the training.

However, when we looked student assessment scores between permanent and temporary teachers, no significant difference was found. The observed small difference in mean scores between permanent and temporary teachers could simply be a random error. Comparing rural and urban teachers, no significant difference was found present.

When we analyze teachers' use of in-classroom student assessment techniques by school authority and project type, we find that ERSP ranks the first with 2.77 and UNRWA ranks the second with 2.63. By this criteria the lowest was the discovery schools with 2.08. Results show clearly that teachers in ERSP schools are significantly more likely to use in-classroom student assessment techniques to help students learn than teachers in non-ERSP schools. In fact, teachers in Discovery schools, MADRASATI schools, are significantly less likely to use student assessment techniques than teachers from all other schools.

Table (11): Composite Score of Students Assessment by Authority & Educational intervention

School Authority & Educational intervention	Average Score	St deviation	N.
Discovery	2.08	1.59	89
Madrasti	2.38	1.63	89
ERSP	2.77	1.67	93
MOE None intervention	2.25	1.63	281
MOD	2.37	1.56	86
Private	2.33	1.69	130
UNRWA	2.63	1.49	75

It was noted that the overall is very low, leaving a large room for further improvement.

3.6 Use of Instructional Technology and Media Aids

The use of technology in the classroom has dramatically increased all over the world in recent years. Jordan is no exception. Results of a separate study showed that computer equipment, e-learning applications (e.g. Edu-Wave program), and media aids are available to some extent in many MOE schools. It is the one of the objectives of our study to ascertain to what degree teachers in Jordanian schools use modern technological resource in their daily teaching routine.

The following table shows the percentages of teachers who used modern instructional technology in class while being observed. This baseline snapshot of teachers' utilization of computer technology shows that very rarely do teachers in any of the core subjects, Arabic, Math, and Science, use computer technology as part of their teaching methodologies.

Table (12): % of Teachers Who Use Computer Technology in the Classroom by Gender and Subject

Observation Checklist	Male teachers	Female teachers	Arabic teachers	Math teachers	Science teachers
✓ Use of computer in teaching	4%	13%	7%	7%	3%
✓ Use of projected screen	4%	11%	7%	7%	3%
✓ Use of computer for assessment	2%	3%	2%	2%	2%
✓ Teacher uses Edu-Wave	2%	5%	3%	3%	2%
✓ Student use of computers	2%	8%	6%	6%	2%
✓ Use of Powerpoint or Excel	3%	8%	5%	5%	2%
✓ Use of laptop computer	1%	3%	2%	2%	1%

From other studies, it was noticed that many teachers were trained in using computers or e-learning contents through Edu-Wave. 54% of male teachers and 62% of female teachers were trained in the ICDL program and 17% of male teachers and 13% of female teachers were trained in INTEL program. However, by measuring daily use of technology in classroom teaching, rarely do teachers in Jordan use the technology or integrate the technology into their curriculum. We fully understand that this is a snapshot to create a baseline for future comparison purpose. However, observing a large number of classes in grades, 5, 9, and 11 on three core subjects, Arabic, Math and Science, provides a good picture of the reality of teachers' utilization of computer technology in the classroom. We will not be able to pin point the exact causes of the low level of use of computer technology in this study, but the problem must be addressed. Is it lack of computer technology for use in classrooms or lack of knowledge and skills in using computer technology in classrooms? These are questions for a further study.

In addition, we observed how teachers used non-computerized instructional teaching tools in the classroom. It was hypothesized that teachers would use more traditional teaching tools since they make little use of computer technology. Table 09 below shows the percentages of teachers who used traditional teaching tools (non-computer technology) in classrooms. It is clear from this table that teachers are much more likely to use the traditional instructional tools to help students learn in class as most of their professional development training on the use of traditional tools in class.

Table (13): % of Teachers Who Use Traditional Teaching Aid in Class by Gender and Subject

	Observation Checklist	Male teachers	Female teachers	Arabic teachers	Math teachers	Science teachers
✓	Use of maps, paper models, etc.	28%	43%	40%	29%	41%
✓	Use of handouts and materials	18%	41%	36%	21%	34%
✓	Responsive to inquires about educational aid in class	25%	44%	36%	28%	43%

It is also worth noting that female teachers are more likely to use teaching aids in the classroom than their male counterparts regardless of whether it is using computer technology or more traditional tools. This has been consistent across the study in almost all aspects of measures.

3.7 Observers' Evaluation Results

In addition to the classroom observations, each observer provided his or her own summary opinions on how each teacher performed in the classroom. The purpose of examining observers' opinions of the observed class on the same domain issues is to let observers provide useful aggregate information on teachers' performance as well to correlate the observers' opinions with the observed evidence to see if observers' opinions are in line with the observation results. In the observation tool, observers specifically were asked to rate relevant domains – classroom management, student-centered active learning and teaching, use of tools and teaching aids, student motivation, and alternative assessment – with a four-point scale – failure, needs improvement, satisfactory, and outstanding. Although only a few items were developed for each domain, these provided valuable information for research reference. Given that these observers are directorate-level subject matter supervisors and inspectors, their opinions are highly regarded.

Classroom management

Four statements under classroom management were given to observers to rate according to the standard four point scale explained above: time management, student behavior management, respectful interactions, and class (lesson) organization. The results from 47 observers are mostly positive (See Table 14). Regarding time management, 42.4% of teachers were rated by observers as satisfactory and 40.4% as outstanding, while only 12.2% were rated as needing improvement and 5% as failures. As for student behavior management, 39.1%of teachers were rated as satisfactory and 49.6% as outstanding respectively and only 8.7% as needing improvement and 2.6% rated as failures in this aspect. On respectful interactions between teacher and students, 28.4% of teachers were rated as satisfactory and a high 64.2% as outstanding while only 6.1% were rated as needing improvement and a mere 1.3% rated as failures. Finally in terms of lesson organization, observers rated 45.9% of teachers as satisfactory and 39.5% as outstanding, and only 12.6% as needing improvement and 2% as failures.

Table (14): Observers' Rating on Classroom Management

	Failure	Needs Improvement	Satisfactory	Outstanding
Time Management	5.0%	12.2%	42.4%	40.4%
Student Behavior Management	2.6%	8.7%	39.1%	49.6%
Respectful Interactions	1.3%	6.1%	28.4%	64.2%
Lesson Organization	2.0%	12.6%	45.9%	39.5%

Regarding the observers' rating on classroom management, we find performance to be greatly overrated. Earlier in the report, the study showed that only 6.1% of teachers were considered as excellent or outstanding classroom managers based on our 10-point scale observation composite score of classroom management. Although we are fully aware of different scales between the observed and rated data, the consistency in all items measuring both scales clearly show the overrating by the observers.

Student-centered active learning and teaching

Three specific statements were given to observers to rate according to the standard four-point scale:

1) use of diverse instructional methods,
2) use of student-centered approach, and
3) encouraging students to participate in class.

On the use of diverse instructional methods, 39.3% of teachers were rated by observers as satisfactory, while only 6.7% as outstanding. Similarly, on the reverse side, 46.6% were rated as needing improvement and 7.3% were rated as failures (see Table 15). Although this shows clearly that majority of teachers may need improvement in using diverse instructional strategies. As for use of student-centered approach, observers rated 50.2% of teachers as satisfactory and 9.5% as outstanding, while they rated 33.4% as needing improvement and 6.8% as failures. In terms of encouraging students to participate in class, observers rated 52.0% of teachers as satisfactory, and 16.6% as outstanding, 28.6% as needing improvement and only 2.8% as failures.

Table (15): Observers' Rating on Student-centered active learning and teaching

	Failure	Needs Improvement	Satisfactory	Outstanding
Diverse Teaching Methods	7.3%	46.6%	39.3%	6.7%
Use of Student-Centered Approach	6.8%	33.4%	50.2%	9.5%
Encouraging Students to Participate	2.8%	28.6%	52.0%	16.6%

The study finds that observers' ratings on student-centered active learning and teaching is well in line with the analysis results based on the 26-point observation composite score of student-centered active learning and teaching. Earlier in the report, the result shows that 16.2% of all observed teachers scored high and were considered as the best teachers in using student-centered active learning and teaching methodology in class.

Teachers' Motivation of Students

There are three areas used for observers to rate teachers' motivation of students:

1) showing intensive interest in motivating students,
2) showing enthusiasm in teaching, and
3) students' enthusiasm in learning.

The results are worth noting (see Table 16 below). On teacher's intensive interest in motivating students, observers rated 48.2% of teachers as satisfactory and 26.3% as outstanding while only 19.3% as needing improvement and 6.1% as failures. Regarding teachers' enthusiasm in teaching, observers rated 47.2% of teachers as satisfactory and 37.0% as outstanding, while only 13.1% as needing improvement and 2.7% as failures. As for students' enthusiasm in learning in class, 52.8% and 20.7% of teachers were rated as satisfactory and outstanding respectively while 23.6% and 2.8% of teachers were rated as needing improvement and failure respectively.

Table (16): Observers' Rating on Encouragement for Student Motivation

	Failure	Need Improvement	Satisfactory	Outstanding
Keen to Motivate Students	6.1%	19.3%	48.2%	26.3%
Teacher's Enthusiasm in Teaching	2.7%	13.1%	47.2%	37.0%
Students' Enthusiasm in Learning	2.8%	23.6%	52.8%	20.7%

Compared to the low performance by teachers in this aspect based on the observation data, the observers' rating is highly overrated. Based on the observation data, only 2% of all sampled teachers were considered as "effective motivators" in class. This clearly shows a large discrepancy between the observation data and observers' opinion.

Teachers Utilization of In-Classroom Student Assessments

The new curriculum standards in Jordan require teachers to use several types of in-classroom student assessments to boost student learning. In addition to quizzes, tests, teachers may also encourage peer review and assessment, individual feedback, reviewing homework, among other options. Two specific summary statements were asked of observers for their teacher performance ratings:

1) use of various in-classroom student assessments, and
2) use of in-classroom student assessment results to improve learning outcomes (See Table 17 below).

On use of various assessments, observers rated 24.8% and 21.8% of teachers as satisfactory and outstanding respectively while 41.6% as needing improvement and 11.8% as failures. In terms of using assessment results to support student learning, 45.3% of teachers were rated as satisfactory, only and 9.4% as outstanding , and 30.6% as needing improvement and 14.7% as failures.

Table (17): Observers' Rating on Utilization of Student Assessment

	Failure	Needs Improvement	Satisfactory	Outstanding
Use of Various Assessment	11.8%	41.6%	24.8%	21.8%
Use of Assessment Results	14.7%	30.6%	45.3%	9.4%

Again it was found that observers' rating on utilization of student assessment methods is overrated in comparison with the result based on the observation data evidence. Based on the composite score of utilizing student assessment techniques in class, we found that only 5.3% of teachers were considered as outstanding scoring six or seven points out of seven total composite score.

3.8 Correlations between NAfKE Scores and SCALT Results

A major question for this study is that does *SCALT score correlate with student achievement in NAfKE?* Although the relationship framework and hypothesis is debated among education scholars around the world, it is critical and important to address this relationship question in this study. To answer this question, we matched students' NAfKE score with their teachers' SCALT score by subjects and grades to test whether or not there was a correlation between the two composite variables. After the analysis, we present the following:

a) Variations in means of NAfKE Scores and SCALT scores by Grade and Subject

Table 18 shows a set of simple average means for students' NAfKE and teachers' SCALT scores by subject and grade. Note that a higher score means high performance while a low score implies poor performance in those subjects and grades. NAfKE scores ranged from 0—100. SCALT scores ranged from 0—26.

Table (18): Students' mean NAfKE scores and their teachers' mean SCALT scores by grade and subject.

Subjects	Grades	Students' NAfKE Scores M (SD)	*n*	Teachers' SCALT Score M (SD)	*n*
Math	5th	28.5 (14.8)	893	11.7 (5.6)	99
	9th	33.5 (15.0)	1003	12.6(5.6)	100
	11th	22.7 (14.0)	1435	12.1(5.2)	99
Science	5th	45.5 (17.6)	878	13.5(5.3)	99
	9th	36.9 (18.4)	962	12.2(5.6)	96
	11th	30.6 (18.0)	997	12.6(6.4)	60
Arabic	5th	40.1 (19.2)	865	13.1(5.2)	98
	9th	39.6 (17.0)	1016	13.2(6.1)	101
	11th	45.8 (17.6)	1419	14.3(5.2)	99

As shown in the above table, students' average NAfKE scores across grades and subjects range between 22.7 and 45.8, which demonstrates that on average, students have low NAfKE performance in the test. However, with subjects or grades, there are variations in NAfKE scores. Most are normally distributed, indicating there is a significant level of reliability in NAfKE results to distinguish students' performance levels within each subject or grade.

As regards teachers' SCALT scores, it has been pointed out earlier, most teachers in Jordan score somewhere in the middle in the SCALT composite (around 64%). However, teachers of different subjects have different SCALT scores indicating that they use variably SCALT methodology in classroom. The study finds that the lowest SCALT scores were reported among Math teachers particularly in 5thth and 11th grade math teachers, (11.7 and 12.1, respectively). The results suggest that SCALT methodology is less likely to be applied by math teachers in Jordan. Are the math teachers less trained, more resistance in applying SCALT methodology, or SCALT is less conducive to Math teaching is not yet known.

As expected, teachers' mean SCALT scores were higher among Arabic teachers, particularly in 11th grade. As mentioned earlier in this report, language and literature might be more conducive to the application of SCALT methodologies than Math and Science. An evidence of that fact might be the observed differences between the lowest and highest scores among Arabic and math teachers in the SCALT composite. On average, the highest mean SCALT score (11th grade Arabic) was 22% higher than the lowest SCALT scores (5th grade math).

b) Correlations between Students' NAfKE and Teachers' SCALT by Grade and Subject

In the correlational analysis, we examined the correlation between students' performance, measured through NAfKE scores) and teachers' utilization of SCALT methodologies in the classroom (measured through the SCALT composite score). Below is the paired correlation.

Table (19): Correlation matrix with students' NAfKE scores and teachers' SCALT Scores

Grade Level	NAfKE Results and Subjects		
	Science-SCALT	**Math-SCALT**	**Arabic-SCALT**
Grade 5	.042	-.004	.030
Grade 9	.108**	.024	.018
Grade 11	.160**	.060*	.182**

Note: ** indicates that p-value is =< 0.01 and * is =<0.05.

The table above shows that in grade 5, there is no statistically significant relationship between SCALT and NAfKE in any of the 3 subjects. However, in grade 11, there is a significant positive relationship between SCALT and NAfKE in every of the 3 subjects. The more SCALT used by teachers, the better NAfKE performance students have. In grade 9, the significance of correlation results vary differently depending on subjects. The findings from the table above that significant correlations between NAfKE and SCALT exist in all subjects in higher grades, but not in lower grades imply that student competencies in critical thinking, problem solving and information decoding skills (measured by NAfKE) can be better improved by application of SCALT methodologies, particularly in higher grades.

Although the magnitude of the correlation findings is relatively small and varied by subject, consistent findings in all subjects of the 11th grade, also suggest that MOE may be able to prioritize SCALT training for teachers in the higher grades and to conduct further studies to determine why the lower grades have not responded to SCALT as positively. One hypothesis is that SCALT methodology may have varying effects including non-effect on student performance at different levels of their social and cognitive development.

4. Conclusion

The classroom observation baseline study carried out in partnership between NCHRD and World Education under MEP project is intended to support MoE policy makers and educators to understand the current status of teaching practices in basic and secondary schools in Jordan. Under ERfKE II program, curriculum standards-based training and interventions have been and will continuously be

implemented to improve teachers' teaching practices in classrooms. This current study has provided a solid baseline and offers many valuable findings regarding teaching practices in Jordanian school classrooms. We are confident that a follow up study will bring about a full realization of measuring changes in teachers' classroom practices.

A random selection of a large sample size, use of scientific methods, development of a reliable classroom observation instrument, training of 65 experienced MOE observers, empirical data analysis strategies, and collaborative team work under the MEP project has made this classroom observation baseline study unique and credible.

Specifically, this baseline study has provided useful information in the following areas:

1) The current profile of teaching practices in MOE schools in Jordan;

2) The extent to which teachers performed well or poorly on the use of student-centered active learning and teaching and other teaching methodologies in the classroom;

3) The variations in the key measures of teaching practices and behavior in class between and within directorates, and other comparison groups such as female and male teachers, urban and rural school, and schools under various authorities and project interventions;

4) Which factors (limiting to teacher characteristics) significantly explain the variation in the measure of student-centered active learning and teaching.

In addition, we integrated the classroom observation data with NAfKE student assessment data to answer the question if the use of student-centered active learning and teaching methodologies correlates with student achievement.

The results of this baseline study suggest that Jordanian teachers have much room for improvement in utilizing student-centered active learning and teaching (SCALT) methodologies, integrating technology into classroom teaching and learning, motivating students in learning, and providing in-classroom student assessment feedback to students. This study also shows that a significant number of excellent teachers in the measured domain areas can be found in many directorates, urban and rural areas, and in all grades. Learning from the high performance teachers is critical to replicate this success through the Jordanian education system. In addition, female teachers' consistent outperformance over their male counterparts and ERSP school teachers' outperformance over other teachers in almost all measured areas in this study deserve a further and thorough study to summarize the positive experience and successful factors.

Some policy implications may result from these findings: 1) Teachers of higher grades above 10th grade should be encouraged to apply SCALT to boost critical thinking, problem solving and communications skills in all subjects; 2) SCALT methods could be applied differently by different subject teachers; 3) SCALT related training programs could be more tailored to subject and grade level needs; 4) there is a need for a further and in-depth study to answer why female teachers consistently better than their counterparts and why ERSP teachers are significantly better than other school teachers in applying the SCALT methodology.

As we have completed the baseline study, we are fully aware of the remaining unanswered inquiries. For example, if all teachers use SCALT methodologies well in Jordan, will students perform much better in learning achievement? Can the 21st century skills be effectively taught by teachers using SCALT in Jordan? Why do female teachers perform significantly better in almost all aspects of the outcome measures than their male counterparts even considering other factors equal? Will the post-service training on SCALT methodologies be sufficient to change the culture of classroom teaching? Recognizing these inquiries, we are continuing our research and evaluation journey in MEP project and strengthening our analytical capacity as education analysts to support ERfKE II program and ultimately improve the quality of Jordan education system.

Appendix 1: Descriptive Statistics of Items Used for Composite Measures

Part II: Teaching and Learning Practices (filled during the classroom observation)

Domain	Sub-domain	Items	Yes (%)	No (%)
"Class Management"		1. The class begins on time	94.5	5.5
		2. The class ends on time	93.5	6.5
		3. Teacher tracks number of absentees at the beginning of the class	48.1	51.9
		4. Teacher appears well organized as he/she starts the class.	86.0	14
		5. Teacher calls students by their names. [teacher knows students well]	93.8	6.2
		6. Teacher interacts with students politely or respectfully.	98.2	1.8
		7. Teacher presents class agenda/objectives to the class (orally or in written form).	61.6	38.4
		8. Teacher conducts group discussions in an orderly manner.	73.5	26.5
		9. Teacher successfully stops students who misbehave in class.	76.9	23.1
		10. Most students [at least 90%] appear to be on task	74.9	25.1
"Student-centered active learning and teaching"	Active Instruction	11. Teacher poses questions often during the class [often means at least 5 times].	97.2	2.8
		12. Teacher gives students time to answer questions [waits at least 5 seconds]	92.6	7.4
		13. Teacher uses examples to illustrate a concept or theory.	93.5	6.5
		14. Teacher let students practice skills taught in this class.	82.2	17.8
	Problem Solving and Inquiry	15. Teacher asks open-ended questions so students can contribute to the discussions	70.6	29.4
		16. Teacher encourages the students to ask questions or express their opinions.	84.9	15.1
		17. Teacher encourages alternative answers or multiple answers to a question.	73.2	26.8
		18. Teacher poses a problem and let students discuss it to bring about a solution.	55.8	44.2
		19. Teacher allows students to interact with each other around content issues	46.9	53.1
		20. Teacher lets students practice scientific	34.9	65.1

		method, drawing data evidence for a conclusive statement.		
	Cooperative learning	21. Teacher divides the students into groups for discussions or problem solving.	32.8	67.2
		22. Teacher assigns group leader and other role-playing responsibilities for each group.	16.4	83.6
		23. Teacher lets student groups elect leader and role-playing responsibilities in each group	19.2	80.8
		24. Teacher discusses learning progress with each group.	31.5	68.5
		25. Teacher walks between groups during the group work.	30.4	69.6
		26. Teacher lets student groups (either representative or whole group) present their group work results	30.6	69.4
		27. Teacher only criticizes wrong or insufficient answers for each presentation.	30.1	69.9
	Learning by Activity	28. The teacher uses debate technique	11.7	88.3
		29. Teacher lets students present different opinions, ideas or problem solutions.	66.1	33.9
		30. Teacher lets students play games for learning purpose	11.6	88.4
		31. Teacher asks the students to prepare projects that require organizing learning materials	14.6	85.4
	Critical Thinking	32. Teacher raises issues and ask for different opinions from students.	53.5	46.5
		33. Teacher gives students chance to express their opinions about the raised issues.	60.5	39.5
		34. Teacher acknowledges students' opinions about the raised issues with positive notation.	64.7	35.3
		35. Teacher summarizes different points of views for a given issue discussed in classroom.	52.4	47.6
		36. Teacher gives students time to assess the raised opinions.	38.1	61.9
"Use Instructional Tools & Media Aids"	Using Computer in class	37. Teacher uses computer in the class during the teaching period.	8.2	91.8
		38. Teacher has a projected screen to show teaching and learning materials.	7.8	92.2
		39. Teacher uses computer to assess students learning.	2.2	97.8
		40. Teacher uses e-content from EduWave or other on-line sources in teaching.	3.4	96.6
		41. Students use computers in classroom for learning purpose.	4.8	95.2

		42. Teacher utilizes PowerPoint or Excel presentations for teaching	5.4	94.6
		43. Teacher uses laptop for teaching purposes	1.9	98.1
	Other Educational Aid	44. Teacher uses non-computerized teaching aid (maps, anatomical models, etc)for teaching the class.	36.0	64
		45. Teacher distributes sheets and educational materials: reports, working-sheets…	29.8	70.2
		46. Teacher answers the students' questions about the educational aid.	34.6	65.4
"Motivation"		47. Teacher praises students' responses, participation in activities or expression of opinions.	83.3	16.7
		48. Teacher sends out praises for the whole class for its overall learning performance.	57.3	42.7
		49. Teacher mocks the whole class when their performance is not satisfactory.	2.4	97.6
		50. Teacher mocks student who makes a mistake or misbehaves.	4.7	95.3
		51. The teacher reviews the previous learning of the students.	87.8	12.2
		52. The teacher links the new learning with the previous one.	87.7	12.3
"Students Assessment"		53. The teacher presents continually feedback.	84.3	15.7
		54. Teacher gives the students a "check list" for self assessment	8.5	91.5
		55. Teacher asks the students to assess their peers on a piece of specific work.	19.5	80.5
		56. Teacher uses a "check list" to assess students learning	14.0	86
		57. Teacher uses a "rating scales" to assess students learning	17.6	82.4
		58. Teacher discuses the homework assigned previously.	36.0	64
		59. Teacher assigns new homework and tells students to hand it in the next day	59.2	40.8

Part III: Summative Evaluation (filled out soon after the classroom observation)

Item	Failure	Need Improvement	Satisfactory	Outstanding
	1	2	3	4
I. Class Management				
1. Teacher manages the time effectively during the class	5.0	12.2	42.4	40.5
2. Teacher manages students' misbehaviors effectively	2.6	8.7	39.1	49.6
3. Teacher and students have respectful interactions.	1.3	6.1	28.4	64.2
4. Class is conducted in a well organized way.	2.0	12.6	45.9	39.5
5. Class is interactive class.	3.7	15.3	39.6	41.5
II. Instructional Strategies				
6. Teacher uses diverse instructional strategies.	7.3	46.6	39.3	6.7
7. Teacher tends to teach using the indirect strategies	21.7	43.0	30.0	5.3
8. Teacher gives most of the class' time to his students during learning.	2.8	28.6	52.0	16.6
9. This class is a good model of student-centered teaching	6.8	33.4	50.2	9.5
III. Tools & Aids				
10. Teacher uses all available tools and aids for teaching.	17.6	43.7	29.6	9.1
11. Teacher uses these tools and aids proficiently.	14.2	31.6	30.5	23.8
IV. Motivation				
12. The teacher is keen to raise learners motivation	6.1	19.3	48.2	26.3
13. Teacher appears enthusiastic in teaching the class.	2.7	13.1	47.2	37.0
14. Students appear enthusiastic in learning .	2.8	23.6	52.8	20.7
V. Students Assessment				
15. Teacher uses various assessments in the class.	11.8	41.6	24.8	21.8
16. Teacher uses assessment results for addressing challenges.	14.7	30.6	45.3	9.4