Self-assessment of dental Students in Wax Carving Course in Syrian Private University

Rana Al-Haj Hussain

Lecturer in operative dentistry, Department of operative dentistry, Faculty of Dentistry, Syrian Private University, Damascus, Syria.

Abstract: *Self-assessment of preclinical formative and summative exercises provides an opportunity for students to become engaged in the learning activity by evaluating their performance. The aim of this study was to show whether dental students were able to critically evaluate their own work in terms of dental anatomy carving test, and whether their capacities for self-assessment could had been improved during the study course. This study included 251 students from the Faculty of Dentistry at Syrian Private University SPU (undergraduates/ first term). All students were subjected to 2 tests, in which they had to carve an upper canine using dental wax block. Three clinical supervisors evaluated the student's quality of work. All students were asked to self-assess their works and record marks for themselves. T-test for paired samples was used to compare marks assigned by supervisors and those recorded by students for each task. Regarding the first task, the mean mark recorded by the supervisors was (6.29±1.49) (Mean± Standard deviation), and that recorded by students was (7.55±1.68). Regarding the second task, the mean mark recorded by the supervisors was (5.69±1.94), and that recorded by students was (7.57±1.93). The mean difference between 1st task supervisors' assessment records and 1st task students' assessment records was significant [P < 0.05]. The mean difference between 2ed task supervisors' assessment records and 2ed task students' assessment records was also significant [P < 0.05]. After splitting the students into 2 categories, it was shown that, regarding good-level students, the mean difference between 1st task supervisors' assessment records and 1st task students' assessment records was significant, and the mean difference between 2nd task supervisors' assessment records with 2nd task students' assessment records was also significant. Regarding poor-level students, the mean difference between 1st task supervisors' assessment records and 1st task students' assessment records was significant, and the mean difference between 2ed task supervisors' assessment records with 2ed task students' assessment records was also significant. This study found that students tend to overestimate the level of their own work during self-assessment session. Making a critical self-assessment by the students to enhance the learning system in the university needs more restricted criteria.*

1. Introduction

Self-assessment (SA) is considered a means to develop critical thinking, self-directed learning, and lifelong learning in dental students [1, 2]. Therefore, having a meaningful student assessment can help lead to the goal of self-directed learning.

When students develop the ability to accurately assess their preclinical performance, it is expected to translate into the long-term ability to self-evaluate and make necessary corrective actions. Therefore, one essential component is the student's ability to recognize good performance and poor performance. In preclinical courses, the faculty assessment is the standard against which students measure their assessment of their performance. One challenge is that novice students cannot reliably perform qualitative evaluations of their performance to the same level as experienced faculty members [1, 3, 4].

Self-assessment of preclinical formative and summative exercises provides an opportunity for students to become engaged in the learning activity by evaluating the strengths and weaknesses of their performance and thus promotes students' taking responsibility for achieving learning objectives.

Asadoorian and Batty [5] defined SA as "an active process of developing an awareness of a personal learning exigency, meaning a pressing need, within one's professional activities to guide the initiation of appropriate learning activities.". This definition supports the theory that training students to accurately self-assess will translate into health care providers' professional learning and implementation of changes required to generate better performance. This is a fundamental goal of preclinical skills development and technique courses, helping the student develop technical skills and modify techniques to achieve desired outcomes. In addition to the preclinical setting, reflective learning through SA has been incorporated at various curricular levels, including the clinical setting, with the goal of creating oral health care providers with the skills to self-appraise [6, 7].

Dental anatomy carving is one of the subjects which is taught to first-year students at the Faculty of

Dentistry/ Syrian Private University (SPU) in Damascus, Syria, that is the entrance to Dentistry
in which students recognize preliminary information of teeth anatomy to learn the structure and internal shape of every tooth. In our opinion, this would enhance the competence and even the practice of professional work later at the dental office.

Through the Faculty's continuous trials to enhance the educational methods in SPU, the aim of the current study was to show whether dental students (first class undergraduates) at the Faculty of Dentistry in SPU were able to critically evaluate their own work, and whether their capacities for self-assessment could had been improved during the study course.

2. Materials and Methods

This study was designed to include 251 registered students from the Faculty of Dentistry at SPU who were at the first class (undergraduates/ first term) attending the dental anatomy carving course.

All students were subjected to 2 tests, or tasks, in which they were asked to carve an upper canine using the dental wax block. The exact appointments of those 2 tests were known by all students. There were previous pre-clinical lectures given to the students before those 2 tests, in which they learned the basics of anatomical dental carving. The time given to the students for every task was 40 minutes, and they had to follow the anatomical standards of teeth being carved (i.e. all anatomical features of mesial, distal, buccal, and lingual surfaces of the canines). The anatomical standards for carving were obtained from those mentioned by Wheeler (1974) (8) which were already explained to students during their previous lectures. All students were allowed, for the whole semester, to keep an educational DVD at home in which a full description of anatomical dental carving was explained in details.

After finishing the tasks by students, a committee of 3 clinical supervisors, who used to teach dental anatomy, had to evaluate the student's quality of work using a continuous marks scale from 0 to 10 that had to be recorded on a separate student's sheet.

The evaluation achieved by the supervisors depended basically on how much the student had followed the anatomical standards in his/her carving of the canines.

After each test, all students were asked to self-assess their works and record marks for themselves using the same mark scale used by the supervisors (0-10).

A period of time, namely 2 weeks, was assigned between the first test and the second test.

T-test for paired samples was used to compare marks assigned by supervisors and those recorded and self-assessed by students with a confidence level of 95%. This was done for each task.

3. Results

Regarding the first task, the supervisors assessed the work of 212 students, and the marks ranged from 2 to 9 with a mean record of (6.29±1.49) (Mean± Standard deviation) (Figure 1).

Out of the whole number of students assessed by supervisors, 8 students refused to self-assess their own works later on, and were, therefore, dropped from the evaluation. The rest of students (i.e. 204 students) recorded their marks as per self-assessed, and they ranged from 0 to 10 with a mean record of (7.55±1.68) (Figure 2).

Regarding the second task, the supervisors assessed the work of 223 students, and the marks ranged from 0 to 10 with a mean record of (5.69±1.94) (Mean± Standard deviation) (Figure 3).

Out of the whole number of students assessed by supervisors in the second task, another 8 students refused to self-assess their own works and were, therefore, dropped from the evaluation. The rest of students (i.e. 215 students) recorded their marks as per self-assessed, and they ranged from 0 to 10 with a mean record of (7.57±1.93) (Figure 4).

T-test for paired samples was used to find any significant statistical differences between the 1st task supervisors' assessment records and the 1st task students' assessment records. The same test was also used to compare the 2ed task supervisors' assessment records with 2ed task students' assessment records, 1st task supervisors' assessment records and 2ed task supervisors' assessment records, and also to compare 1st task students' assessment records with 2ed task students' assessment records. However, and due to some technical drops in the records' list, only 184 records were assigned to the statistical test (Table 1).

The results of the statistical test showed that the mean difference between 1st task supervisors' assessment records and 1st task students' assessment records was significant [t=11.89, Sig. (2-tailed)= 0.000 that is < 0.05]. The mean difference between 2ed task supervisors' assessment records and 2ed task students' assessment records was also significant [t=15.01, Sig. (2-tailed)= 0.000 that is < 0.05]. The mean difference between 1st task students' assessment records and 2ed task students' assessment records was not significant [t= -0.57, Sig. (2-tailed)= 0.56 that is > 0.05]. It was also shown that the mean difference between 1st task supervisors' assessment records and 2ed task supervisors' assessment records was significant [t= 3.91, Sig. (2-tailed)= 0.000 that is < 0.05] (Table 2).

In order to find out more critical differences between groups, the students records were split into 2 categories; namely, records of students who got 5 or more marks by the supervisors, and were considered as good-level students, and records of students who got less than 5 marks by the supervisors, and those were, therefore, considered as poor-level students.

Evaluation results of the good-level students

For the first task, the supervisors assessed the work of 146 students, and the marks ranged from 6 to 9 with a mean record of (7.08±1.00)(Mean± Standard deviation) (Figure 5).

Out of the whole number of students assessed by supervisors, 4 students refused to self-assess their own works and were, therefore, dropped from the evaluation. The rest of students (i.e. 142 students) recorded their marks as per self-assessed and they ranged from 0 to 10 with a mean record of (8.01±1.50) (Figure 6).

Regarding the second task, the supervisors assessed the work of 132 students, and the marks ranged from 6 to 10 with a mean record of (7.02±0.93)(Mean± Standard deviation) (Figure 7).

Out of the whole number of students assessed by supervisors in the second task, 6 students refused to self-assess their own works and were, therefore, dropped from the study. The rest of students (i.e. 126 students) recorded their marks as per self-assessed and they ranged from 1 to 10 with a mean record of (8.34±1.36) (Figure 8).

T-test for paired samples was used to explore any statistical differences between 1st task supervisors' assessment records and 1st task students' assessment records. The same test was also used to compare 2ed task supervisors' assessment records with 2ed task students' assessment records (Table 3).

The results of the statistical test showed that the mean difference between 1st task supervisors' assessment records and 1st task students' assessment records was significant [t=7.94, Sig. (2-tailed)= 0.000 that is < 0.05]. The mean difference between 2nd task supervisors' assessment records with 2nd task students' assessment records was also significant [t=9.95, Sig. (2-tailed)= 0.000 that is < 0.05].

Evaluation results of poor-level students

For the first task, the supervisors assessed the work of 66 students, and the marks ranged from 2 to 5 with a mean record of (4.55±0.72) (Mean± Standard deviation) (Figure 9).

Out of the whole number of students assessed by supervisors in the first task, 4 students refused to self-assess their own works and were, therefore, dropped from the evaluation. The rest of students (i.e. 62 students) recorded their marks as per self-assessed and they ranged from 1 to 9 with a mean record of (6.5±1.6) (Mean± Standard deviation) (Figure 10).

Regarding the second task, the supervisors assessed the work of 91 students, and the marks ranged from 0 to 5 with a mean record of (3.76±1.29) (Mean± Standard deviation) (Figure 11).

Out of the whole number of students assessed by supervisors in the first task, 3 students refused to self-assess their own works and were, therefore, dropped from the evaluation. The rest of students (i.e. 88 students) recorded their marks as per self-assessed and they ranged from 0 to 10 with a mean record of (6.5±2.13) (Figure 12).

T-test for paired samples was used to explore any statistical differences between 1st task supervisors' assessment records and 1st task students' assessment records. The same test was also used to compare 2ed task supervisors' assessment records with 2ed task students' assessment records (Table 5).

The results of the statistical test showed that the mean difference between 1st task supervisors' assessment records and 1st task students' assessment records was significant [t=9.34, Sig. (2-tailed)= 0.000 that is < 0.05]. The mean difference between 2ed task supervisors' assessment records with 2ed task students' assessment records was also significant [t=13.5, Sig. (2-tailed)= 0.000 that is < 0.05].

4. Discussion

The current study addressed the relationship between the marks of self-assessment recorded by students and those recorded by faculty supervisors in a dental anatomy course including wax curving of a tooth sculpture.

In general, all students in this study were somehow inaccurate about assessing their own works due to the significant differences between their own–recorded marks and the supervisors' marks. However, the so-called poor-level students were the furthest away, in their self-assessment, from the supervisors.

Due to some technical drops in the records' list, only 184 records were assigned to the statistical test to compare the first and second task between students and supervisors. This occurred because some students had records for first task but not for the second one and vice versa. Another reason might be that some students were assessed by supervisors for the 1st task and/or for the 2nd one while other students were not. However, this had no negative impact on the results because only available data were statistically assessed at the end.

Our findings are consistent with studies that have suggested that student self-assessment, although a potentially valuable pedagogical tool, falls short due to students' reluctance or inability to accurately self-assess (9, 10).

It should be mentioned that the first carving test was not the first carving task given to the students. It was predisposed by 2 different tasks during the course. Thus, it was not a kind of surprise for the students to self-assess their works without previous knowledge.

Regardless of the students' self-assessments, and the repetition of the assigned tasks, the supervisors' assessments were almost constant during the study throughout both tests. This reflects a high reliability of the supervisors in this study.

The conviction of students that they fulfilled the requested criteria in every test may explain why they overestimated themselves with the related scores out of 10. The students, however, usually tend to feel that they deserve more than they really deserve in reality. Thus, the unchanged self-assessments' records of students (i.e. as overestimated) could be explained by the students' feelings that their levels were always steady.

The decrease of mean records related to the supervisors between the first task (6.32) and the second one (5.72) with a notable significance (P=0.000) reflects the reliability of recording the students' levels without any bias or preferences. This gave an extra strength to the study.

Our trial to split the students into 2 categories; namely good-level and poor-level students, and processing their data, did not change the result that they kept overestimating themselves in a significant way (P=0.000).

Regarding the poor-level students, it was obvious that they also overestimated their own works. Moreover, the mean differences between the poor-level students' assessments and the supervisors' assessments for both tests (1.87 and 2.68 for the first and second test respectively) were higher than those noted between the good-level students' assessments and the supervisors' assessments (0.93 and 1.31 for the first and second test respectively). This might be explained by the false conviction of the poor-level students that giving a close shape of the carved tooth was enough to reach the requirement of the task. They could have thought also that the action of repeating a carving task could have been enough to get improved. We believe that those category of students did not get the real target of the carving task, which was giving a carved tooth sculpture that is to be accurate in all scales and measurements as per known according to Wheeler (8).

This study is consistent with the literature on students' self-assessment of their laboratory projects in terms of confirming modest faculty consistency with each other and weak consistency between student and faculty assessment of projects. It was confirmed that students gave themselves higher scores than faculty members did and those students with the lowest marks both overestimated their scores more and used a wider range of scores (1).

Making a definite assessment of the students' abilities to produce a remarkable work in the course of dental carving might not be that precise through only 2 tasks. Therefore, we recommend many tasks in the future for better assessment (e.g. 4-5 tasks). The continuity of assessing the students all over the semester would also provoke the students to improve their works.

We believe that the students should have seen their faults and errors after the first task, so that they could have fixed the gaps they did, and got themselves more ready for the next one. This would enhance the quality of work.

It is our conviction that good carving is closely related to having an innate talent, which affect the quality of work by students. Thus, from our point of view, it is highly recommended to include a tooth sculpture carving test as a prerequisite for admission in every dental school, including SPU.

5. Conclusion

This study found that students, in general, tend to overestimate the level of their own work (i.e. carving a dental sculpture in terms of a test) during self-assessment session even if they are classified into good-level and poor-level students. However, poor-level students were the furthest away from the supervisors' assessments than good-level students.

Making a critical self-assessment by the students to enhance the learning system in the university needs more tasks to achieve in terms of frequent tests, in which they find out their faults and errors for further improvement.

We suggest a "tooth sculpture carving test" as a prerequisite for admission in every dental school.

6. References

[1]. Chambers DW, LaBarre EE. The Effects of Student Self-Assessment on Learning in Removable Prosthodontics Laboratory. Journal of dental education. 2014;78(5):668-80.

[2]. DeVore L, Fried J, Dailey J, Qori C. Dental hygiene self-assessment: a key to quality care. Journal of dental hygiene: JDH/American Dental Hygienists' Association. 1999;74(4):271-9.

[3]. Eva KW, Regehr G. Self-assessment in the health professions: a reformulation and research agenda. Academic Medicine. 2005;80(10):S46-S54.

[4]. Guenzel PJ, Knight GW. Long-term outcomes for remedial students. The Journal of prosthetic dentistry. 1992;67(6):865-9.

[5]. Asadoorian J, Batty HP. An evidence-based model of effective self-assessment for directing professional learning. Journal of dental education. 2005;69(12):1315-23.

[6]. Cho GC, Chee WW, Tan DT. Dental students' ability to evaluate themselves in fixed prosthodontics. Journal of dental education. 2010;74(11):1237-42.

[7]. Navickis MA, Bray KK, Overman PR, Emmons M, Hessel RF, Cowman SE. Examining clinical assessment practices in US dental hygiene

programs. Journal of dental education. 2010;74(3):297-310.

[8]. Wheeler RC. Dental anatomy, physiology, and occlusion: WB Saunders Company; 1974.

[9]. Chambers DW, LaBarre EE. Why professional judgment is better than objective description in dental faculty evaluations of student performance. Journal of dental education. 2014;78(5):681-93.

[10]. Davis DA, Mazmanian PE, Fordis M, Van Harrison R, Thorpe KE, Perrier L. Accuracy of physician self-assessment compared with observed measures of competence: a systematic review. Jama. 2006;296(9):1094-102.

Table 1. Paired samples assigned to the statistical analysis

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	1st Task Student's Assessment	7.557	184	1.5805	.1165
	1st Task Supervisor's Assessment	6.329	184	1.4605	.1077
Pair 2	2ed Task Student's Assessment	7.639	184	1.9280	.1421
	2ed Task Supervisor's Assessment	5.728	184	1.9360	.1427
Pair 3	1st Task Student's Assessment	7.557	184	1.5805	.1165
	2ed Task Student's Assessment	7.639	184	1.9280	.1421
Pair 4	1st Task Supervisor's Assessment	6.329	184	1.4605	.1077
	2ed Task Supervisor's Assessment	5.728	184	1.9360	.1427

Table (2): Results of T-test for comparing students' tasks with supervisors' tasks

	Paired Differences					t	df	Sig. (2-tailed)
	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
				Lower	Upper			
Pair 1 — 1st Task Student's Assessment - 1st Task Supervisor's Assessment	1.2283	1.4005	.1032	1.0246	1.4320	11.897	183	.000
Pair 2 — 2ed Task Student's Assessment - 2ed Task Supervisor's Assessment	1.9103	1.7254	.1272	1.6594	2.1613	15.019	183	.000
Pair 3 — 1st Task Student's Assessment - 2ed Task Student's Assessment	-.0815	1.9167	.1413	-.3603	.1973	-.577	183	.565
Pair 4 — 1st Task Supervisor's Assessment - 2ed Task Supervisor's Assessment	.6005	2.0796	.1533	.2981	.9030	3.917	183	.000

Table (3): Paired samples assigned to the statistical analysis

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	1st Task Student's Assessment	8.011	142	1.5059	.1264
	1st Task Supervisor's Assessment	7.077	142	.9970	.0837

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	2ed Task Student's Assessment	8.337	126	1.3665	.1217
	2ed Task Supervisor's Assessment	7.024	126	.9484	.0845

Table (4): Results of T-test for comparing students' tasks with supervisors' tasks

	Paired Differences					t	df	Sig. (2-tailed)
	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
				Lower	Upper			
Pair 1	1st Task Student's Assessment - 1st Task Supervisor's Assessment							
	.9331	1.3994	.1174	.7009	1.1653	7.946	141	.000

	Paired Differences					t	df	Sig. (2-tailed)
	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
				Lower	Upper			
Pair 2	2ed Task Student's Assessment - 2ed Task Supervisor's Assessment							
	1.3135	1.4815	.1320	1.0523	1.5747	9.952	125	.000

Table (5): Paired samples assigned to the statistical analysis

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	1st Task Student's Assessment	6.500	62	1.6094	.2044
	1st Task Supervisor's Assessment	4.621	62	.6317	.0802

		Mean	N	Std. Deviation	Std. Error Mean
Pair 2	2ed Task Student's Assessment	6.500	88	2.1335	.2274
	2ed Task Supervisor's Assessment	3.818	88	1.2368	.1318

Table (6): Results of T-test for comparing poor-level students' tasks with supervisors' tasks

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	1st Task Student's Assessment - 1st Task Supervisor's Assessment	1.8790	1.5829	.2010	1.4770	2.2810	9.347	61	.000

Table (7): Results of T-test for comparing poor-level students' tasks with supervisors' tasks

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	2ed Task Student's Assessment - 2ed Task Supervisor's Assessment	2.6818	1.8634	.1986	2.2870	3.0766	13.501	87	.000

Figure (3): Records of the 2ed task of carving according to the supervisors' assessments
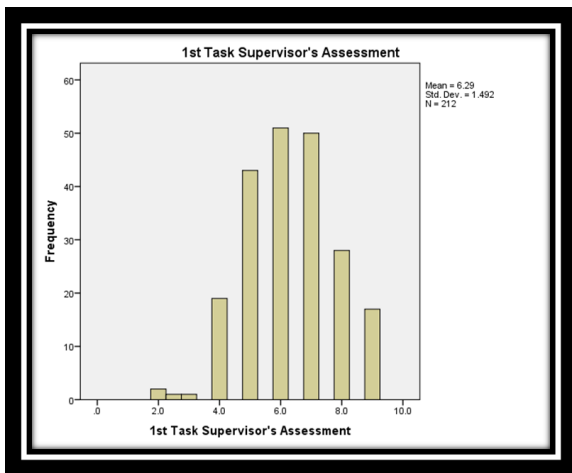
Figure (1): Records of the 1st task of carving according to the supervisors' assessments
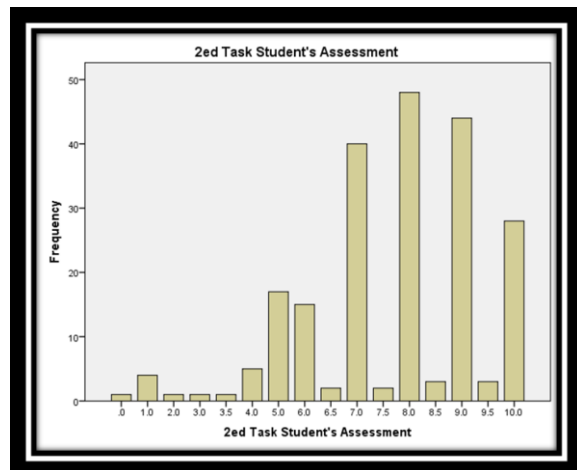
Figure (4): Records of the 2ed task of carving according to the students' assessments

Figure (2): Records of the 1st task of carving according to the students' assessments
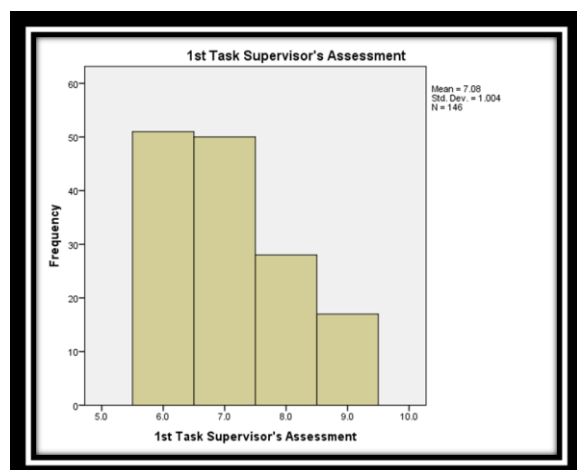
Figure (5): Records of the 1st task of carving according to the supervisors' assessments for good-level students
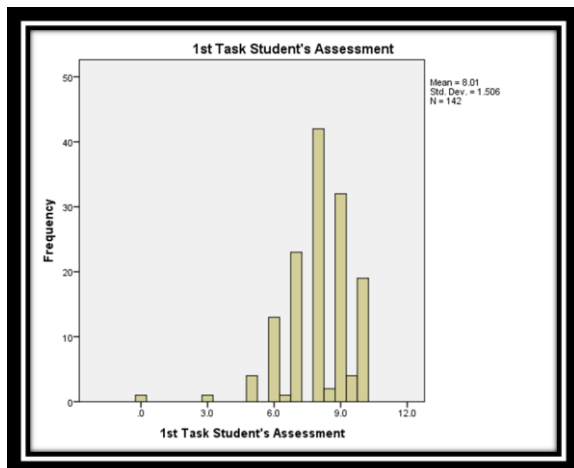
Figure (6): Records of the 1ˢᵗ task of carving according to the good-level students' assessments
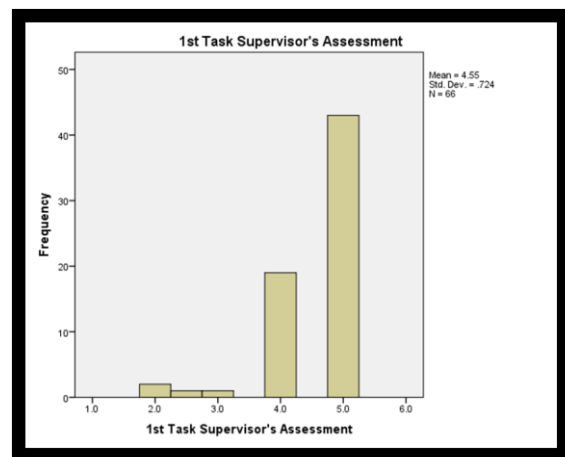

Figure (7): Records of the 2ed task of carving according to the supervisors' assessments for good-level students

Figure (8): Records of the 2ed task of carving according to the good-level students' assessments

Figure (9): Records of the 1st task of carving according to the supervisors' assessments for poor-level students
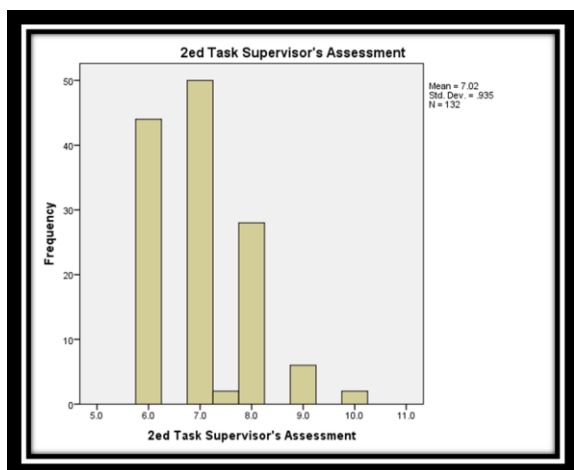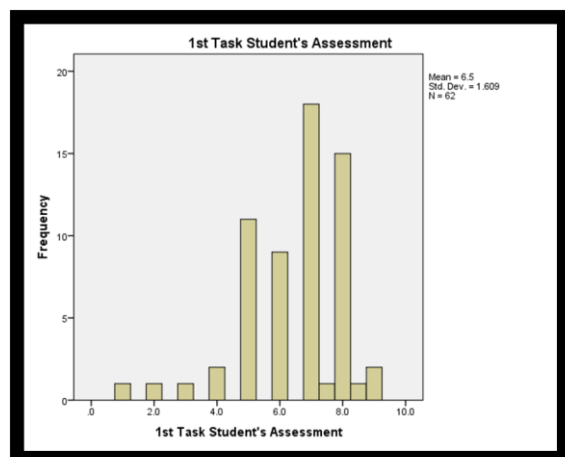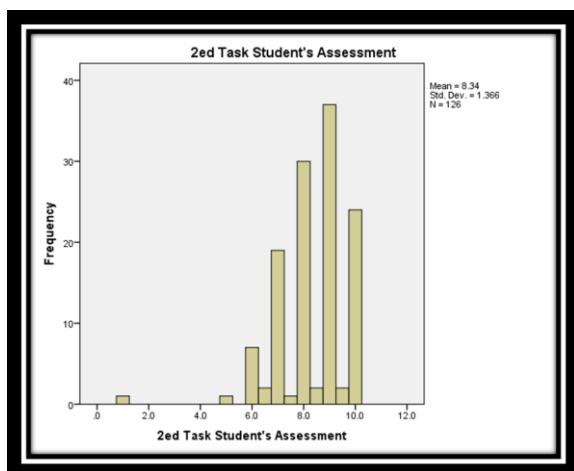

Figure (10): Records of the 1st task of carving according to the poor-level students' assessments
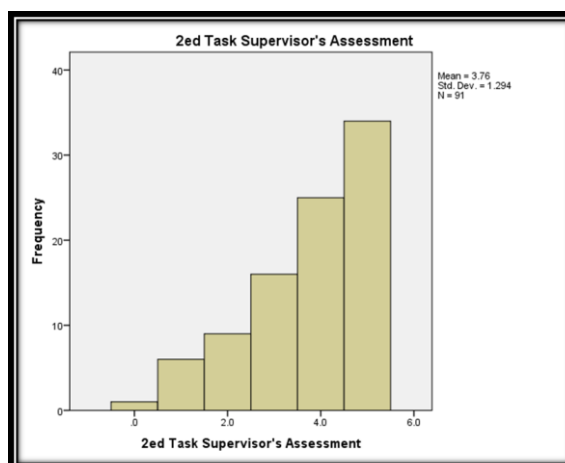

Figure (11): Records of the 2st task of carving according to the supervisors' assessments for poor-level students
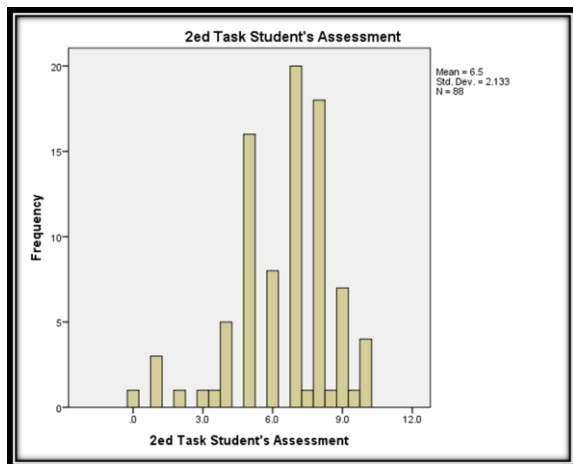
Figure (12): Records of the 2ed task of carving according to the poor-level students' assessments