CRDP
Center for
Educational Research
& Development

IMPLEMENTATION OF FORMATIVE ASSESSMENT

Center for Educational Research and Development-2021

Contents

Introduction

In Lebanon, there has been great progress in reaching increased enrollment in schools, especially in primary education. It is worth noting that, in the academic year 2015-2016, enrollment in public education increased 11 000 students than the previous year. The Center of Education for Research and Development in Lebanon (CERD) and the General Directorate of Education are assiduous in delivering quality education services and learning environments throughout the continuum of formal or non-formal schooling pathways to ensure meaningful and grade-appropriate learning for children and youth. In this sense, interventions are continuously designed to meet this goal. (United Nations, 2017)

The S2R2 program will help expand access to schools for all children in Lebanon, with an increased focus on improving the quality and inclusiveness of the teaching and learning environment, and on strengthening the national education system, its policies, planning, and monitoring capacities. S2R2 strongly focuses on the quality of education. MEHE is committed to improving the curriculum, as well as the quality of teaching, the learning environment, and learning materials. MEHE's initiatives will include formative assessments at schools; student support program; the provision of psychosocial and academic counselors to give teachers and schools guidance on performance (World bank, 2016). CERD, under the tutelage of the Minister of Education, is responsible for informing and supporting improvements in education quality through, among other tasks, conducting educational research and national educational statistics, reforming the national curriculum, strategic planning nationwide, and providing in-service training to education staff.

Mehe is committed to reforms that will develop a modern, effective and coherent school system in Lebanon and CERD is working on a new curriculum framework. A key part of these reforms is introducing a National Student Learning Assessment Framework (NSLAF). The NSLAF sets out a plan to provide a coherent framework of assessment of students in Lebanon (CERD, 2020) one of the main parts of the NSLAF targeted the formative assessment and its important role in education.

The formative assessment framework has been developed by CERD under the 4th component of the S2R2 program, based on this framework, Mathematics, and Arabic digital sample lessons were authored and designed to inform teachers' instructions and support them in scaffolding students according to their performance.

CERD has recommended the implementation of the formative assessment process to test the effect of using formative assessment strategies on the achievement of grade 3 students in mathematics and Arabic.

Implementation

Sampling methodology

Simplified sampling methodology

In this study, and to optimize the selections of the 180 schools, we used the "quota sampling method", which is a non-probabilistic sampling method, based on logic, common sense, and knowledge of the field. (Py, 2013). Quota sampling is a method of stratified sampling in which the selection within strata is non-random. Selection is normally left to the discretion of the interviewer, and it is this characteristic that destroys any pretensions towards randomness. In our study, and in the vast majority of CRDP's studies and in order to optimize the selection of the schools, we took the following quotas into consideration: governorates and the school foreign language of instruction (French, English and trilingual). The quotas calculus was based on the school population record published by CRDP in their 2020-2021 statistical bulletin (CRDP,2021)

The choice of the 180 schools is represented in the table below:

		Foreign language of instruction			
		French	English	Trilingual	**Total**
Governorates	Beirut	3	3	1	**7**
	Mont-Lebanon (with Beirut Suburbs)	5	7	0	**12**
	Mont-Lebanon (without Beirut Suburbs)	7	12	1	**20**
	North	35	1	0	**36**
	Beqaa	6	11	1	**18**
	South	3	15	7	**25**
	Nabatiyeh	0	7	10	**17**
	Akkar	31	0	0	**31**
	Baalbek-Hermel	9	4	1	**14**
Total		**99**	**60**	**21**	**180**

Table 1: distribution of schools in the selected sample

Process

The process of implementation was divided into 2 phases to control the variables and assure the correct implementation and make sure that all teachers involved are ready for the process and aware of all its steps to ensure a smooth and reliable data collection.

Phase 1:

Training sessions were given to teachers, during which the process of the study was explained.

First training sessions were scheduled before the first phase of implementation. Teachers were divided into groups of 60 teachers (30 schools from which one math teacher and one Arabic teacher were invited).

During these sessions, the research process was explained in detail. Expectations and deadlines were set and discussed with trainees.

All teachers showed interest in being part of this research and participated in the discussions. They shared their expectations, limitations and worries which were all clarified by the trainers. The team ensured that all teachers are present and ready to implement the process as planned.

The presentation included a timeline in which the trainers explained each step of the first phase in addition to the required work from teachers from delivering the assigned objective to performing the first assessment, grading the papers, and uploading the grades.

After training sessions were done, documents that are related to the first phase of the experiment were shared with teachers. These documents include the first assessment document, the evaluation criteria to be applied so that all teachers follow the same grading schema as well as a sheet on which grades must be uploaded. The team kept a continuous contact and follow up so that all teachers will be able to deliver the required documents on time. This follow up continued during the grades submission by schools in order to confirm the validity of the data sent.

Phase 2:

The second phase of the implementation started with training sessions. Teachers were divided into groups of 30 teachers. Each group attended a six-hour-session in which the trainers recalled the steps of the implementation through an interactive activity that focused on embedding the "formative assessment" in its steps. Trainers guided teachers to define the formative assessment through a rich discussion about their experiences. All types of formative assessment were displayed and explained in detail. Teachers were very interested in sharing their point of view and exchanging expertise with fellow teachers.

Teachers were then shown the steps of implementing the "formative assessment" in delivering the assigned objective. A demo lesson took place where the trainers explained the parts of the lesson that will be covered during the coming phase and focused on all aspects of formative assessment embedded in the lesson.

Deadlines and expectations were discussed and finalized as well.

After training sessions, documents related to the second phase were shared with the teachers and the implementation took places in classes.

Teachers submitted their grades for the second phase of the implementation and the team confirmed the data through direct contact with the teachers.

Results comparisons:

In our study, and in order to verify if there's a significant improvement of the student's results before and after the formative assessment implementation, and after verifying the normality of our quantitative variables using the Kolmogorov-smirnov (KS) test of normality, we used the paired samples T-test to compare the means of the different questions in the Arabic and the Math test as well as the total scores of both subject.

We have to note that the Paired sample t-test or the dependent t-test compares the means between two related groups on the same continuous, dependent variable considering that the variables are normally distributed.
In testing the hypotheses of our research, there are two ways to do this work: either use the test statistic or use the p-value. The latter approach (which is also called the observed significance level) is based on a probability called the *p-value*. Assuming the null hypothesis is true; the p-value is the probability of obtaining a sample result that is at least as unlikely as what is observed (Anderson et al., 2003, p. 348). The current study uses the dependent t-test as a statistical tool and the p-value approach to test the research hypotheses at a maximum significance level of 5% ($\alpha = 0.05$)

The main descriptive statistics results of the questions and the total scores are represented in the table below:

Pair	Mean	Std. Deviation	Coefficient of variation
phase1_.Q1_arabic	2.1920	0.93067	0.42
phase2_.Q1_arabic	2.2492	0.85130	0.38
phase1_.Q2_arabic	1.6452	1.02708	0.62
phase2_.Q2_arabic	1.8033	1.03513	0.57
phase1_.Q3_arabic	1.2692	0.73622	0.58
phase2_.Q3_arabic	1.1959	0.72754	0.61
phase1_.Q4_arabic	4.0316	4.08201	1.01
phase2_.Q4_arabic	5.2835	4.44088	0.84
phase1_Total_arabic	**9.1379**	**5.81907**	**0.64**
phase2_Total_arabic	**10.5319**	**6.23781**	**0.59**
phase1_.Q1_math	1.31	0.873	0.67
phase2_.Q1_math	1.50	0.780	0.52
phase1_.Q2_math	1.67	0.701	0.42
phase2_.Q2_math	1.77	0.596	0.34
phase1_.Q3_math	1.29	0.851	0.66
phase2_.Q3_math	1.46	0.800	0.55
phase1_.Q4_math	1.24	0.903	0.73
phase2_.Q4_math	1.06	0.908	0.86
phase1_.Q5_math	1.30	0.843	0.65
phase2_.Q5_math	1.23	0.821	0.66
phase1_.Q6_math	0.90	0.916	1.02
phase2_.Q6_math	1.00	0.858	0.86
phase1_.Q7_math	1.13	0.816	0.72
phase2_.Q7_math	1.17	0.826	0.71
phase1_.Q8_math	0.93	0.849	0.92
phase2_.Q8_math	1.20	0.836	0.70
phase1_.Q9a_math	0.81	0.856	1.06
phase2_.Q9a_math	0.71	0.891	1.25
phase1_.Q9b_math	0.42	0.763	1.81
phase2_.Q9b_math	0.41	0.755	1.82
phase1_Total_math	**10.99**	**5.823**	**0.53**
phase2_Total_math	**11.52**	**5.476**	**0.48**

Table 2: main descriptive statistics

Paired sample T-test results:

Pair comparison	t	df	p-value	Significant mean difference
phase1_.Q1_arabic - phase2_.Q1_arabic	-3.973	2993	0.000	yes
phase1_.Q2_arabic - phase2_.Q2_arabic	-9.891	2993	0.000	yes
phase1_.Q3_arabic - phase2_.Q3_arabic	6.979	2993	0.000	yes
phase1_.Q4_arabic - phase2_.Q4_arabic	-24.482	2993	0.000	yes
phase1_Total_arabic - phase2_Total_arabic	**-23.221**	**2993**	**0.000**	**yes**
phase1_.Q1_math - phase2_.Q1_math	-11.727	2991	0.000	yes
phase1_.Q2_math - phase2_.Q2_math	-7.705	2991	0.000	yes
phase1_.Q3_math - phase2_.Q3_math	-10.547	2991	0.000	yes
phase1_.Q4_math - phase2_.Q4_math	10.141	2991	0.000	yes
phase1_.Q5_math - phase2_.Q5_math	4.191	2991	0.000	yes
phase1_.Q6_math - phase2_.Q6_math	-6.438	2991	0.000	yes
phase1_.Q7_math - phase2_.Q7_math	-2.869	2991	0.004	yes
phase1_.Q8_math - phase2_.Q8_math	-17.566	2991	0.000	yes
phase1_.Q9a_math - phase2_.Q9a_math	5.723	2991	0.000	yes
phase1_.Q9b_math - phase2_.Q9b_math	0.566	2991	0.572	**no**
phase1_Total_math - phase2_Total_math	**-7.343**	**2991**	**0.000**	**yes**

Table 3: Paired samples T-test general results