International Journal of Innovative Research and Scientific Studies, 8(2) 2025, pages: 12-25

IJIRSS International Journal of Innovative Research & Scientific Studies

ISSN: 2617-6548

URL: www.ijirss.com

Assessing the quality of science teachers' practices in employing formative assessment as a learning tool to enhance children's primary science learning

Rommel Mahmoud AlAli[1*], Ali Ahmad Al-Barakat[2]

[1]*The National Research Center for Giftedness and Creativity, King Faisal University, Al-Ahsa, Kingdom of Saudi Arabia.*
[2]*Faculty of Educational Sciences, Yarmouk University, Irbid, Jordan.*

Corresponding author: Rommel Mahmoud AlAli (*Email: ralali@kfu.edu.sa*)

Abstract

This study examined the quality of science teachers' practices in employing formative assessment as a learning tool to enhance children's primary science learning. The research sample consisted of 30 randomly selected male and female teachers from Jordanian primary schools during the second semester of the 2023/2024 academic year. To assess these practices, a verbal rating scale was developed as a classroom observation instrument, incorporating ten criteria for assessing the quality of formative assessments. Each criterion included several performance indicators, totaling 38 indicators that represent formative assessment practices. The findings revealed that the overall quality of these practices was consistently low, with all practices classified within the "low" practice degree. This indicates a significant gap in adherence to formative assessment quality standards in science learning environments. Additionally, non-parametric analysis showed no significant correlation between the quality of these practices and the teachers' gender or teaching experience. The study concludes with a set of recommendations aimed at improving the quality of formative assessment practices in science learning environments.

Keywords: Formative assessment, Learning environments, Practices, Science learning, Science teachers.

DOI: 10.53894/ijirss.v8i2.5090
Funding: This work was financially supported by the Deanship of Scientific Research, King Faisal University (Grant number: KFU250838).
History: Received: 23 January 2025 / **Revised:** 22 February 2025 / **Accepted:** 27 February 2025 / **Published:** 4 March 2025

Competing Interests: The authors declare that they have no competing interests.
Authors' Contributions: Both authors contributed equally to the conception and design of the study. Both authors have read and agreed to the published version of the manuscript.
Transparency: The authors confirm that the manuscript is an honest, accurate, and transparent account of the study; that no vital features of the study have been omitted; and that any discrepancies from the study as planned have been explained. This study followed all ethical practices during writing.
Acknowledgements: The authors extend their sincere gratitude to the Deanship of Scientific Research at King Faisal University for its financial support of this research.
Publisher: Innovative Research Publishing

1. Introduction

Current trends in education emphasize the importance of providing science teachers with a range of skills to enhance the quality of learning outcomes for students. To ensure effective learning experiences for students, paying attention to the teaching methods used by educators is crucial. Particularly important is the ability of teachers to assess student progress within the classroom. Global efforts to reform science education stress the significance of integrating new and innovative assessment techniques throughout the teaching and learning process. These approaches promote not only teaching and learning environments but also empower students to take charge of their own educational objectives. This shift in responsibility helps lessen the load on teachers when it comes to assessments. Good assessment methods are known for being versatile and easy to use for teachers [1]. Global guidelines for science education stress that how well students learn is greatly impacted by how teachers assess them in class. This underscores the importance of students gaining the knowledge and skills outlined in science education curricula [2, 3].

Assessing students through formative assessment is pivotal in science education environments, as it plays a role in all aspects of teaching and learning processes. This type of assessment is focused on achieving the desired learning goals by recognizing and addressing students' strengths and weaknesses. To ensure the effectiveness of assessment practices, it is crucial to adhere to transparent criteria that serve as both guidance and support for teachers to enhance their professional skills and ultimately improve student learning outcomes (Menéndez, et al. [4]). Menéndez, et al. [4] suggest that incorporating assessment is crucial in the field of science education to help meet learning goals by consistently recognizing and addressing student requirements over time. The implementation of formative assessment strategies and practices hinges on adhering to benchmarks that enhance teachers' expertise and their professional performance, consequently elevating the quality of student education.

Schildkamp, et al. [5] stress the importance of formative assessment in education as a means to enhance student performance by recognizing and meeting their learning requirements throughout the teaching and learning journey. This method involves providing consistent feedback to help teachers tailor their teaching methods and provide better assistance to students effectively. Schildkamp, et al. [5] emphasize the importance of adhering to clear and well-defined standards to guide formative assessment practices, which is crucial for enhancing teachers' professional performance and improving student learning outcomes.

Kulasegaram and Rangachari [6] highlight that numerous science education reform movements stress the need to adopt integrated standards to ensure high-quality teacher performance, with a particular focus on formative assessment of student learning. To meet these standards, teachers must gain experience in providing regular, diverse feedback and fostering students' active engagement with this feedback to achieve high-quality learning outcomes [7, 8]. Global standards for formative assessment quality advocate for varied assessment strategies to enhance science learning [9]. These strategies emphasize learner-centered approaches within the teaching and learning process, recognizing that effective assessment is crucial for promoting meaningful learning. Such assessments are instrumental in assessing students' abilities to think critically, acquire and construct knowledge, and apply it in various real-life contexts [10, 11].

To advance student-centered learning through formative assessment, teachers need to shift from traditional assessment methods that prioritize memorization and recall to more effective classroom practices that support achieving all intended science learning outcomesHuang, et al. [10] and Schildkamp, et al. [5]. Ozan and Kincal [12] emphasize that formative assessment provides feedback to both teachers and students, allowing teachers to refine instructional practices and helping students monitor and enhance their learning. As contemporary teaching methods evolve, formative assessment has gained prominence, advocating for adjustments in practices and policies to ensure accountability in achieving educational reforms, a concept known as assessment-based accountability. Adarkwah [13] posits that assessment encompasses all activities undertaken by teachers to assess students and by students for self-assessment, generating feedback that improves both teaching and learning. Assessment methods and practices ought to be interactive, formative and done in cooperation, with the goal of enhancing students' understanding and proficiency as their attitudes, in specific subject areas.

The prevailing viewpoint, on evaluation sees learning and assessment as intertwined processes that mutually benefit and elevate each other. The methods used to gather assessment data are crucial in determining what teachers should teach and what students should learn, necessitating a variety of approaches to capture this information accurately [14]. This perspective also emphasizes the importance of students applying scientific knowledge in real-world contexts, mirroring those they will encounter outside the classroom. It underscores the need for assessments to align with instructional programs, ensuring that students derive meaningful insights from their experiences [15-18].

Key science learning outcomes, such as conducting scientific investigations and understanding the natural world, can be assessed using various methods and tools AlAli [19] and Meng [18]. Chand and Pillay [20] differentiate between summative assessment, which assesses learning after it has occurred, and formative assessment, which is integrated into the learning process to diagnose student needs, plan instruction, and provide feedback that empowers students to improve their learning and succeed.

To effectively employ formative assessment for learning, the classroom assessment cycle involves four key stages: identifying specific learning goals, preparing a student-centered teaching plan, collecting and analyzing evidence through various methods, and modifying instruction based on the findings Kanjee and Bhana [21]. Tobler [22] introduced a method integrating formative assessment into primary science through the 5E learning cycle strategy, which includes:

- Engagement: Presenting activating questions to stimulate interest and assess prior knowledge, introducing new concepts.
- Exploration: Engaging students in activities and discussions to observe phenomena, compare ideas, and reveal their understanding and misconceptions.
- Explanation: Students articulate their experiences and solve problems using a structured model, demonstrating their understanding and highlighting difficulties.

- Elaboration: Connecting experiences to new situations and summarizing key concepts through formative assessment activities.
- Assessment: Assessing understanding through brief tests to identify weaknesses, enabling targeted support. While formative assessment is continuous, this stage uses grades or scores to gauge learning levels.

Swanson and Clarke-Midura [23] emphasize that formative assessment involves teacher-led processes such as observation, discussion facilitation, and assignment evaluation. Hodgson and Watts [24] highlight that effective assessment focuses on three critical questions: Where are students headed? Where are they now? How will they get there? Addressing these questions is essential for student achievement in science, requiring continuous evaluation, active learning, and opportunities for applying knowledge in real-life contexts.

Chemeli [25] proposed five key strategies for developing effective daily formative assessment activities across all learning domains and grade levels: (1) Identifying learning objectives and success criteria; (2) Organizing effective classroom discussions, questions, and tasks; (3) Providing feedback to students; (4) Encouraging student responsibility for their own learning; and (5) Activating students as resources for each other. Educational research Van Der Steen, et al. [26]; Schildkamp, et al. [5] and Menéndez, et al. [4] underscores the importance of both planned and interactive formative assessment in enhancing science learning and teaching environments. Planned formative assessment allows teachers to gather general information about student progress and provide feedback on teaching effectiveness, while interactive formative assessment, conducted during classroom interactions, enables real-time observation, differentiation, and responsive teaching, fostering effective feedback, active engagement, adaptive instruction, motivation, self-confidence, and self-assessment.

Acknowledging the crucial role of formative assessment, ten U.S. states have adopted the State Collaborative on Assessment and Student Standards (SCASS), which outlines specific standards for effective formative assessment practices. These standards include:

1. Learning Outcomes: Linking lesson topics to previous or future learning, and presenting clear and relevant learning outcomes to students.
2. Success Criteria: Involving students in developing success criteria that align with learning outcomes, ensuring they are clear, relevant, and appropriate for various student levels.
3. Extracting Evidence Through Tasks and Activities: Creating tasks aligned with learning outcomes, reviewing student work to gauge progress, and using responses to adjust teaching.
4. Extracting Evidence Through Questioning Strategies: Asking timely questions to assess progress, allowing sufficient response time, and using questions to gather learning evidence.
5. Feedback Loops During Questioning: Asking questions that promote student participation and building on responses.
6. Descriptive Feedback: Providing individualized descriptive feedback and opportunities for application.
7. Peer Assessment: Offering clear peer assessment opportunities, organizing activities, and using criteria that impact the quality of student work.
8. Self-Assessment: Providing opportunities for self-assessment, ensuring clarity of tasks, and using criteria to enhance student work quality and teaching development.
9. Collaboration: Encouraging teacher-student partnerships, student cooperation, and using feedback to enhance learning and confidence.
10. Using Evidence to Reshape Teaching: Collecting and analyzing learning evidence to adjust teaching practices.

The implementation of formative assessment standards aims to enhance student performance through personalized learning opportunities [27, 28]. Effective application of these standards depends on thorough teacher preparation, both pre-service and in-service, to develop expertise in formative assessment [29, 30]. Numerous field studies have shown that providing teachers with formative assessment skills leads to improvements in student performance [27-33]. This has prompted countries like the UK to emphasize teacher training in assessment, especially formative assessment, using frameworks such as the UK Training and Development Agency's progression from beginner to advanced levels [34].

Similarly, the Jordanian Ministry of Education emphasizes the importance of formative assessment within its national standards for science teacher preparation [35]. Core expectations include a deep understanding of assessment strategies and methods; the ability to connect learning outcomes, instruction, and assessment; designing and selecting diverse assessment tools, including technological ones, aligned with learning objectives; accurately documenting student progress and communicating with parents and school administration; engaging students in self-assessment and progress monitoring; and analyzing performance, providing feedback, and creating activities to foster growth based on assessment results.

The research gap in applying formative assessment standards in science education at schools in Jordan has attracted attention globally and locally due to the lack of studies in this area, despite efforts to enhance teacher training and implement these standards effectively in Jordanian classrooms to improve student learning outcomes. This emphasizes the importance of conducting research to recognize methods in use and obstacles encountered by educators in order to improve the implementation of formative assessment for superior educational outcomes.

2. Research Problem

In order to tackle the issues of the century effectively and adapt to the changing times, it's crucial to instill a robust appreciation for science among students, especially in subjects like science, mathematics, and technology. While education in science plays a role in shaping and advancing contemporary societies, the initiatives undertaken in Jordan starting from 2014 to transition towards a knowledge-driven economy and enhance practices through curriculum revisions and teacher training have not yielded the anticipated enhancements in student academic performance. The inadequacy is clear in the decreasing achievements of students from Jordan on scientific evaluations and the dropping pass rates on high school tests in science and math education.

These concerns bring up questions, about how effective science educators are in educational settings when it comes to utilizing formative assessment techniques and practices and leveraging assessment outcomes to boost student learning. Studies have consistently demonstrated that when formative assessment is applied correctly and efficiently it plays a role, in enhancing student achievements [5]. In the field of education research. Take a look, at this; various studies have shown how formative assessment can really make a difference [27-31, 36-38]. In addition, to the point made by Guhn, et al. [37] that effective formative assessment practices are closely connected to improved student performance in science. This relationship is underscored by assessments like the Trends, in International Mathematics and Science Study (TIMSS).

There is a pressing requirement to review the methods of assessment, especially formative assessment practices in science education within this framework, and grasp the difficulties that science teachers encounter when implementing these methods and practices—with regard to differences in teacher gender and teaching background levels. The primary goal of this research is to delve into these concerns by shedding light on the characteristics of formative assessment practices within science settings and evaluating how the effectiveness of these practices may be influenced by teacher gender and experience.

The research issue in this study is the difficulties in enhancing science education results in Jordan despite endeavors that may suggest deficiencies in executing formative assessment strategies and practices effectively. Hence, it is essential to delve into these strategies and practices and comprehend the impact of factors on their efficiency to offer practical suggestions for improving science education in Jordan. Consequently, the study explores the questions:

1. To what extent do science teachers utilize formative assessment practices to enhance children's primary science learning?
2. Does the use of formative assessment practices by science teachers to enhance children's primary science learning vary according to their gender (male or female) and teaching experience (less than 10 years or 10 years and more)?

3. Significance of the Study

The importance of this research lies in the evaluation of student learning in today's approaches that view assessment as a key element in improving learning standards. Through an emphasis on formative assessment, the study seeks to evaluate how science teachers' methods and practices affect student learning outcomes.

This study offers information, for policymakers at the Ministry of Education on how science teachers perform by using formative assessment techniques. These results may guide initiatives to enhance teacher education programs pre service and, in service to enhance assessment methods and boost student learning achievements in the run.

Additionally, the study offers standardized instruments that stakeholders can use to develop teachers' skills in assessing student learning. These instruments, based on internationally recognized standards, have the potential to be applied across various educational contexts, supporting the continuous improvement of educational practices.

4. Research Methodology

4.1. Study Design

The study utilized a descriptive-analytical approach because of its ability to offer a thorough grasp of the research issue by systematically gathering and analyzing data to describe current phenomena and identify patterns and trends, within the data. By utilizing this methodology, the study presented an accurate depiction of the subject matter while critically examining the underlying factors and correlations. This approach established a robust foundation for drawing meaningful conclusions and making well-informed recommendations.

4.2. Participants

30 teachers were chosen for a study in Irbid Governorate in Jordan during the semester of the 2023/2024 year. The group was evenly split between female and male teachers teaching science at the level. Among them were 13 teachers with less than 10 years of experience and 17 teachers with over a decade of experience. This sample was chosen due to the effective cooperation of the teachers, which facilitated accurate data collection. The sample size is deemed appropriate for the study, as it relies on classroom observation. It allows for the collection of balanced and comprehensive data on formative assessment practices, providing a detailed analysis of the effectiveness of these practices in various educational settings.

4.3. Instrument of the study

A standardized Classroom Observation Checklist, adapted from an instrument approved by ten American states Martinez, et al. [39] and Wylie and Lyon [40] was employed. This instrument was selected for its relevance to the study's goals and was modified to fit the specific context of this research. The instrument is structured as a verbal rating scale encompassing ten criteria, each associated with a set of measurable indicators. Each criterion is rated across four graduated performance levels, ranging from Beginning to Extending, which can be directly observed and assessed within the classroom environment.

The final version of the observation checklist includes 38 indicators distributed across ten criteria as follows:

1. Learning Outcomes Criterion: This criterion consists of three indicators focusing on formative assessment by linking the lesson topic with prior or future learning, presenting and formulating learning outcomes for students, and connecting these outcomes to the lesson topic.
2. Success Criterion: This includes five indicators that assess the quality of formative assessment practices. These indicators focus on student involvement in developing success criteria, alignment with learning outcomes,

appropriateness for different student levels, connection to all learning activities and tasks, and the use of suitable language.

3. Criterion for Extracting Evidence of Student Learning Through Tasks and Activities: This criterion addresses four indicators reflecting the quality of formative assessment through the relevance of tasks and activities to learning outcomes, clarity, teacher review of student work to assess progress, and the use of student responses to adapt teaching.
4. Criterion for Extracting Evidence of Student Learning Through Questioning Strategies: This criterion comprises four indicators focused on the teacher's performance in asking timely questions to assess progress, managing wait times, using questions to gather evidence of learning, and employing student responses to make inferences and adapt instruction.
5. Feedback Loops During Questioning Standard: This includes two indicators related to how the teacher encourages student participation through questioning and builds on student responses.
6. Criterion for Descriptive Feedback: This criterion includes two aspects related to providing personalized feedback and offering students opportunities to implement this feedback effectively.
7. Peer Assessment Criterion consists of five aspects related to assessment methods; these indicators encompass peer evaluation opportunities availability clarity, in assigned tasks influence on work quality task organization and criteria utilized by students, for peer assessment purposes.
8. Criterion for Self-Assessment: The information consists of five indicators designed to help students improve their self-assessment skills. These indicators focus on providing opportunities for self-assessment, the clarity of task organization, the influence on work quality, and the criteria employed by students for self-assessment.
9. Collaboration Standard: This includes five indicators assessing the teacher's performance in fostering classroom partnerships, promoting student cooperation, utilizing student responses and questions to deepen learning, exploring diverse viewpoints, and enhancing student confidence in their learning abilities.
10. Using Evidence to Reshape Teaching Standard: This consists of three indicators related to methods of collecting evidence about student learning, analyzing and reasoning with this evidence, and reshaping teaching processes accordingly.

To assess the quality of formative assessment practices using the verbal rating scale, each indicator is assessed across four performance levels:

1. Beginning Level: Represents low performance, where the teacher is classified as a "beginner." At this level, the observer assigns one point for the teacher's performance.
2. Growing Teacher Level: Indicates improvement over the beginning level, with the teacher recognized as a "growing teacher." At this level, the observer assigns two points for the teacher's performance.
3. Progressing Teacher Level: Reflects further improvement, where the teacher is identified as a "progressing teacher." At this level, the observer assigns three points for the teacher's performance.
4. At the tier of teaching performance is the Exemplary Teacher Level, where the teacher is recognized as an educator and awarded four points by the observer for their exceptional performance.

4.4. Validity and Reliability

The study ensured the validity of the observation checklist through a rigorous process. A group of eleven specialists from universities in Jordan, who have expertise in science curriculum development and assessment methods, carefully reviewed the items on the observation checklist and shared their thoughts on how well they matched the checklist categories. Taking into account their insights, these experts confirmed that the tool was suitable for the study's goals. The finalized study tool comprised 38 indicators.

To further enhance the observation checklist's validity and reliability, the researchers conducted a pilot test with 20 teachers and incorporated their feedback into the final version. Construct validity was verified through several measures, including Macdonald's Omega, Composite Reliability (CR), convergent validity, and discriminant validity. Macdonald's Omega and CR are frequently employed to evaluate the reliability of tools.

Table 1.
Selected indicators and construct validity coefficients.

Constructs	**Items**	**Factor loading**	**MacDonald's Omega**	**CR**	**AVE**
Learning outcomes	3	0.60-0.80	0.878	0.881	0.504
Success criterion	5	0.60-0.83	0.893	0.889	0.586
Extracting evidence of student learning through tasks and activities	4	0.60-0.85	0.935	0.936	0.596
Extracting evidence of student learning through questioning strategies	4	0.68-0.88	0.889	0.893	0.624
Feedback loops during questioning	2	0.76-0.90	0.948	0.951	0.701
Descriptive feedback	2	0.65-0.85	0.924	0.925	0.570
Peer evaluation	5	0.70-0.86	0.935	0.932	0.599
Self-assessment	5	0.62-0.82	0.891	0.887	0.585
Collaboration	5	0.67-0.86	0.898	0.895	0.593
Using evidence to reshape teaching	3	0.71-0.87	0.912	0.903	0.605

The produced results ranged from 0.881 to 0.951 and 0.878 to 0.948, respectively, which are above the suggested threshold of 0.700, indicating coherence. The average variance extracted (AVE) values spanned from 0.504 to 0.701, surpassing the halfway point of 50%. The square root of the AVE for each construct was greater than the correlations between that construct and others in the model, confirming the discriminant validity of the observation checklist using Fornell and Larcker's method. These results collectively demonstrate that the observation checklist is both reliable and valid [41, 42].

4.5. Data Collection

After obtaining consent from the participating science teachers, classroom observations were conducted, each lasting approximately 45 minutes per session. To ensure a comprehensive evaluation, a diverse range of topics was covered across the observed classes. For each teacher in the study sample, one observer completed an observation checklist, while a second observer independently filled out another checklist. Consequently, two separate observation checklists were completed for each teacher during each class session. This dual-observer approach was employed to enhance reliability and consistency in data collection, a practice emphasized in qualitative research methodologies [41]. To assess the degree of agreement between observers in completing the observation checklist, the "Intraclass Correlation Coefficient" (ICC) was calculated to assess the consistency of observers' evaluations of formative assessment practices. Additionally, Cohen's Kappa coefficient was utilized to determine the level of agreement between observers' assessments for each teaching-learning situation. The Kappa coefficient is particularly effective in assessing observer agreement as it adjusts for the likelihood of random agreement [43]. The resulting Kappa coefficient of 0.88 indicated a high level of consistency in data collection.

4.6. Data Analysis

The means and standard deviations of the observers' assessments of the quality of formative assessment practices by science teachers were calculated for each of the ten criteria and for each specific indicator associated with those criteria in the observation tool. The maximum possible score for each practice was 4, and the minimum was 1. These means were then converted into percentage averages for analysis.

To present and discuss the results related to the first research question, the following interpretive scale was employed: an arithmetic mean below 2 indicated a low practice rating, a mean between 2 and 2.99 indicated an average practice rating, and a mean between 3 and 4 indicated a high practice rating. To address the second research question, non-parametric tests were utilized, including the Mann-Whitney U test for pairwise comparisons.

5. Research Results

5.1. Results of the First Question

The first research question aims to determine how science teachers utilize formative assessment practices to enhance primary science learning for children. To address this, researchers employed an observation checklist during classroom visits. Observations were conducted in science classrooms with a sample of 30 male and female teachers. The observers recorded formative assessment practices across these various classroom settings. Each observer independently assessed the quality of these practices, guided by international standards outlined in the observation checklist.

The classroom observation utilized a verbal rating scale tool, which included ten criteria linked to specific performance indicators. Each indicator's quality was assessed across four performance levels:

1. Beginner: This level represents low performance, where the teacher is classified as a "Beginner." The observer assigns one point for performance at this level.
2. Growing Teacher: The second level indicates improvement from the first, where the teacher is recognized as a "Growing Teacher." At this level, the observer assigns two points.
3. Progressing Teacher: The third level reflects further improvement, where the teacher is identified as a "Progressing Teacher." Three points are assigned for performance at this level.
4. Extending Teacher: The fourth level represents the highest performance, where the teacher is deemed an "Extending Teacher." Four points are assigned for performance at this level.

Table 2.
Means and standard deviations of participants' quality scores for formative assessment practices.

Rank	Criteria No.	Criteria of formative assessment practices	Mean	St. Dev.	Percentile score	Rating
1	1	Learning outcomes	2.30	0.47	%57.5	Medium
2	5	Questioning feedback loops	1.98	0.31	%49.5	Low
3	4	Eliciting evidence of student learning through questioning strategies	1.97	0.53	%49.25	Low
4	3	Eliciting evidence of student learning through tasks and activities	1.96	0.66	%49.0	Low
5	9	Collaboration	1.95	0.49	%48.75	Low
6	6	Descriptive feedback	1.53	0.43	%38.25	Low
7	10	Using evidence to reshape teaching	1.46	0.32	%36.5	Low
8	7	Peer assessment	1.41	0.26	%35.25	Low
9	2	Success criteria	1.39	0.51	%34.75	Low
10	8	Self-assessment	1.38	0.23	%34.5	Low
Overall			1.73	0.42	%43.33	Low

Based on the classroom observations, the arithmetic means and standard deviations of the practice scores for each main criterion of formative assessment were calculated separately for each participant in the study sample. Additionally, the level of practice was assessed according to the established scale used in this study. The results are presented in Table 2 and Figure 1.

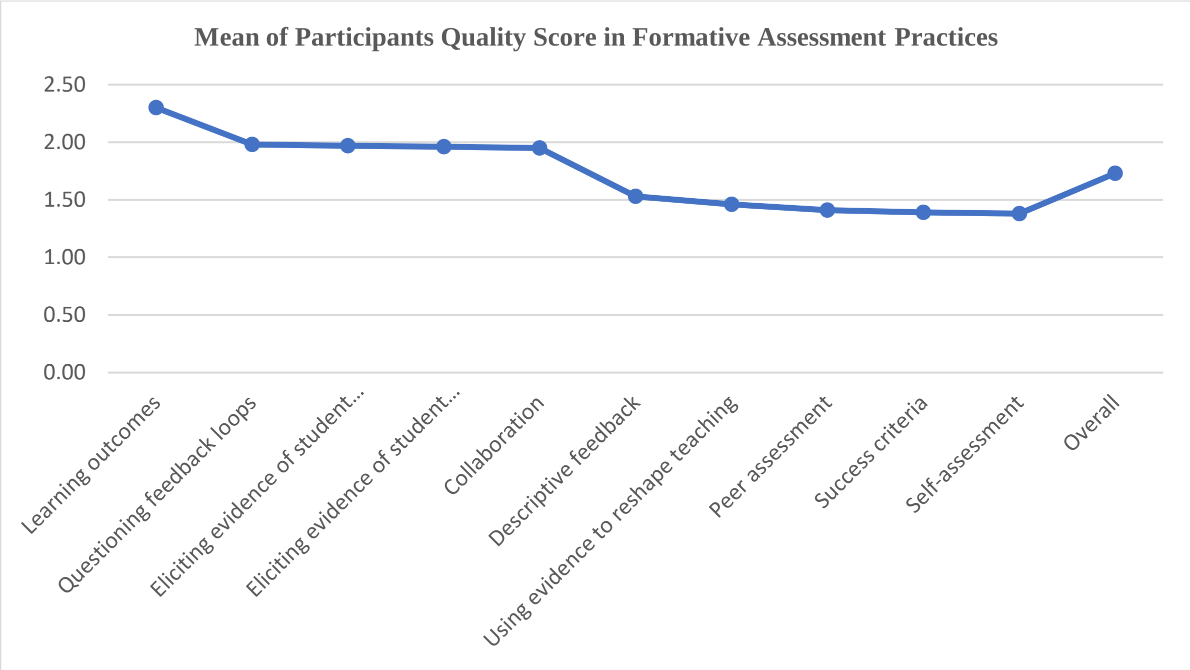

Figure 1.
Mean and Standard Deviation of Participants' Quality Scores in Formative Assessment Practices.

The observation results presented in Table 2 and Figure 1 indicate that the overall performance score for the quality of formative assessment practices in science learning environments was 1.73, equivalent to 43.33%. This score is considered low and categorizes the teachers' performance at the "Beginner" level, which is significantly below the minimum quality threshold adopted in this study—the "Progressing" level, with an acceptable score of 3.00. The data further reveal that the criterion for learning outcomes received the highest arithmetic mean within the second level, suggesting some improvement among teachers. However, the remaining criteria recorded closely clustered arithmetic means, ranging from 1.95 to 1.98. This indicates that the quality of formative assessment practices in science learning environments was consistently rated as "low" across all criteria and for the tool as a whole.

Table 2 also shows that the other criteria were distributed within a quality score below the average of 2.00, equivalent to less than 50%. This highlights a significant weakness in the quality of formative assessment practices in science learning environments. Ideally, these areas should achieve higher arithmetic means and reflect the highest levels of classroom practice. Effective formative assessment in science education depends on adherence to global standards, especially given their interconnected nature. The close alignment in the arithmetic means of these criteria suggests that students' learning of science, particularly assessment for learning, did not effectively engage with the core concepts and skills fundamental to science education. This deficiency poses a substantial issue for primary school science education, as foundational science learning at this stage is crucial for students' understanding of physics, chemistry, and biology in later grades.

The study extends beyond merely presenting the overall results for the observation checklist's fields by providing a detailed analysis of the arithmetic means and standard deviations for each indicator individually, as organized by the main criterion in which it was included. These findings are presented in Table 3.

Table 3 illustrates that all criteria for the quality of formative assessment practices scored below 50%, with the exception of the first criterion, "Learning Outcomes," which achieved a score of 57.5%. Additionally, the data reveals that eight out of the 38 performance indicators recorded percentage marks exceeding 50%. These indicators include:

- Linking the lesson topic with previous or future learning (Criterion 1), which exhibited an arithmetic mean of 2.71, corresponding to a percentage mark of 65.75%.
- Linking tasks and activities to student learning outcomes and Clarity of tasks and activities (Criterion 3), which recorded arithmetic means of 2.21 and 2.14, respectively, indicating some improvement among the study participants.
- Using questions to gather evidence about student learning and Asking questions to assess student progress (Criterion 4), which recorded percentage marks higher than 50%, with averages of 2.35 and 2.10, respectively.
- Asking questions to obtain responses from students (Criterion 5), which achieved an average of 2.25, translating to a percentage of 56.25%.
- Searching for different points of view among students (Criterion 9), which attained an average of 2.03, is equivalent to a percentage of 58.30%.

- Methods of collecting evidence about student learning in the classroom situation (Criterion 10), which recorded an average of 2.05, corresponding to 50.25%.

Table 3.
Arithmetic Means and Standard Deviations of Participants' Quality Scores for Formative Assessment Practices.

No.	Criteria	Indicators	Mean	St. Dev.	Percentile score	Rating
1	Learning outcomes	• Connecting the lesson topic to prior or future learning	2.71	0.53	65.75%	Medium
		• Presenting learning outcomes to students	2.22	0.51	47.50%	Low
		• Formulating and aligning learning outcomes with the lesson topic	1.96	0.58	43.75%	Low
2	Success criteria	• Engage students in the development of success criteria	1.98	0.31	25.00%	Low
		• Ensure alignment of success criteria with learning outcomes	1.99	0.31	25.00%	Low
		• Tailor success criteria to student proficiency levels	1.98	0.31	25.00%	Low
		• Integrate success criteria with learning activities and tasks	1.96	0.31	25.00%	Low
		• Appropriately formulate the language of success criteria	1.97	0.31	25.00%	Low
3	Extracting evidence of student learning through tasks and activities	• Align tasks and activities with student learning outcomes	2.21	0.55	55.00%	Medium
		• Ensure clarity of tasks and activities	2.14	0.71	53.75%	Medium
		• Conduct teacher reviews of student work	1.90	0.89	46.25%	Low
		• Utilize student responses to adjust teaching strategies	1.62	0.49	37.50%	Low
4	Extracting Evidence of Student Learning Through Questioning Strategies	• Pose questions to assess student progress	2.35	0.48	61.25%	Medium
		• Implement appropriate wait times	1.65	0.54	41.25%	Low
		• Utilize questions to collect evidence regarding student learning	2.10	0.50	51.25%	Medium
		• Leverage student responses to modify instruction	1.75	0.59	41.25%	Low
5	Feedback loops During Questioning	• Pose questions to elicit student responses	2.25	0.49	56.25%	Medium
		• Expand upon students' responses	1.65	0.48	44.35%	Low
6	Descriptive feedback	• Deliver descriptive feedback	1.85	0.48	44.35%	Low
		• Offer students opportunities to apply descriptive feedback	1.21	0.37	28.20%	Low
7	Peer assessment	• Facilitate opportunities for students to assess their peers	1.56	0.37	28.20%	Low
		• Ensure clarity in peer assessment tasks	1.56	0.37	28.20%	Low
		• Assess the impact of peer assessment on the quality of students' work	1.40	0.29	25.85%	Low
		• Organize peer assessment tasks effectively	1.40	0.29	25.85%	Low
		• Define the criteria used by students to assess their peers	1.40	0.29	25.85%	Low
8	Self-assessment	• Facilitate opportunities for students to engage in self-assessment of their learning	1.45	0.29	25.85%	Low
		• Ensure clarity in self-assessment tasks	1.45	0.29	25.85%	Low

No.	Criteria	Indicators	Mean	St. Dev.	Percentile score	Rating
		• Organize self-assessment tasks effectively	1.45	0.29	25.85%	Low
		• Assess the impact of self-assessment on the quality of students' work and the enhancement of teaching practices	1.37	0.14	24.80%	Low
		• Define the criteria utilized by students for self-assessment	1.34	0.29	25.85%	Low
9	Cooperation	• Foster a partnership between the teacher and students within the classroom environment	1.29	0.56	47.95%	Low
		• Encourage cooperation among students	1.18	0.35	30.95%	Low
		• Enable the teacher to utilize students' responses and questions to deepen understanding	1.27	0.59	44.80%	Low
		• Promote the exploration of diverse perspectives among students	2.03	0.57	58.30%	Medium
		• Enhance students' confidence in their learning capabilities	1.18	0.49	34.25%	Low
10	Utilizing Evidence to Redesign Instruction	• Methods for collecting evidence of student learning within the classroom context	2.05	0.68	50.25%	Medium
		• Analysis and reasoning of collected evidence	1.01	0.33	28.20%	Low
		• Redesigning instructional processes based on the evidence gathered	1.07	0.42	29.70%	Low

In light of the above, it can be concluded that, aside from the eight indicators showing arithmetic averages indicative of improvement in teacher performance, the remaining 30 indicators demonstrate a low growth level in the quality of formative assessment practices, as they fell below the 50% threshold. Furthermore, the analysis of Table 2 reveals that the quality of primary science teachers' performance in formative assessment practices did not attain the third level (three points) across all indicators. Consequently, the overall performance fails to meet the established quality standards.

5.2. Results of the Second Question

The second research question aims to determine whether there are statistically significant differences at the level of (p=0.05) n the quality of formative assessment practices in science learning environments attributed to the teacher's gender (male/female) and teaching experience (less than 10 years vs. 10 years or more). To address this, non-parametric tests, specifically the Mann-Whitney U test, were utilized. The results of this test are presented in Table 4.

Table 4.
Results of the Mann-Whitney U Test for Assessing the Quality of Formative Assessment Practices According to Gender and Teaching Experience.

Variable	Variable categories	Participants	Average Rank	W value (Sum of ranks)	U value	Sig.
Gender	Male	15	13.30	92.00	26.000	0.072
	Female	15	17.70	118.00		
Teaching Experience	Less than 10 years	13	17.84	101.00	28.000	0.134
	More than 10 years	17	13.91	109.00		

Table 4 indicates that the average rank for the quality of formative assessment practices in learning environments was 13.30 for male participants, whereas female participants had a higher average rank of 17.70. These results indicate that women teachers tend to apply assessment methods and practices more efficiently than male teachers do overall; although when running a Mann–Whitney U test analysis, a U value of 26.0 and a p value of 0.072 revealed that this variance is not statistically significant.

Participants who had less than 10 years of teaching experience averaged a rank of 17.84 for the quality of formative assessment practices compared to 13.91 for those with 10 or more years of experience. This indicates that teachers with less experience tend to be more involved in utilizing formative assessment methods. However, the Mann-Whitney U test results, with a U value of 28.00, also show no statistically significant difference in the average ranks of formative assessment practices based on teaching experience.

6. Discussion

6.1. Discussion of the Results of the First Question

The analysis of the quality of formative assessment practices in science education reveals several critical deficiencies:

1. Deficiencies in Initial Teaching Practices: The study highlights a significant inadequacy in leveraging formative assessment to connect prior learning experiences with new content. Students showed a lack of clarity about expected learning outcomes, which were not well aligned with the cognitive demands of the science curriculum. This disconnect indicates that formative assessments are not being used effectively to build on students' existing knowledge.
2. Lack of Assessment Quality Standards: The absence of clear assessment quality standards or "success criteria" impairs the ability to measure student performance against established learning outcomes. The study found that performance in this criterion was particularly poor, with some indicators failing to exceed a 25% threshold. This suggests a widespread lack of awareness among teachers regarding the importance of setting and communicating clear success criteria to assess student learning effectively.
3. Weakness in Evidence Extraction: The quality of formative assessment practices in extracting evidence of student learning through tasks and activities was found to be insufficient. Issues included misalignment between tasks and learning outcomes and a lack of clarity in task instructions. Teachers also demonstrated inadequate performance in reviewing student work and using this feedback to inform instructional adjustments.
4. Insufficient Use of Questioning Strategies: Formative assessment practices based on questioning strategies were not effectively employed to assess student progress. Problems such as inadequate wait times for responses and a lack of evidence collection about student learning were common. Moreover, there was a lack of types of feedback (such as informative, supportive, corrective, and interpretive) that play a role in encouraging deeper learning and improving student comprehension.
5. There is an absence of peer assessment opportunities, as the research showed a lack of formative assessment practices in connection with peer assessment. Students had limited or poorly defined opportunities to assess each other’s work. They were not provided with opportunities for peer assessment during the observations made; this hinders the development of critical thinking and collaborative skills, which are crucial for their growth.
6. Lack of Self-Assessment Standards: There was a lack of guidelines and standards for quality self-assessment practices in the setting that hindered students' ability to effectively evaluate their own learning progress and achieve positive outcomes.
7. Lack of Collaboration in Education Settings: A research study highlighted the cooperation and teamwork between educators and students in various settings. There was an oversight in recognizing and tackling the learning obstacles that students encountered. This affected teachers’ capacity to address these issues adeptly. Observations revealed a lack of support for student queries, hindering the exploration of varying viewpoints and adversely affecting students’ confidence in their learning skills.
8. Insufficient attention to learning challenges; the research also pointed out a lack of assessment methods to help students with learning difficulties in science subjects. There was a shortage of efforts in gathering and using data on student learning, which is crucial for customizing teaching methods to support students and help them tackle their learning obstacles.

The results underscore areas of focus regarding the use of assessment methods and practices in science education settings. They indicate a need for substantial improvement and restructuring to enhance the quality of teaching and learning outcomes.

Based on these discoveries, it seems that the outcomes could be linked to a lack of understanding among the individuals involved in the study regarding their knowledge of how formative assessment practices work in education. This lack of knowledge might stem from training initiatives provided to science teachers that fail to equip them with the necessary competencies for successfully incorporating formative assessment strategies into their teaching methods. Consequently, this outcome underscores a critical issue: the absence of a personal philosophy among science teachers grounded in contemporary educational philosophies.

The lack of a coherent professional vision and mission among teachers suggests that they do not embody the fundamental values essential for diverse teaching and learning contexts. Furthermore, they lack clearly defined professional goals and may not possess a comprehensive understanding of their roles, as well as the roles of learners, within the framework of contemporary educational thought. This interpretation aligns with Amari [44] study, which indicated that science teachers often lack a well-defined philosophy of science education, tending instead to view assessment primarily as a means of assigning grades and measuring students' memorized knowledge.

Moreover, it is evident that possessing theoretical knowledge about formative assessment in science education, without practical application, does not equip teachers with the ability to effectively utilize these assessments in real learning environments. Existing teacher preparation programs for science educators are not sufficiently integrated and fail to emphasize the use of assessment for learning; instead, they focus on assessing the extent of students' information retention through rote memorization. Many science teachers believe that traditional paper-and-pencil tests represent the most effective method for assessing student learning in science, given the subject’s reliance on mathematical problem-solving.

This weakness in the preparation of science teachers highlights the absence of training programs grounded in sound educational principles that promote the effective implementation of formative assessment practices. Ideally, these practices should become ingrained behaviors within science learning environments. Supporting this assertion, numerous teachers have conveyed to researchers during informal discussions that they do not engage in performance reviews or develop self-improvement plans. Many also reported lacking advanced educational qualifications and exhibiting low motivation to attend workshops and training sessions related to effective teaching, particularly in formative assessment. Additionally, they expressed limited awareness of contemporary educational developments and successful practices implemented by their colleagues.

The results from the first research question further illustrate the insufficient attention given by authors of science textbooks to the incorporation of formative assessment strategies. This lack of emphasis is evident in the behavior of primary science teachers, who often adhere strictly to the prescribed content of these textbooks. Classroom observations revealed that science teachers frequently follow the assessment procedures outlined in these texts in a literal and step-by-step manner, underscoring the need for a more flexible and innovative approach to formative assessment in science education.

Given the prevalent incorrect assessment practices, it can be asserted that students' learning of science fails to uphold the principles of effective learning, particularly in fostering suitable environments for the exploration and discovery of scientific ideas. This shortfall represents a significant obstacle to students' ability to construct and comprehend scientific concepts and skills, understand their implications, and apply them in real-life situations. Additionally, these flawed practices, which have adversely impacted formative assessment, have deprived students of the opportunity to engage in individual learning tasks, cultivate independent learning, and ensure their grasp of these tasks. As a result, this has negatively influenced students' development of self-confidence and responsibility for their own learning, especially considering that success in primary science heavily relies on the extent to which teachers provide opportunities for self-directed learning.

Furthermore, the results indicate that the lack of assessment through exercises and learning tasks as indicators of students' understanding of science contradicts prevailing global trends in science education. These trends emphasize assessing a student's level in science by observing their performance in learning tasks and activities. The failure to align with these practices may contribute to students' inability to develop relational understanding, encourage critical thinking, and link conceptual knowledge with procedural skills. This outcome is also at odds with the Ministry of Education's initiatives aimed at advancing education toward a knowledge-based economy, which stresses the importance of integrating formative assessment in educational environments. This approach is intended to empower students to reflect on their performance in authentic learning contexts and engage in meaningful tasks that promote higher-order thinking skills (Ministry of Education, 2005).

The study's results also highlighted that formative assessment practices failed to sufficiently prioritize the role of peer assessment and its importance in fostering students' learning of science. This approach stands in contrast to constructivist theories, which assert that while learning is fundamentally a self-directed process, it also requires students to build knowledge through interaction with others during the completion of learning tasks. Such interactions enable students to plan, direct, and manage their learning effectively. This outcome contradicts the findings of field studies, which have emphasized the role of formative assessment in helping students develop critical thinking skills, access and construct knowledge, assess and critique information, and apply it in various real-world contexts [19, 36, 45-48].

Given the above, the poor quality of formative assessment practices strongly indicates a deficiency in teachers' ability to fulfill their roles in addressing and overcoming challenges in students' science learning. This inadequacy negatively impacts students' understanding of science, particularly as ignoring students' errors and the difficulties they encounter is one of the most significant factors impeding their learning progress. These shortcomings in formative assessment practices can likely be attributed to science teachers' limited knowledge of how to address and correct physical and chemical misconceptions. The researchers observed that study participants often did not attempt to correct such errors, either before or after lessons, nor did they consider providing students with opportunities to reflect on and correct their misunderstandings. This weakness is directly tied to the dominance of traditional instructional approaches among the participants, which emphasize the delivery of knowledge without fostering a deeper understanding or engagement with the material. Such practices are at odds with constructivist thought, which views assessment as a tool for facilitating learning through understanding rather than simply recalling information.

These findings suggest that the lack of exposure to effective error-diagnosis techniques in the academic training of science teachers—particularly concerning disciplines like physics and chemistry—has significantly contributed to the inadequacies in their formative assessment practices. This gap has resulted in a failure to address and correct student errors, undermining the primary role of formative assessment in enhancing the quality of students' learning outcomes in science.

The current study's findings align with existing educational research and international reports, which have similarly highlighted that poor student performance in science is closely linked to deficiencies in formative assessment practices. These practices, which are not learner-centered—a crucial element in the teaching-learning process—were found lacking across all components of the educational experience. This is particularly concerning given that student assessment is a key component of effective teaching and learning [19, 26, 36, 49]. The results are consistent with those of Bouchaib [50] who indicated that teachers had limited abilities in employing various assessment methods and tools.

However, the findings of the current study diverge from those reported in previous research. For instance, Dayal [51] observed that teachers effectively utilized a variety of assessment strategies in their instruction. Similarly, studies by Pang [52] and AlAli [19] documented an increase in the adoption of assessment for learning practices among teachers. These contrasting outcomes highlight the variability in the implementation and effectiveness of formative assessment practices across different educational contexts, suggesting that the success of these practices may depend heavily on the specific conditions and approaches employed within individual classrooms.

6.2. Discussion of the Results of the Second Question

The analysis of the data revealed no significant effect of teacher gender on the quality of formative assessment practices, as indicated by the Mann-Whitney test results. This outcome can be attributed to the Ministry of Education's emphasis on aligning learning practices with the knowledge economy in mathematics education rather than focusing on effective formative assessment practices. Consequently, both male and female science teachers receive similar academic and educational preparation, leading to comparable formative assessment practices across genders.

Despite educational reforms in the Jordanian system, there has been no notable improvement in the quality of formative assessment practices among science teachers, regardless of gender. Informal discussions with science teachers suggest that this stagnation may be due to the lack of targeted training programs aimed at enhancing formative assessment skills and the limited incorporation of contemporary assessment trends in educational supervision. This observation aligns with the findings of Obeido [53] and Basilio and Bueno [54] who noted that educational supervision has not effectively facilitated the adoption of modern assessment concepts among science teachers.

Additionally, the study found no significant impact of teaching experience on formative assessment practices. Several factors may explain this outcome, including the generally weak educational qualifications of science teachers across experience levels, the superficial nature of workshops and training courses provided by the Ministry of Education (often limited to new teachers), and a decline in teaching motivation with increased experience. As a result, science teachers, irrespective of their experience, tend to implement formative assessment in similar ways.

This finding is consistent with studies by Moyo, et al. [55] and Jett [56] which reported no statistically significant differences between teacher experience and the degree of formative assessment use. However, it contrasts with research by Van Der Steen, et al. [26] and Ryan [57] which found that formative assessment practices tend to improve with more teaching experience. This discrepancy may arise from differences in data collection methods; the current study employed classroom observations with a verbal rating scale, whereas other studies used questionnaires.

7. Conclusions and Implications

The educational landscape in the century has seen notable changes in the field of science education related to assessment practices that prioritize learning outcomes through formative assessments to improve overall learning quality. Many countries have shifted from using assessments primarily for accountability and grading to employing them as tools for fostering student development in science education.

The results of this study reveal a critical issue: the alignment between instructional tasks and students' learning outcomes was inadequately implemented. This misalignment led to ineffective use of student responses in assessing learning progress and resulted in assessment practices that failed to engage students in activities that build upon their prior experiences to acquire new knowledge. Science teachers often do not collect sufficient information about students' understanding of scientific concepts, which prevents them from addressing misconceptions crucial for deepening both conceptual and procedural knowledge. This issue arises from a lack of clarity regarding formative assessment among teachers, leading to a focus on rote memorization rather than conceptual understanding. Such challenges hinder broader educational reform efforts in Jordan, which aim to utilize formative assessment as a tool for building and applying knowledge in real-life contexts. Furthermore, it can be concluded that the fundamental philosophy of formative assessment remains unclear to science teachers, regardless of their gender or experience, leading to ineffective implementation and diminished potential for enhancing student learning.

To address these challenges, a multifaceted approach is essential. Implementing effective training workshops for science teachers, supported by ongoing classroom support, is crucial to ensure the practical application of acquired knowledge and skills. Trainers should be selected based on their expertise to guarantee high-quality professional development. Additionally, school leaders and educational supervisors should be empowered to conduct regular and continuous assessments of teachers to monitor and support their progress. Concurrently, the curriculum must be revised to ensure alignment with formative assessment requirements across all its components. The Ministry should develop standardized assessment tools that include quality standards and key performance indicators for formative assessment, potentially using the classroom observation tool from this study as a model.

Creating a learning environment requires addressing issues like student conduct problems and overcrowded classrooms while ensuring access to educational materials. Integrating formative assessment development programs and projects into the Ministry's strategic and executive plans will support these efforts.

However, this study encounters several limitations. The relatively small sample size may affect the generalizability of the results to the broader population of science teachers. The focus solely on formative assessment practices may limit understanding of interactions with other types of assessments. Furthermore, there is a challenge in comprehending the development of formative assessment practices over time due to the absence of long-term data. Moreover, the potential observer bias introduced by the aspect of classroom observations could impact result reliability. Additionally, the study focused on a subset of science teachers and research tools, potentially limiting the generalizability of the findings to educational settings or geographical areas.

We should look into these limitations in studies by using more varied samples to improve the applicability of the results. Studying changes over time through research could help us understand how formative assessment methods evolve with advancements and changes. Broadening the scope of research to cover a range of assessment types and subject areas will provide insight into assessment practices.

Furthermore, utilizing a variety of data-gathering approaches, such as formal interviews, could potentially reduce any observer bias and delve into the nuances of formative assessment techniques and practices within different educational environments. By delving into educators' perspectives and attitudes towards formative assessment across situations, we can gain a deeper insight into how it influences student academic progress. Furthermore, future research should focus on enhancing educational materials with activities that promote self-assessment and peer assessment skills within science learning contexts.

References

[1] S. Kim, M. Raza, and E. Seidman, "Improving 21st-century teaching skills: The key to effective 21st-century learners," *Research in Comparative and International Education,* vol. 14, no. 1, pp. 99-117, 2019. https://doi.org/10.1177/1745499919832895

[2] K. Neumann, A. Kauertz, and H. E. Fischer, "Quality of instruction in science education," *Second International Handbook of Science Education,* pp. 247-258, 2012.

[3] A. Duraiappah *et al.*, "The international science and evidence-based education assessment," *npj Science of Learning,* vol. 6, no. 1, p. 7, 2021. https://doi.org/10.1038/s41539-021-00019-9

[4] I. Y. C. Menéndez, M. A. C. Napa, M. L. M. Moreira, and G. G. V. Zambrano, "The importance of formative assessment in the learning teaching process," *International Journal of Social Sciences and Humanities,* vol. 3, no. 2, pp. 238-249, 2019.

[5] K. Schildkamp, F. M. van der Kleij, M. C. Heitink, W. B. Kippers, and B. P. Veldkamp, "Formative assessment: A systematic review of critical teacher prerequisites for classroom practice," *International Journal of Educational Research,* vol. 103, p. 101602, 2020. https://doi.org/10.1016/j.ijer.2020.101602

[6] K. Kulasegaram and P. K. Rangachari, "Beyond "formative": Assessments to enrich student learning," *Advances in Physiology Education,* vol. 42, no. 1, pp. 5-14, 2018. https://doi.org/10.1152/advan.00114.2017

[7] H. Tairab and S. Belbase, "Current trends in science curriculum reforms in response to stem education: international trends, policies and challenges in reforming science teacher education programs in the stem era: International and comparative perspectives." Cham: Springer International Publishing, 2023, pp. 265-281.

[8] R. M. Pringle and R. M. Pringle, "Reforms in science education: A response to changing societal contexts," *Researching Practitioner Inquiry as Professional Development: Voices from the Field of Science Teaching,* pp. 9-22, 2020.

[9] J. Stanja, W. Gritz, J. Krugel, A. Hoppe, and S. Dannemann, "Formative assessment strategies for students' conceptions—The potential of learning analytics," *British Journal of Educational Technology,* vol. 54, no. 1, pp. 58-75, 2023. https://doi.org/10.1111/bjet.13274

[10] L. S. Huang, R. Khatri, and A. Alhemaid, "Enhancing learning through formative assessment: evidence-based strategies for implementing learner reflection in higher education. in global perspectives on higher education: From crisis to opportunity." Cham: Springer International Publishing, 2023, pp. 59-73.

[11] S. Narciss and J. Zumbach, "Formative assessment and feedback strategies in international handbook of psychology learning and teaching ". Cham: Springer International Publishing, 2022, pp. 1-28.

[12] C. Ozan and R. Y. Kincal, "The effects of formative assessment on academic achievement, attitudes toward the lesson, and self-regulation skills," *Educational Sciences: Theory and Practice,* vol. 18, no. 1, pp. 85-118, 2018. https://doi.org/10.12738/estp.2018.1.0132

[13] M. A. Adarkwah, "The power of assessment feedback in teaching and learning: A narrative review and synthesis of the literature," *SN Social Sciences,* vol. 1, no. 3, p. 75, 2021. https://doi.org/10.1007/s43545-021-00073-7

[14] R. Ajjawi, J. Tai, T. L. Huu Nghia, D. Boud, L. Johnson, and C.-J. Patrick, "Aligning assessment with the needs of work-integrated learning: The challenges of authentic assessment in a complex context," *Assessment & Evaluation in Higher Education,* vol. 45, no. 2, pp. 304-316, 2020. https://doi.org/10.1080/02602938.2019.1563789

[15] A. Hargreaves and M. Fullan, *Professional capital: Transforming teaching in every school*. New York: Teachers College Press, 2012.

[16] J. Hattie and G. Yates, *Visible learning and the science of how we learn*. New York: Routledge, 2014.

[17] J. Hiebert and A. K. Morris, "Teaching, rather than teachers, as a path toward improving classroom instruction," *Journal of teacher Education,* vol. 63, no. 2, pp. 92-102, 2012. https://doi.org/10.1177/0022487111433643

[18] S. Meng, "Enhancing teaching and learning: Aligning instructional practices with education quality standards," *Research and Advances in Education,* vol. 2, no. 7, pp. 17-31, 2023.

[19] R. M. AlAli, "Assessment for learning at Saudi Universities: An analytical study of actual practices," *J. Institutional Res. South East Asia,* vol. 19, no. 1, pp. 20-41, 2021. https://doi.org/10.53586/jirsea.2021.v19i1.3

[20] S. Chand and K. Pillay, "Understanding the fundamental differences between formative and summative assessment," *Global Scientific and Academic Research Journal of Education and literature,* vol. 2, no. 2, pp. 6-9, 2024.

[21] A. Kanjee and J. Bhana, *Using formative assessment to improve learning and teaching: Practical guidelines for teacher during the COVID-19 pandemic*. Oxford University Press. https://doi.org/10.13140/RG.2.2.17255.47528, 2020.

[22] K. Tobler, "Keeley probes as a tool for uncovering student ideas: How do teachers use formative assessment probes to plan and adapt instruction?," Master's Thesis, Portland State University, 2015.

[23] H. Swanson and J. Clarke-Midura, "Integrating formative assessment and feedback into scientific theory-building practices and instruction," *Assessment in Education: Principles, Policy & Practice,* vol. 28, no. 2, pp. 118-134, 2021. https://doi.org/10.1080/0969594X.2021.1875670

[24] D. Hodgson and L. Watts, "Four principles of good assessment practice: A teaching and learning approach to designing and assessing student work," *Advances in Social Work and Welfare Education,* vol. 18, no. 1, pp. 8-22, 2016.

[25] J. Chemeli, *Impact of the five key formative assessment strategies on learner's achievement in mathematics instruction in secondary schools*. Kenya: A Case of Nandi County, 2019.

[26] J. Van Der Steen, T. van Schilt-Mol, C. Van der Vleuten, and D. Joosten-ten Brinke, "Designing formative assessment that improves teaching and learning: What can be learned from the design stories of experienced teachers?," *Journal of Formative Design in Learning,* vol. 7, no. 2, pp. 182-194, 2023. https://doi.org/10.1007/s41686-023-00106-9

[27] A. C. Alonzo, "An argument for formative assessment with science learning progressions," *Applied Measurement in Education,* vol. 31, no. 2, pp. 104-112, 2018. https://doi.org/10.1080/08957347.2017.1408630

[28] P. Atjonen, S. Pöntinen, S. Kontkanen, and P. Ruotsalainen, "In enhancing preservice teachers' assessment literacy: Focus on knowledge base, conceptions of assessment, and teacher learning," *Frontiers in Education,* vol. 7, p. 891391, 2022. https://doi.org/10.3389/feduc.2022.891391

[29] P. Atjonen, S. Kontkanen, P. Ruotsalainen, and S. Pöntinen, "Pre-service teachers as learners of formative assessment in teaching practice," *European Journal of Teacher Education*, vol. 47, no. 2, pp. 267-284, 2024. https://doi.org/10.1080/02619768.2024.2338840

[30] G. L. Black, "Say cheese! A snapshot of elementary teachers' engagement and motivation for classroom assessment," *Action in Teacher Education*, vol. 36, no. 5-6, pp. 377-388, 2014. https://doi.org/10.1080/01626620.2014.977689

[31] G. Conderman, E. Pinter, and N. Young, "Formative assessment methods for middle level classrooms," *The Clearing House: A Journal of Educational Strategies, Issues and Ideas*, vol. 93, no. 5, pp. 233-240, 2020. Strategies, Issues & Ideas, 93(5), 233–240. https://doi.org/10.1080/00098655.2020.1778615

[32] C. DeLuca, T. Chavez, A. Bellara, and C. Cao, "Pedagogies for preservice assessment education: Supporting teacher candidates' assessment literacy development," *The Teacher Educator*, vol. 48, no. 2, pp. 128-142, 2013. https://doi.org/10.1080/08878730.2012.760024

[33] J. P. Leighton, "Students' interpretation of formative assessment feedback: Three claims for why we know so little about something so important," *Journal of Educational Measurement*, vol. 56, no. 4, pp. 793-814, 2019. https://doi.org/10.1111/jedm.12237

[34] TDA, "Training and development agency for schools annual report and accounts for 2010−2011," Retrieved: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/229179/1071.pdf, 2011.

[35] Ministry of Education, *Science curriculum guidelines for the elementary stage*. Amman, Jordan: Ministry of Education, 2024.

[36] V. Atasoy and G. Kaya, "Formative assessment practices in science education: A meta-synthesis study," *Studies in Educational Evaluation*, vol. 75, p. 101186, 2022. https://doi.org/10.1016/j.stueduc.2022.101186

[37] M. Guhn, A. Gadermann, and A. D. Wu, "Trends in international mathematics and science study (TIMSS). In encyclopedia of quality of life and well-being research." Cham: Springer International Publishing, 2024, pp. 7309-7311.

[38] R. L. Staberg, M. I. M. Febri, Ø. Gjøvik, S. A. Sikko, and B. Pepin, "Science teachers' interactions with resources for formative assessment purposes," *Educational Assessment, Evaluation and Accountability*, vol. 35, no. 1, pp. 5-35, 2023. https://doi.org/10.1007/s11092-022-09347-3

[39] F. Martinez, S. Taut, and K. Schaaf, "Classroom observation for evaluating and improving teaching: An international perspective," *Studies in Educational Evaluation*, vol. 49, pp. 15-29, 2016. https://doi.org/10.1016/j.stueduc.2016.07.003

[40] C. Wylie and C. Lyon, "Using the formative assessment rubrics, reflection and observation tools to support professional reflection on practice," *Formative Assessment for Teachers and Students (FAST) State Collaborative on Assessment and Student Standards (SCASS) of the Council of Chief State School Officers*, 2013.

[41] R. AlAli and A. Al-Barakat, "Young children's attitudes toward science learning in early learning grades," *Asian Education and Development Studies*, vol. 13, no. 4, pp. 340-355, 2024. https://doi.org/10.1108/AEDS-02-2024-0036

[42] R. AlAli and S. Saleh, "Towards constructing and developing a self-efficacy scale for distance learning and verifying the psychometric properties," *Sustainability*, vol. 14, no. 20, p. 13212, 2022. https://doi.org/10.3390/su142013212

[43] S. R. Walter, W. T. Dunsmuir, and J. I. Westbrook, "Inter-observer agreement and reliability assessment for observational studies of clinical work," *Journal of Biomedical Informatics*, vol. 100, p. 103317, 2019. https://doi.org/10.1016/j.jbi.2019.103317

[44] S. I. Amari, "Information geometry and its applications," vol. 194: Springer, 2016.

[45] D. Alt, L. Naamati-Schneider, and D. J. Weishut, "Competency-based learning and formative assessment feedback as precursors of college students' soft skills acquisition," *Studies in Higher Education*, vol. 48, no. 12, pp. 1901-1917, 2023. https://doi.org/10.1080/03075079.2022.2130403

[46] X. Xu, W. Shen, A. Islam, and Y. Zhou, "A whole learning process-oriented formative assessment framework to cultivate complex skills," *Humanities and Social Sciences Communications*, vol. 10, no. 1, pp. 1-15, 2023. https://doi.org/10.1057/s41599-023-01542-7

[47] Z. Yan, M. M. Chiu, and E. C. K. Cheng, "Predicting teachers' formative assessment practices: Teacher personal and contextual factors," *Teaching and Teacher Education*, vol. 114, p. 103718, 2022. https://doi.org/10.1016/j.tate.2022.103718

[48] W. Zeng, F. Huang, L. Yu, and S. Chen, "Towards a learning-oriented assessment to improve students' learning—a critical review of literature," *Educational Assessment, Evaluation and Accountability*, vol. 30, pp. 211-250, 2018. https://doi.org/10.1007/s11092-018-9281-6

[49] N. M. Dah, M. S. A. M. Noor, M. Z. Kamarudin, and S. S. S. A. Azziz, "The impacts of open inquiry on students' learning in science: A systematic literature review," *Educational Research Review*, p. 100601, 2024. https://doi.org/10.1016/j.edurev.2024.100601

[50] B. Bouchaib, "Exploring teachers' assessment practices and skills," *International Journal of Assessment Tools in Education*, vol. 4, no. 1, pp. 1-18, 2016. https://doi.org/10.21449/ijate.283029

[51] H. C. Dayal, "How teachers use formative assessment strategies during teaching: Evidence from the classroom," *Australian Journal of Teacher Education (Online)*, vol. 46, no. 7, pp. 1-21, 2021. https://doi.org/10.14221/ajte.2021v46n7.1

[52] N. S.-K. Pang, "Teachers' reflective practices in implementing assessment for learning skills in classroom teaching," *ECNU Review of Education*, vol. 5, no. 3, pp. 470-490, 2022. https://doi.org/10.1177/20965311221114524

[53] M. Obeido, "Science teachers' beliefs about employing alternative assessment in teaching basic stage students laboratory skills," Unpublished PhD Thesis, Yarmouk University, Irbid, Jordan, 2016.

[54] M. B. Basilio and D. C. Bueno, "Instructional supervision and assessment in the 21st century and beyond," *Online Submission*, vol. 4, pp. 1-8, 2021.

[55] S. E. Moyo, C. Combrinck, and S. Van Staden, "Evaluating the impact of formative assessment intervention and experiences of the standard 4 teachers in teaching higher-order-thinking skills in mathematics in frontiers in education," vol. 7: Frontiers Media SA, 2022, p. 771437.

[56] P. M. Jett, *Teachers valuation and implementation of formative assessment strategies in elementary science classrooms*. University of Louisville, 2009.

[57] A. Ryan, "Assessment practices for learning among primary school mathematics teachers in Hebron government schools from their point of view," *Journal of the Islamic University for Educational and Psychological Studies*, vol. 23, no. 1, pp. 300-272, 2015.