

Performance Assessment of the Calculus Students by Using Scoring Rubrics in Composition and Inverse Function

Nawal H. Shirawia¹, Aida B. Qasimi¹, Mohammad A. Tashtoush^{1,2,*}, Noha M. Rasheed², Mohamad A. S. Khasawneh³ and Emad A. Az-Zo'bi⁴

¹Department of Education, Faculty of Education and Arts, Sohar University, Sohar, Oman

²Department of Basic Sciences, AL-Huson University College, AL-Balqa Applied University, AL-Salt, Jordan

³Special Education Department, Faculty of Education, King Khalid University, Abha, Saudi Arabia

⁴Department of Mathematics and Statistics, Mutah University, Al Karak 61710, Jordan

Received: 24 Aug. 2023, Revised: 16 Dec. 2023, Accepted: 9 Jul. 2024

Published online: 1 Sep. 2024

Abstract: Performance assessment is a valuable method of evaluating students' understanding and abilities in educational settings. It involves measuring what students know and can do by having them demonstrate their skills, knowledge, and competencies through various tasks, projects, or activities. Performance assessments are designed to assess a student's ability to apply their knowledge and skills in real-world contexts, moving beyond simple memorization or regurgitation of facts. This study aimed to assess the performance of Calculus students for some mathematical tasks, the procedures they follow in Composition Functions and finding an inverse function, and to determine the interpretation of these procedures. The study used case study approach; the sample consisted of three students whom were enrolled in the Calculus course during the spring semester 2021/2022. The students were exposed to a quiz in which their performance was assessed by using two scoring rubrics (holistic and analytical), to determinations individual interviews were conducted with them that were analyzed and discussed, behind their failure to complete some parts of solving the quiz and their failure to verify their solution. The results of the study showed that the analytical rubric is the best than of the holistic rubric; the study recommended that students should pay attention to the interpretation of their procedures when performing mathematical tasks, invite teachers to use scoring rubrics to assessment the performance of their students.

Keywords: Assessment Strategies, Educational Psychology, Pedagogical Methods, Performance Assessment, Rubric, Functions.

1 Introduction

The past few years have witnessed significant development in the field of assessment and its growing importance among education experts for educational reform and improvement. Thomas Kalvin, the president of the International Organization for Measurement and Evaluation, addressed the issue of educational reform and development through various assessment processes during the 27th conference of the organization. He emphasized that most past reform efforts focused on educational inputs. At the same time, recent trends have emerged that prioritize educational outcomes and the extent to which students acquire knowledge, skills, behaviors, and attitudes.

According to [1], assessment is gathering evidence about students' mathematical knowledge, their ability to use mathematical knowledge, and their attitudes towards mathematics. It involves extracting judgments from this evidence for various purposes. Assessment has become a primary source of evidence on which teachers base their inferences about what students know or need to learn. Furthermore, student assessment should not be conducted solely for evaluation but also for guiding and supporting student learning and understanding patterns of knowledge.

Researchers and experts in education believe that assessment requires collecting authentic data from various sources and methods which ensure appropriate decisions. However, traditional evaluation depended on

* Corresponding author e-mail: tashtoushzz@su.edu.om

biased assessment tools and approaches [2,3]. Such focuses accumulate information in the mind of the learners. This does not suit the current cognitive view that should be embedded and required in education. Yet, assessment is not compared with determining learners' success or failure, they are gaining to higher scores or educational stages. Instead, it is looked at to enhance the learning process and correct its trajectory [4,5]. Modern assessment differs from transition evaluation where the first focus on the balanced personality of the learners. It evolves focusing on students' deep comprehending of the content, skills of self learning, and interacting with the surrounding [6,7,8].

Performance assessment assesses students' comprehension and abilities focusing not only on students' know but how they can demonstrate their skills, and vailities through various activities tasks, or projects. Scoring rubrics are guiding models that arise certain criteria to detect students' performance in matching to these criteria. They play a significant role in the assessment process in education. [6] pointed out some criteria of rubric. The assessment rubric should be clear and transparent for teachers to assign students' expectations. Rubrics should also help teachers to be consistent in evaluating students' performance without bais. Rubrics should provide students with feedback to develop their skills. They should also provide stuents with self-assessment procedures to evaluate their progress. Finally, scoring rubric whould align with the learning outcomes (LOS) of the course. [9,10].

In the Sultanate of Oman, the Ministry of Education has developed "Mathematics Assessment Document, 2012" as a guidance for teachers mathematics. It determines assessment as judgment process of the value of things, subjects, or individuals. In this sense, it requires criteria or standards to estimate this value. The document categorizes into several types of assessment. 1. Continuous Assessment which contributes to determing strengths and weaknesses in the educational program, for reviewing, developing and modifying, its content. 2. Formative Assessment provides learners and teachers with outcome results to enahce the educational process. It assesses the objectives achievement and utilize feedback on dailty based activities. 3. Summative assessment judges the learners' mastery of LOs which helps in taking decisions for graduation or.promotion. 4. Another type that MAD contains involves students' self assessment where they can apply some evaluating criteria to check their own work. 5. Still peers can participate in the essement of their friend based on some criteria. This type is called peer assessment. 6. MAD lists student portfolio as an essement tool where teachers can assess students' performance over various activities along the curriculum. 7. Modern assessment requires the evaluation of students' performance in light of the learning outcomes. 8. Methodoligical consistency assesment supports students with activities that align with the LOs.

Finally, 9. Standardization requires grading students fairly according to the gradting rubric [11].

Recently, performance-based assesement received poularity as a scoring rubric. It is viewed as schadual cointed by specialist teachers to aid them analyzing the performancce of students over tasks students' performance tasks. Scoring rubric differs traditional approaches in providing authenticity and confidency grading to students tasks [12,13,14].

Scoring rubrics provide performance indicators of the quality of tasks. They also shifted from the basis of traditional assessment approaches to making self-judgments and realistic. Several contemporary scholars set four types of score rubrics [15,16,17] classify scoring rubrics into four main types. Holistic scoring rubrics that provide an overall assessment. They estimate the competence of students as a whole. These rubrics are not for classroom due to their holistic focuses on overall competence. Analytic scoring rubrics divided assessment into dimensions that evaluate separate aspects of performance [18]. Analytic rubrics also provide more information to teachers and students about students' strengths and weaknesses in multiple criteria of performance. Single-trait scoring rubrics involve zero down the features for judging performance in the task to a main dimension. It helps focus on a separate aspect of performance. These rubrics yet require pre-determining the main elements for good task performance. Finally, the multi-trait scoring rubrics provide assessing students' performance over multi dimensions [19,20].

Both analytic scoring and holistic rubrics are capable of gathering data about students' performance and cognitive skills. However, holistic scoring rubrics provide separate grading of students' performance [21,22]. These reburies are compatible with standardized tests. On the other hand, analytic scoring rubrics judge provide an independent grading; they offering a scoring for each category and the dimension of students' performance [23,24,25].

Based on the foundation that traditional assessment methods in most educational institutions have proven ineffective in measuring students' skills and knowledge, it became necessary to develop and adopt modern assessment methods, diversify student assessment approaches, and focus on performance based on performance criteria. These approaches aim to simultaneously measure learning outcomes and processes [15,20,26]. Therefore, this study considers performance assessment rubrics indicators of performance quality in specific tasks. They can gather information about students' task performance, improve their cognitive skills, and make judgments on student performance more realistic. Many education experts in this field recognize the use of performance assessment rubrics as providing convincing justifications for the feasibility and effectiveness of this type of assessment.

Problem Statement

Scoring rubrics are educational concepts and a type of assessment based on performance that plays a significant role in evaluating the process of teaching and learning mathematics. This type of assessment provides teachers with information about students' understanding of knowledge and skills and their ability to apply them in learning mathematics. It also enables teachers to integrate classroom teaching with performance-based assessment and scoring rubrics, enriching their educational and pedagogical experience.

Since associations are the fundamental building blocks of studying calculus, this study aims to provide a background on some scoring rubrics that fall under performance-based assessment. These rules can benefit teachers and university instructors, as they are the critical element in the assessment process in the classroom. Two scoring rubrics were adopted for assessing mathematical tasks: the first is the holistic scoring rubric, and the second is the analytical scoring rubric. Each of them has its aspects. The holistic scoring rubric provides a basis for comparing two performances, but does not rely on task analysis or provide diagnostic information about students' task performance. On the other hand, the analytical scoring rubrics provide more detailed grading, and its results are described as more accurate, although it may focus more on certain aspects of performance than others.

Due to significant deficiencies in the assessment methods and practices employed in our educational institutions, and based on field observations by researchers who are mathematics teachers, as well as the second researcher's participation in international assessment tests, and the first researcher's supervision of pre-service mathematics teachers' training in the educational field, and considering their recognition of the importance of modern assessment methods, this study is aligned with global and Arab movements and local initiatives in the Sultanate of Oman. It aims to develop assessment methods based on students' performance. The idea of this study emerged from the researchers' firsthand experience and struggles in trying to understand students' thinking processes when they engage in tasks involving different levels of operations on functions, particularly the composition of functions and finding inverse functions. They also seek to explain the low level of performance in these tasks. Specifically, this study aims to answer the following two questions:

1. *Do the two scoring rubrics provide consistent interpretations of scores at the same level?*
2. *Does students' knowledge of the details of the scoring rubrics contribute to improving their performance?*

Study Importance

The importance of the study stems from the significance of the topic it addresses. The theoretical importance lies in the fact that it is one of the studies that call for the examination of associations in general, and specifically the operations of composing associations and

finding inverse associations. It involves using methods and techniques to assess students' performance in Calculus, based on alternative evaluation that requires searching for multiple sources of evidence, building conclusions, and judging what students know to achieve realistic assessment. Researchers hope that this study will enrich the theoretical and research literature in mathematics and fill a gap in this area of research. The practical importance of this study lies in its potential to benefit students, faculty members, and specialists in developing the assessment process. It enables them to become familiar with performance-based assessment tools and the use of scoring rubrics.

Procedural Definitions

- Performance-based Assessment:** It is a type of assessment designed to measure a student's ability to perform specific tasks and judge their achievement using assessment tools that estimate their level.
- Scoring Rubrics:** A method of evaluating a student's performance based on meaningful criteria and judging their performance level in a single task or multiple tasks.
- Holistic Scoring Rubrics:** A scale that provides a general overview and estimation of a student's performance in a specific task.
- Analytic Scoring Rubrics:** A scale that categorizes a student's performance in a specific task into multiple levels, where each level is measured separately, and then an overall judgment is made based on all levels.

Limitations

- Time Boundaries:** This study was conducted during the second semester of the academic year 2021/2022.
- Spatial Boundaries:** The study was conducted at Sohar University in Oman.
- Human Boundaries:** The study was conducted on students enrolled in the Calculus course.
- Subject Boundaries:** This study addressed the overall and analytical scoring rubrics, their results using data collection tools, procedures, the nature of the community and the sample, and the operations of composing associations and finding inverse associations.
- The study is defined by its psychometric properties, including acceptable validity and reliability, for the purposes of scientific research to achieve the study's objectives.

Literature Review

Through reviewing the theoretical and educational literature that addressed the importance of using scoring rubrics in teaching and learning mathematics, this study discussed some previous studies related to the subject of the current study, which could be useful to mention in the current study. [27] Conducted a study to investigate the effectiveness of using analytical scoring rubrics to assess student's performance in geometric proof by employing

cognitive behavior theories of geometric knowledge. A group of evaluators assessed the performance of 241 students in geometric proof using a scoring rubric that included multiple criteria through various tests measuring students' ability in geometric proof. The study results showed that analytical scoring rubrics yielded better results than traditional assessment methods.

Similarly, [28] examined the impact of broad-spectrum assessment on teaching methods and instructional strategies followed by teachers in Pennsylvania. The study identified factors influencing teachers' beliefs and usage of broad-spectrum assessment and used a questionnaire administered to 168 teachers from 20 different schools. The study revealed that while teachers recognized the importance and value of the scoring rubrics in the assessment guide, they did not use them. They attributed this to developing their own scoring rubrics. Furthermore, teachers adhered to traditional methods and did not adopt advanced scoring rubrics in assessment.

[29] Conducted a study aiming to investigate the effect of using scoring rubrics to assess performance in achievement and attitudes of tenth-grade students in mathematics in Jordan. Two scoring rubrics, holistic and analytical, were developed, and an achievement test and a questionnaire to measure students' attitudes were administered to 128 students divided into three groups: the first experimental group was assessed based on an analytical scoring rubric, the second experimental group was assessed based on a holistic scoring rubric, and the control group was assessed using traditional methods. The study results showed the effectiveness of holistic and analytical assessment methods on students' performance. The study recommended incorporating scoring rubrics to assess student's performance and achievement in mathematics and train teachers to use them.

In the same context, [10] conducted a study investigating the possibility of using scoring rubrics as a multidimensional assessment approach to evaluate mathematical power. The study included 62 students from three eighth-grade classes in Turkey and employed a case study methodology. The study results showed that using scoring rubrics effectively contributed to the student's growth in problem-solving skills, decision-making abilities, communication skills, and the evaluation of practical and conceptual knowledge. This demonstrated the potential for assessing learning outcomes and processes.

Furthermore, [30] aimed to examine the effectiveness of performance-based assessment in developing mathematical thinking and problem-solving abilities among secondary school students. The study included a sample of 74 female students from the eleventh grade in a school in Jordan. An experimental group was evaluated using performance-based assessment, while a control group was assessed using traditional methods. The researcher utilized tests for mathematical thinking and problem-solving to collect data. The study results

indicated statistically significant differences favoring the experimental group regarding mean scores in mathematical thinking and problem-solving tests. The researcher recommended further studies on alternative assessment methods, developing specialized assessment guidelines, and diversifying classroom exercises and homework to include performance tasks that stimulate logical thinking.

[12] conducted a study aimed to investigate the impact of using scoring rubrics to assess performance on students' achievement and attitudes towards mathematics. Two scoring rubrics, holistic and analytical, were developed for this purpose. The study included a sample of 128 students divided into three groups: the first experimental group was assessed using analytical scoring rubrics, the second experimental group was assessed using holistic scoring rubrics, and the control group was assessed using traditional methods. Achievement tests and an attitude scale towards mathematics were administered after the study implementation. The study results revealed statistically significant differences attributed to the application of the assessment method in favor of the first and second experimental groups compared to the control group, while no statistically significant differences were found between the two experimental groups.

In a related context, [31] conducted a study aimed at investigating the impact of using analytical performance evaluation criteria on the academic achievement of third-grade students. The study included a sample of 46 students from schools in KSA, divided into two groups: experimental and control. A test was administered to assess academic achievement. The results showed the effectiveness of analytical performance evaluation criteria in assessing students' performance, and the study recommended the use of these criteria in teaching mathematics to elementary stage students.

[9] Conducted a study to examine the effect of scoring rubrics on performance evaluation and its impact on the academic achievement and attitudes of eleventh-grade female students towards mathematics. The purposive sample consisted of two groups: experimental (29 students) and control (28 students). An achievement test of multiple-choice type and an attitudes scale were prepared. The results of the study indicated the effectiveness of using scoring rubrics in evaluating students' performance, and the study recommended the utilization of scoring rubrics in assessing students' performance.

The study of [32] aimed to investigate the effectiveness of using analytical performance evaluation criteria in solving mathematical problems and enhancing the academic achievement of seventh-grade students in KSA. The researcher used an experimental approach with two groups: control and experimental. The study included a sample of 46 students. The researcher developed a student activity guide and a teacher guide for teaching ratio and proportion lessons based on analytical

performance evaluation criteria, along with an achievement test. The results showed the effectiveness of the experimental treatment in improving students' achievement in the achievement test with a significant effect size. The study suggested conducting further research on analytical performance evaluation criteria to enhance learning in other mathematical subjects and different educational stages, as well as comparing them with holistic performance evaluation criteria.

While previous studies mentioned focused on students in different educational stages, covering various topics such as algebraic equations, operations on numbers, ratio and proportion, fractions, and geometry, the current study focused on university-level students in advanced topics related to composite functions and finding inverse functions. Additionally, it also stood out in its nature as a case study, while most previous studies followed either an experimental or descriptive approach. However, this study aligns with previous studies utilizing tests as a data collection tool. This study contributes to supporting previous studies that emphasized the importance of using scouring rubrics in evaluating students' performance in mathematics.

2 Methodologies

Research Design

The case study method is one of the methods of descriptive studies and a method of descriptive research that provides the researcher with quantitative and qualitative data about multiple factors related to an individual, institution, or a small number of individuals.

Participants

The study population consisted of students enrolled and attending Sohar University in Oman during the spring semester 2021/2022. The study sample was purposefully selected from students enrolled in Calculus course. It consisted of the first three students who voluntarily approached the researcher expressing their willingness to participate in this study.

Instruments

Mathematical Achievement Test: After reviewing the theoretical and research literature [9,32,33,34] a Mathematical Achievement Test was developed, the test consisted of three mathematical tasks that covered the process of combinatorial permutations, the nature of combinatorial permutations (whether they are switchable in general or not), finding the inverse permutation and its relationship to the process of combining it with itself, the test tasks were structured as follows: Direct Task: Students answered this task by identifying the given data and applying the rule directly in two steps. Reverse Task: Students used the given data and performed the steps of the previous task in reverse order, linking the results. Comprehensive Task for both Concepts: Students

analyzed the given data, found the inverse permutation, and linked it to the concept of combining it with itself when verifying the correctness of the answer.

To ensure the instrument's validity, it was presented to a group of experts in pure mathematics, mathematics curricula, and teaching methods. Their opinions and suggestions regarding the tasks, language accuracy, question nature, difficulty level, and alignment with the study objectives were considered. Based on their feedback, the test items were revised until the test reached its final form. The opinions of the students were also obtained through a survey sample of students who were selected from outside the study sample. Their feedback was used to modify the wording of the third task, which students had difficulty understanding regarding the method of verifying the solution's accuracy.

To establish the reliability of the instrument, the method of inter-rater agreement was followed by calculating the correlation coefficient between the evaluations of the two researchers for the survey sample. The coefficient was found to be 0.92. This coefficient was chosen because it is difficult to conduct a test-retest reliability coefficient for the students' responses. It is not preferable to reapply the test to the same group with a time interval between the applications, as it may be influenced by the students' recall factor of the test tasks, leading to an increased reliability coefficient. Alternatively, students might become familiar with the test, resulting in higher scores in the second application and reducing the reliability coefficient.

Scoring Rubrics: This study relies on evaluating students' performance to develop the assessment system for university instructors in general and specifically for instructors at Sohar University. Two scoring rubrics were applied to the three tasks given to students: holistic and analytical. A set of descriptors was developed to describe different performance levels on the tasks using the holistic scoring rubric, which includes the following categories: (Perfect, Median, Novice, Weak). Tables 1, 2 and 3 describe each level along with the corresponding score for each performance level on the three tasks.

In the analytical scoring rubrics, estimates have been provided to describe students' different levels of performance in three tasks, which include: (Task Understanding, Solution Planning, Solution Execution, and Solution Verification). Tables 4, 5 and 6 describe each performance level for the three tasks and the corresponding Scoring.

To verify the validity of the rubrics, they were presented to a group of experts in mathematics curriculum and teaching methods. Their opinions and suggestions regarding the criteria on which the rubrics were built, the nature of the assessment, and the extent to which they achieve the study's objectives were taken into consideration. Based on their feedback, some criteria in the scoring rubrics were modified until they reached their final form. To verify the reliability of rubrics, the researchers evaluated the responses of the survey sample

Table 1. Student performance using the holistic scoring rubrics on the first task

Performance Level	Performance Details	Scoring
Perfect	Utilizing rules and algorithms correctly and in a proper sequence to find the matching $(f \circ g)(x)$	2
Median	Improper utilization of rules and algorithms to find the structure, or an incorrect sequence in using the rules in opposite side $g(f(x))$ not $f(g(x))$	1
Weak	Erroneous utilization of rules and algorithms without considering any sequencing to find the structure, or no solution exists	0

Table 2. Student performance using the holistic scoring rubrics on the second task

Performance Level	Performance Details	Scoring
Perfect	Proper utilization of rules and algorithms involves finding $(g \circ f)(x)$ and establishing the relationship between $(f \circ g)(x)$, followed by performing a comparison $(f \circ g)(x)$	2
Median	Incorrect utilization of rules and algorithms, failing to establish any connection between direct and inverse usage, thus neglecting to mention the relationship between them.	1
Weak	Incorrect application of the inverse procedure without addressing any relationship between it and the original procedure, or no solution exists.	0

Table 3. Student performance using the holistic scoring rubrics on the third task

Performance Level	Performance Details	Scoring
Perfect	Proper utilization of rules and algorithms with a correct sequence and validating the correctness of the solution using two valid methods	3
Median	Proper utilization of rules and algorithms with a correct sequence and validating the correctness of the solution using one valid method	2
Novice	Proper utilization of rules and algorithms with a correct sequence, but without verifying the accuracy of the solution, or errors in the operations resulting in an incorrect final answer	1
Weak	Incorrect utilization of rules and algorithms with an incorrect sequence (resulting in an incorrect solution), without verifying the accuracy of the solution, or no solution exists	0

twice, with a one-week interval between evaluations. The results of each evaluator were compared independently, and the agreement coefficient was measured using Holsti's equation [35], which yielded 0.95. Additionally, the Pearson correlation coefficient was calculated for the evaluations of student performances according to the holistic and analytical rubrics, resulting in a coefficient of 0.98.

Interviews: An interview was conducted with each student individually. The interview was of a semi-structured nature, consisting of pre-prepared questions with the possibility of introducing new questions based on the interview conditions and the students' responses to the prepared questions. A set of questions was asked to identify the key criteria in the student evaluation process. The interview consisted of eight questions, and the students provided explanations and clarifications regarding their procedures within a timeframe of 20-30 minutes per student.

To verify the validity of the interview questions, interviews were conducted with a sample of survey participants. The questions were also presented to a group

of experts in mathematics curriculum and teaching methods. Their opinions and suggestions regarding the questions' alignment with the study objectives were considered. Based on their feedback, some questions were modified until the final form of the interview questions was achieved.

Data Gathering Procedures

During the two weeks prior to conducting the study, students were exposed to an objective explanation of the concept of pairing and finding the inverse pairing. The pilot sample of students was presented with a short test containing the three tasks of the study during one of the lectures. The students' feedback on the test was taken into consideration, leading to adjustments in the test duration, which was increased from 15 to 20 minutes. Additionally, based on their feedback, a modification was made to the third task due to students' expressed lack of direct understanding of the purpose of verifying the answer's correctness. Following these adjustments, the test was administered to the study sample under natural conditions as a short classroom test. Prior to taking the test, the

Table 4. Student performance using the analytic scoring rubrics on the first task

Performance Level	Performance Details	Scoring
Task Understanding	Clear and accurate understanding is demonstrated through writing the rules for both $f(x)$ and $g(x)$ and identifying the task's objective, which is to find $(f \circ g)(x)$	2
	Clear identification of the given information without specifying the objective $(f \circ g)(x)$, or specifying the objective without writing the given data	1
	Failure to specify the given data or the objective (indicating a lack of understanding of the problem)	0
Solution Planning	The required strategy is complete: $f(g(x))$ followed by a correct sequence of steps	2
	The strategy only addresses a portion of $g(x)$, or the strategy does not lead to solving both $g(x)$ and $f(x)$ separately.	1
	There is a lack of strategy.	0
Solution Execution	Correct manner: Applying f to g .	2
	Incorrect solution execution: Applying g to f .	1
	Failure to execute the solution.	0
Solution Verification	Complete verification: Repeating the solution to ensure its accuracy.	2
	Partial verification: Incorrectly repeating the solution.	1
	Failure to verify.	0

Table 5. Student performance using the analytic scoring rubrics on the first task

Performance Level	Performance Details	Scoring
Task Understanding	Clear and accurate understanding is demonstrated through writing the rules for the two mappings, $f(x)$ and $g(x)$, and specifying the objective of the task, which is to find $(g \circ f)(x)$. The understanding is directly related to the first task.	2
	Specifying the given data and identifying the first objective $(g \circ f)(x)$ without understanding its relationship to the first task, or specifying the given data and the first objective $(g \circ f)(x)$ with a misunderstanding of the second objective (its relationship to the first task)	1
	No given data, no objective specified, and no understanding of the relationship between the objective and the first task.	0
Solution Planning	Complete plan: Sequential steps to find $(g \circ f)(x)$ and compare it to $(f \circ g)(x)$.	2
	A plan to find $(g \circ f)(x)$ without connecting it to the first task.	1
	No plan provided.	0
Solution Execution	Correct solution execution: Finding $(g \circ f)(x)$, comparing it to $(f \circ g)(x)$, and writing the complete commentary	2
	Correct solution execution without linking it to the first task, or incorrect solution execution with an incorrect link to the first task.	1
	Failure to execute the solution.	0
Solution Verification	Complete verification: Repeating the process of finding $(g \circ f)(x)$ and ensuring its relationship to the first task.	2
	Incorrect repetition of the solution.	1
	Failure to verify the solution.	0

students had been provided with detailed explanations of the two correction methods: the holistic and analytical correction methods.

After the test, individual interviews were scheduled with the three students. Their answers to the interview questions were recorded, and each student's performance in the test tasks was assessed according to the holistic and analytical correction methods. A comparison was then made between the assessment results and the notes taken for discussion with each student. During the interview, the previously recorded observations were discussed, and

additional questions that arose during the interview were posed to clarify the students' problem-solving procedures. The student's performance was evaluated again after the interview, and a comparison was made between their evaluation and performance in the test.

Data Analysis

The researchers recorded the students' responses to each task independently. Each researcher read the student's answers for each task and identified the issues related to the study questions. They then formed

Table 6. Student performance using the analytic scoring rubrics on the first task

Performance Level	Performance Details	Scoring
Task Understanding	Clear, accurate, and comprehensive understanding is demonstrated through writing the rule for $f(x)$, specifying the objective as finding $f^{-1}(x)$, and mentioning the validation rules for confirming the solution	2
	Partial understanding: Providing the given data without specifying the objective.	1
	No understanding demonstrated.	0
Solution Planning	The strategy is complete: Symbol substitution, followed by crucial steps, and then an overview for verification	2
	The strategy is incomplete: Steps are provided without symbol substitution, and there is no plan for verification.	1
	No strategy or verification plan provided.	0
Solution Execution	Correctly: Applying the steps in a proper sequence.	2
	Completely incorrect: Making errors in executing the steps or in the symbols used.	1
	Failure to execute the solution.	0
Solution Verification	Complete verification: Checking in at least one direction or both directions.	2
	Partial verification: Not reaching a conclusive result.	1
	No verification conducted.	0

preliminary generalizations and confirmed their validity based on the students' responses to the test. The analysis of the interview transcripts relied on portions of the data related to the study questions and issues related to the students' understanding of the intended meaning of the assigned task and the reasons that prevented them from completing certain parts of the solution.

3 Results

Results of the First Question: "Do the two scoring rubrics provide consistent interpretations of scores at the same level?"

To answer this question, the researchers monitored the analysis results of students' responses to the test tasks. These results are presented in Tables 7 and 8.

The inter-rater agreement coefficient between the researchers was 0.92, based on 3 different ratings out of a total of 18. This coefficient indicates a high level of agreement and is acceptable for result reliability. Additionally, the Pearson correlation coefficient between the first researcher's evaluations of student performance and the second researcher's evaluations was 0.98. This high correlation value indicates a very strong relationship between the evaluators, almost reaching perfect positive correlation [36, 37].

Students' Responses Results

In the first task, it was observed that all three students successfully completed the task in terms of holistic rubrics, with a significant convergence in the analytical rubrics. However, none of them reached the highest level of scores in this criterion. In the second task, the same level of achievement was observed in the holistic rubrics, with a significant convergence in the analytical rubrics.

However, none of them reached the final answer by connecting the task with its predecessor in both rubrics. As for the third task, none of the students were able to reach the highest level of performance in this task, with slight variation among them in both the holistic and analytical rubrics.

It can be concluded that only the first task achieved the highest level of scores in the holistic rubrics. The students justified this by the clarity of the task procedures. In the remaining tasks, students' estimates in the analytical rubrics were higher than those in the holistic rubrics. Additionally, the students expressed satisfaction with this scoring method, as it provided them with a sense of fairness compared to the holistic rubrics by considering the details of their responses.

Results of the Second Question: "Does students' knowledge of the details of the scoring rubrics contribute to improving their performance?"

Based on the students' responses to the interview questions, including a specific question about the extent of their benefit from knowing the details of the scoring rubrics they would be exposed to in their task performance, the students unanimously affirmed that it was comforting to be familiar with these details. They stated that it helped them gain confidence in themselves during task performance. However, it did not significantly or tangibly contribute to raising their response levels, mainly because they were not accustomed to this approach, particularly the analytical rubrics.

Interviews Responses Results

The interviews focused on students' responses to the tasks in terms of completion and verification. The students' responses reflected the following aspects:

Table (7): Evaluators' by using holistic scoring rubric

Student	First Evaluator				Second Evaluator			
	Task 1	Task 2	Task 3	Total	Task 1	Task 2	Task 3	Total
First	2	1	1	4	2	1	0	3
Second	2	1	1	4	2	0	2	4
Third	2	1	2	5	2	1	2	5

Max score for the 1st and 2nd tasks is (2), for the 3rd task is (3) and overall exam is (7)

Table (8): Evaluators' by using analytic scoring rubric

Student	First Evaluator				Second Evaluator			
	Task 1	Task 2	Task 3	Total	Task 1	Task 2	Task 3	Total
First	6	5	6	17	6	4	6	16
Second	5	4	5	14	6	4	5	15
Third	6	6	6	18	5	6	6	17

Max. score for the 1st and 2nd tasks is (2), for the 3rd task is (3) and overall exam is (7)

- The students' focus is mainly on their response as an action and its execution, aiming to achieve a final result.
- Lack of full confidence in their actions may lead students to not complete the task.
- Students' desire to satisfy their teacher may prevent them from writing their procedures or ideas about the action, unless they have complete confidence in them.
- Verifying the correctness of the solution is not a significant concern for students, either due to their confidence in the result, the perceived insignificance of this step, or its absence as a direct requirement in the task.
- There is a tendency to focus on each task independently without linking it to another task that may be closely related.
- Students consider the process of verifying the correctness of the solution as difficult and requiring unusual skills that they do not possess.
- Students feel satisfied when they are informed about the details of the evaluation process that their responses will undergo. However, this awareness is reflected more in their attention and commitment to following principles and rules rather than their focus on verification and evidence.

4 Discussions

Discussion of the First Question: "Do the two scoring rubrics provide consistent interpretations of scores at the same level?"

The results lean in favor of the analytical rubric in terms of scoring. The reason for this may be that this method of scoring takes into consideration all the details of the procedures, as well as the levels of understanding, perception, solution, and verification by the student. This generates a conviction in the student that their right will not be completely lost if they fail in some aspects. They trust in some aspects and try in others, even if they are not confident in their ability to accomplish them.

Additionally, this type of assessment pays attention to details that the student would not have been concerned with if they had not been exposed to the grading method, such as writing a solution plan and verifying its correctness.

This result aligns with the goals of modern assessment systems by calling for the highlighting of mathematics that students can know and perform, and providing each student with the opportunity to demonstrate their mathematical abilities according to their capabilities. This positively reflects on students' performance in tasks. Furthermore, introducing the student to the assessment method and the criteria by which their performance will be judged can have a positive impact in alleviating student anxiety. The use of scoring rubrics integrates learning and assessment processes, resulting in students acquiring knowledge, understanding, and skills using various teaching styles and strategies [12,28,38,39,40] to achieve that knowledge through a multidimensional assessment of their performance.

Discussion of the Second Question: "Does students' knowledge of the details of the scoring rubrics contribute to improving their performance?"

It is natural for a student's confidence in their actions and their value to increase when they become familiar with the scoring rubrics details they will be assessed against. However, it is necessary for them to become accustomed to taking such actions that will be considered in the scoring process for performance before they are exposed to specific tasks. This way, the value of their actions in responding to these tasks becomes greater, and they develop a conviction that if they do not achieve satisfactory results in some aspects, they will do so in other aspects. This may be attributed to the fact that performance-based learning is more focused on the student, as performance tasks help students understand their strengths to enhance them and identify weaknesses to address them. Additionally, performance tasks evaluate processes and learning outcomes, providing an accurate assessment of the student's actual performance.

Additionally, introducing students to the assessment method and criteria used to judge their performance can have a positive impact on alleviating their anxiety, as indicated by the results of some studies [3, 18, 36, 40].

Moreover, the use of scoring rubrics aims to identify the individual student's level of performance, and the student should feel that in order to reach a level of excellence, they must be diligent and persevering. This positively impacts the student's achievement. It is evident from the above that using scoring rubrics to assess student performance promotes integration between the processes of learning and assessment. This integration has a positive impact on students' achievement by acquiring knowledge, understanding, and skills through employing multiple learning styles and teaching strategies [16, 21, 33, 39, 40] to reach that knowledge through a multidimensional assessment of their performance.

Discussion of the Interview Responses

The students affirmed their satisfaction with their familiarity with the scoring rubrics that would be applied to evaluate their task responses through the interview responses conducted with them [41, 42, 43]. One student stated that they may have unintentionally considered what was mentioned in the details of the assessment method, but they would purposefully do so later, especially considering its importance in the utilized evaluation considerations.

5 Conclusions

This study plays a crucial role in the field of performance-based assessment and educational research, as it provides an analytical and methodological framework for understanding and assesses students' educational performance in Calculus. It also contributes to the development of performance-based assessment-dependent educational programs [44, 45]. This is achieved through the following: Student Performance Assessment: The use of various assessment tools, such as the scoring rubrics, helps develop assessment models to measure students' achievement in Calculus, as well as in various mathematical fields. It helps determine the progress of students in understanding Calculus, enhances their strengths, and addresses their weaknesses. School Performance Analysis: Different assessment tools can be used to assess students' performance based on statistical data and quantitative information available in identifying the factors that affect their success. Educational Effectiveness Assessment: Various assessment tools can be used to employ statistical analysis techniques to measure the effectiveness of educational programs and teaching methods, as well as to identify effective ways to achieve educational objectives. Mathematics Curriculum Enhancement: The use of various assessment tools helps in the development of Calculus curricula specifically, and mathematics curricula

in general, making them more up-to-date and suitable for students' needs and the requirements of the job market [46]. Finally, Development better teaching methods that encourage active learning and enhance students' understanding of mathematical concepts.

6 Recommendations

Based on the positive results obtained from this study, the researchers recommend the following: Increase teachers' awareness and understanding of the value of encouraging students to explain their procedures while performing assigned tasks. Create an environment that facilitates this in the classroom and focus on verifying the correctness of solutions to increase students' confidence in themselves and their procedures. Encourage teachers to push students to deepen their understanding and awareness of the steps involved in their procedures, as this has an impact on their learning and progress. Encourage mathematics teachers at Sohar University and various Omani universities to prepare and use different types of performance scouring rubrics to assess their students' performance based on performance-based assessment criteria. Call upon curriculum and textbook authors, especially those responsible for the Calculus course and other mathematics courses in general, to make necessary additions and modifications based on the results of this study. This will increase opportunities for students to justify their procedures through tasks that require it or incorporate it to verify the solution. Encourage planners and developers of the Calculus course, as well as other mathematics courses, to focus on performance scouring rubrics by enriching the curriculum with performance-based tasks that are built on assessment criteria. Conduct similar studies on other tasks that require students to justify their procedures and demonstrate logic and sequencing.

7 Summary

Performance assessment is significant in education because it provides a more authentic and comprehensive evaluation of students' abilities, preparing them for real-world challenges. Rubrics are essential tools in this process, offering clarity, consistency, and the opportunity for constructive feedback to both educators and students, ultimately enhancing the educational evaluation process. Performance-based learning is more student-centered. Performance tasks help students understand their strengths to enhance them and identify their weaknesses for improvement. They also evaluate learning processes and outcomes, providing an accurate assessment of students' actual performance.

Acknowledgments

None

Conflicts of Interest Statement

The authors declare that there is no conflict regarding the publication of this paper.

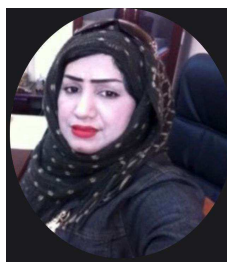
References

- [1] NCTM. Principles and Standards of School Mathematics, The National Council of Teachers of Mathematics, Inc. (2000).
- [2] Az-Zo'bi, E. A. On the reduced differential transform method and its application to the generalized Burgers-Huxley equation. *Applied Mathematical Sciences*, **8**, 8823–8831, (2014). <https://doi.org/10.12988/ams.2014.410835>
- [3] Az-Zo'bi, E. A., Dawoud, K. A., & Marashdeh, M. Numeric-analytic solutions of mixed-type systems of balance laws. *Applied Mathematics and Computation*, **265**, 133–143, (2015). <https://doi.org/10.1016/j.amc.2015.04.119>
- [4] Tami S., & Roger D. Performance-Based Assessment of Secondary Mathematics Student Teachers. *Action in Teacher Education*, **22(3)**, 86-95, (2000).
- [5] Tashtoush, M., Alali, R., Wardat, Y., AL-Shraifin, N., & Toubat, H. The Impact of Information and Communication Technologies (ICT)-Based Education on the Mathematics Academic Enthusiasm. *Journal of Educational and Social Research*, **13(3)**, 287-296, (2023).
- [6] Al-Subaie, N. & Al-Khudair, A. The reality of electronic evaluation tools in loghati alkhaledah course from the point of view of teachers the intermediate school in Al-Ahsa governorate during Corona Pandemic (COVID-19). *International Journal of Psychological and Educational Studies*, **12(1)**, 160-193, (2023).
- [7] Assessment Strategies and Tools Manual (Theoretical Framework). Prepared by the National Assessment Team, Examinations Directorate, Ministry of Education, Jordan, (2004).
- [8] Tashtoush, M., Wardat, Y., Aloufi, F., & Taani, O. The Effectiveness of Teaching Method Based on the Components of Concept-Rich Instruction Approach in Students Achievement on Linear Algebra Course and Their Attitudes Towards, *Journal of Higher Education Theory and Practice*, **22(7)**, 41-57, (2022).
- [9] Al-Ruwaili, E. The scoring rubrics for performance assessment and its impact on the achievement and attitudes of the 11th grade students towards mathematics, *Studies: Educational Sciences*, **43(5)**, 1903-1914, (2016).
- [10] Sahin, S., & Baki, A. A new model to assessment mathematical power, *Procedia- Social and Behavioral Sciences*, **9**, 1368–1372, (2010).
- [11] Mathematics Assessment Document for Students for Grades (5-10). Prepared by the General Directorate of Educational Assessment, Ministry of Education, Oman, (2012).
- [12] Abu Obeid, A. The Scoring Rubrics for performance assessment and its impact on the achievement and attitudes of the 11th grade students towards mathematics, *Psychological and Educational Studies*, **7**, 25-57, (2011).
- [13] Moscal, B. Recommendations for developing classroom performance assessments and scoring rubrics, *Practical Assessment, Research and Evaluation*, **8(14)**, 1-9, (2003).
- [14] Tashtoush, M., Wardat, Y., Alali, R., AL-Shannaq, M., Saleh, S., & AL-Saud, K. Conceptual Understanding of Systems of Linear Equations: Difficulties and Challenges. *Information Sciences Letters*, **12(12)**, 2491-2503, (2023).
- [15] Cohen, A. Assessing Language Ability in the Classroom, 2nd ed. Heinle and Heinle, (1994).
- [16] Hart, D. Authentic Assessment: A Handbook for Educators, Reading, MA: Addison Wesley Publishing Company, (1994).
- [17] Alarabi, K., & Wardat, Y. UAE-based teachers' hindsight judgments on physics education during the COVID-19 pandemic. *Psychology and Education Journal*, **58(3)**, 2497-2511, (2021).
- [18] Tashtoush, M., Wardat, Y., Aloufi, F., & Taani, O. The Effect of a Training Program Based on (TIMSS) to Developing the Levels of Habits of Mind and Mathematical Reasoning Skills among Pre-service Mathematics Teachers. *EURASIA Journal of Mathematics, Science and Technology Education*, **18(11)**, Article No em2182, (2022).
- [19] Al-Shirawia, N. & Tashtoush, M. Differential Item Functioning Analysis of an Emotional Intelligence Scale for Human Resources Management at Sohar University. *Information Sciences Letters*, **12(11)**, 2937-2952, (2023).
- [20] Shirawia, N., Alali, R., Wardat, Y., Tashtoush, M., Saleh, S., & Helali, M. Logical Mathematical Intelligence and its Impact on the Academic Achievement for Pre-Service Math Teachers. *Journal of Educational and Social Research*, **13(6)**, 242-257, (2023).
- [21] Stanley, T. Performance-Based Assessment for 21st-Century Skills, 1st ed. Taylor and Francis Group, (2014).
- [22] Wardat, Y., Tashtoush, M., Alali, R., & Saleh, S. Artificial Intelligence in Education: Mathematics Teachers? Perspectives, Practices and Challenges. *Iraqi Journal for Computer Science and Mathematics*, **5(1)**, (2024).
- [23] Tashtoush, M., Wardat, Y., & Elsayed, A. Mathematics Distance Learning and Learning Loss During COVID-19 Pandemic: Teachers' Perspectives. *Journal of Higher Education Theory and Practice*, **23(5)**, 162-174, (2023).
- [24] Mertler, C. Designing Scoring Rubrics for your Classroom, *Practical Assessment, Research, and Evaluation*, **7(25)**, 1-8, (2001).
- [25] Nitko, A. Educational Assessment of Students, 3rd ed. Merrill, (2001).
- [26] McLellan, S. When Students Teach: Performance Based Assessment, *Transformative Dialogues: Teaching & Learning Journal*, **2(2)**, 1-12, (2008).
- [27] McBride, B. & Carifio, J. Empirical Results of using an Analytic versus Holistic Scoring Method to Score Geometric Proofs, ERIC Document Reproduction No. ED 401307, (1995).
- [28] Lumely, D. & Yan, W. The Impact of state Mandated, Large – Scale Writing Assessment in Pennsylvania, ERIC Document Reproduction No. ED 453220, (2001).
- [29] Al-Absi, M. The scoring rubrics for performance assessment and its impact on the achievement and attitudes of tenth

- grade students towards mathematics, *Journal of Educational Sciences*, **12**(12), 133-157, (2007).
- [30] Balawneh, F. The effect of the performance-based assessment on developing mathematical thinking and problem-solving ability among secondary school students, *An-Najah University Journal for the Humanities*, **24**(8), 2227-2270, (2010).
- [31] Al-Maliki, A. The effect of using analytical rubric on the academic achievement of third graders of primary school, *Journal of Educational and Psychological Sciences*, **15**(3), 285-298, (2011).
- [32] Sarhani, M. The Effectiveness of Using Analytical rubric to Solve Mathematical Problems in Developing Academic Achievement for 7th Grade Students in KSA, *Journal of the College of Education*, **171**(1), 665-684, (2016).
- [33] Tashtoush, M. *Calculus 1 With Examples*, 1st Edition, Text book in Calculus, Dar AL-Amal for Publishing and Distributing, Jordan, (2013).
- [34] Rasheed, N. & Tashtoush, M. The Impact of Cognitive Training Program for Children (CTPC) to Development the Mathematical Conceptual and Achievement. *Journal of Higher Education Theory and Practice*, **23**(10), 218-234, (2023).
- [35] Holsti, R. *Content Analysis for Social Sciences and the Humanities* Addison, Welay Publishing Company, (1969).
- [36] Tashtoush, M., Alshunaq, M., & Albarakat, A. The Effectiveness of Self-Regulated Learning (SRL) in Creative Thinking for CALCULUS Students, *PalArch's Journal of Archaeology of Egypt/ Egyptology*, **17**(7), 6630-6652, (2020).
- [37] Wardat, Y., Tashtoush, M., Alali, R., & Jarrah, A. ChatGPT: A Revolutionary Tool for Teaching and Learning Mathematics. *EURASIA Journal of Mathematics, Science and Technology Education*, **19**(7), 1-18, Article No: em2286, (2023).
- [38] Rasheed, N. & Tashtoush, M. The Fertility and its Relation with Some Demographic, Economic and Social Variables in Jordan. *Turkish Journal of Computer and Mathematics Education*, **12**(11), 5088-5095, (2021).
- [39] Fannakhosrow, M., Nourabadi, S., Huy, D., Trung, N., & Tashtoush, M. A Comparative Study of Information and Communication Technology (ICT)-Based and Conventional Methods of Instruction on Learners' Academic Enthusiasm for L2 Learning, *Education Research International*, **2022**, Article ID 5478088, (2022).
- [40] Abu Saleh, M. & Awad, A. *Introduction to Statistics*, Jordan Book Center, Amman, (1997).
- [41] Az-Zo'bi, E. A. (2018). A reliable analytic study for higher-dimensional telegraph equation. *The Journal of Mathematics and Computer Science*, **18**(04), 423-429. <https://doi.org/10.22436/jmcs.018.04.04>
- [42] Linda, M. Writing to the Rubric: Lingerin Effects of Traditional Standardized Testing on Direct Writing Assessment, *Phi Delta Kappan*, **80**(9), 673-679, (1999).
- [43] Al-Shirawia, N., Al-Kiyumi, M., Al-Shibli, F., & Tashtoush, M. Building a scale of emotional intelligence for human resource management by using the graded-response model, *Migration Letters*, **20**(8), 500-524, (2023).
- [44] Az-Zo'bi, E., Kallekh, A., Rahman, R., Akinyemi, L. Bekir, A., Ahmad, H., Tashtoush, M., & Mahariq, I. Novel topological, non-topological, and more solitons of the generalized cubic p-system describing isothermal flux. *Optical and Quantum Electronics*, **56**(1), Article ID 84, (2024).
- [45] Wardat, Y., Tashtoush, M., Alali, R., & Saleh, S. (2024). Artificial Intelligence in Education: Mathematics Teachers' Perspectives, Practices and Challenges. *Iraqi Journal for Computer Science and Mathematics*, **5**(1), (2024).
- [46] Tashtoush, M., Wardat, Y., AlAli, R., & AL-Saud, K. The Impact of Cyberbullying on Student Motivation to Learn: In-Sights from Abu Dhabi Emirate Schools, *Humanities and Socail Science Letters*, **11**(4), (2023).



Nawal H. Shirawi is graduated from the University of Islamic Sciences in Malaysia. She holds a Ph.D. in Measurement and Evaluation, a Bachelor in Mathematics and Physical Education. Since 2008, she has been an employee at Sohar University. She served as a program coordinator for several educational disciplines. She has published several research papers on measurement and evaluation. She got a grant for research, funded by The Research Council at the Ministry of Higher Education, Research and Innovation for the academic year 2021, for building an emotional intelligence scale for human resources management at Sohar University, Sultanate of Oman.



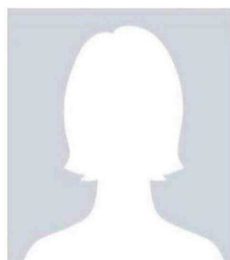
Aida B. Qasimi holds a Ph.D. in Education with a specialization in Educational Foundations and Administration from Cairo University, obtained in 2009. She has been an Assistant Professor at Sohar University College of Education and Arts since September 2016.

She is also the Head of the Educational Rehabilitation Department. She is a member of the Editorial Board of the *Journal of the College of Education and Arts for Humanities and Social Sciences* at Sohar University. Additionally, she is a member of the University's Advisory Committee, a reviewer for post-graduate programs, and an academic researcher in educational sciences.



Mohammad Tashtoush is an Assistant Professor in the Faculty of Education & Arts, Sohar University, Oman, and Lecturer of Math & STAT at AL-Balqa Applied University, JUST, AL-Hussein bin Talal University, Jordan. He received his Ph.D. in

Curriculum and Instruction in Mathematics Education from Yarmouk University, Jordan, in 2019, and M. Sc. & B. Sc. in Pure Mathematics. His research interest includes a range of interdisciplinary areas: Applied Mathematics, Mathematical Physics, Curriculum and Instruction, Curriculum Development, ICT in Education, Achievement, Teaching Methods, Assessment, Learning Losses, Blended Learning, TIMSS, PISA, STEM and ChatGPT. He has published more than 40 peer-reviewed articles in different journals indexed in global databases, he is a reviewer for more than 20 indexed journals and a member of the editorial board for three journals.



Noha Rasheed hold a bachelor's and master's degree in statistics from Yarmouk University and holds a PhD in mathematics curricula and teaching methods from Yarmouk University, Jordan, 2018. She has sustained experience in university teaching in several universities in

Saudi Arabia and Jordan. Her research interests are in teaching strategies and methods, evaluation. Students, achievement, and educational technology. She has published many classified international scientific journals.



Mohamad A. S. Khasawneh hold a PhD in Special Education from Amman Arab University, 2009. He is currently working as an Assistant Professor in Special Education at King Khalid University from 1/9/2015 until now. He worked as an assistant

professor of special education at the University of Hail from 1/9/2009 to 1/9/2014. He is a licensed Photosensitivity Syndrome (SSS) screener by the Irlen Helen Center. And a licensed examiner to measure and diagnose learning difficulties from the Jordanian National Institute for Learning Difficulties. He is also a scientific advisor to Dar Al-Fikr for Publishing and Distribution, Amman. He is an academic scholar in Wasit University Journal.



Emad Az-Zo'bi is a professor of applied mathematics/differential equations at Mutah University. He received his PhD in Mathematics from the University of Jordan in 2011. Professor Az-Zo'bi was listed by Stanford University and Scopus as one of the top scientists in 2021, and has

published about 80 papers in highly indexed journals.