ROSSIISKAYA ISTORIYA
ISSN: 0869-5687
No. 1 (2026)
https://rossiiskaya.com

Difficulties Facing Secondary Education Teachers in Implementing New Evaluation Strategies: A Field Study at Sheikh Omar Al-Mukhtar Secondary School in M'Sila

Souad Arioua
Mohamed Boudiaf University of M'Sila
souad.arioua@univ-msila.dz
Khadra Hallab
Mohamed Boudiaf University of M'Sila
khadra.hallab@univ-msila.dz
Abdellaziz Nekbil
Laboratory of Pedagogical Issues and Related Matters, Higher Teacher Training School (ENS) Bousaada
Nekbil.abdellaziz@ens-bousaada.dz

Received: 03/02/2026 Accepted: 07/04/2026 Published: 25/06/2026

Study Abstract
The current study aimed to identify the difficulties facing secondary education teachers in implementing new evaluation strategies through a field study conducted in M'Sila. To achieve this objective, a descriptive methodology was adopted. A purposive sample of 56 male and female teachers was selected from Sheikh Omar Al-Mukhtar Secondary School in the Province of M'Sila. A questionnaire was constructed by the researchers, and its psychometric properties were verified after its administration to a pilot sample of 25 male and female teachers. Following the administration of the questionnaire to the primary sample, a set of statistical methods was utilized to test and validate the hypotheses. Ultimately, the study addressed the following questions and reached the subsequent results:
1_ What are the difficulties facing teachers in utilizing new evaluation strategies?
2– Do these difficulties vary according to the study variables (gender, educational qualification, experience, and specialization)?
Keywords: Difficulties, educational evaluation strategies, secondary education teachers.

Research Problem and Questions

The educational field is highly sensitive, serving as the core pillar of society and the foundation upon which other economic, social, cultural, and political sectors are constructed. Consequently, educational objectives remain a paramount concern occupying the attention of scientists, educators, and thinkers. Continuous research is underway to develop and improve the pedagogical-learning process, aiming to elevate the learner to the highest academic standards and foster significant holistic growth across various dimensions of personality, thereby ensuring sustainable development.

With the rapid advancements characterizing contemporary societies, there is an urgent need to focus on the status of education and the dilemmas arising from its interaction with the modern reality of learners, their capacities, available resources, guiding policies, and the procedures implemented to fulfill them. In this context, global educational systems face profound challenges due to accelerated developments. Consequently, developing nations have begun initiating a comprehensive overhaul of their educational frameworks, particularly because traditional education in its current configuration no longer fulfills the needs and aspirations of learners. This has prompted the search for novel approaches and modern strategies that contribute to educational advancement and align with developed nations and contemporary international trends (Al-Mana'sah, 2020, p. 2).

Furthermore, the instructional process is viewed as a science governed by established foundations and rules that facilitate the understanding, interpretation, and prediction of phenomena within the learning environment. This serves as a prerequisite for managing instructional dynamics and directing them optimally. Teaching is also considered both an art and an educational system featuring distinct inputs and outputs, whose optimization depends heavily on investigating the impacts of instructional variables. The success of this process relies fundamentally on the competence, skills, and professional inclination of the teacher, which collectively stimulate student learning and prepare them to acquire appropriate educational experiences. The teacher represents the focal axis and primary cornerstone for the success of the educational process and the attainment of desired outcomes. The teacher's role extends beyond lesson planning, selecting instructional media, and lecturing, to encompass comprehensive planning, organization, implementation, and evaluation (Sabah, 2022, p. 6).

The instructional process is inextricably linked to evaluation; instruction without evaluation remains an unmeasurable activity whose efficacy and objective alignment cannot be verified. Evaluation is a foundational component of the curriculum, exerting a direct influence on its various dimensions, including educational objectives, content, and teaching methodologies. It is implemented concurrently with planning processes—such as lesson preparation—and executive operations, such as content delivery and testing. Furthermore, evaluation is practiced continuously and cumulatively by assessing instruction at various intervals, both during and after teaching. This process is not restricted to rendering judgments on academic achievement, but extends to remediating observed deficiencies to overcome learning barriers. Educational evaluation occupies a position of distinct significance in any curriculum due to its role in enhancing the quality of teaching and learning. Numerous educators, including Burton (2000) and Hanes (1997), have affirmed a robust correlation between evaluation methods, instructional strategies employed by teachers, and learning approaches adopted by students. Moreover, West (1993) noted that teaching methods, regardless of their efficacy or learner-centered design, cannot fulfill their purposes unless continuous and sufficient data regarding the learner’s progress or stagnation are provided—a requirement that cannot be met without a functional evaluation system and its associated strategies (Amziane, 2026, pp. 406–407).

Evaluation stands as a pillar of the pedagogical system, driving continuous quality enhancement across various educational tiers. It aims to prepare learners to adapt to and interact with scientific advancements across diverse life spheres, enabling them to contribute to development

within an era characterized by technological acceleration and knowledge explosion. This cannot be realized without refining effective learning methods, upgrading teachers' professional competencies, and adopting robust student assessment practices aligned with targeted goals. The advancement of educational evaluation commands significant interest from educational institutions at various levels due to its role in steering the educational process and guiding students toward documented progress.

When a learner comprehends the targeted material and recognizes their points of strength and weakness, this constitutes the most significant outcome of evaluation (Hassan & Mansi, 2000, p. 92). It provides the insights necessary to introduce curricular adaptations and diagnose learning difficulties across subjects for remedial intervention (Mansi, n.d., p. 17). Additionally, it offers clear evidence regarding the cumulative and overlapping nature of knowledge, particularly through annual or comprehensive end-of-term examinations (Hassan & Mansi, 2000, p. 93).

Educational evaluation also plays an important role in individual student guidance and counseling (Harizi, 1994, p. 42). Reflecting this significance, Clark (2012) emphasized that incorporating evaluation enhances student learning by developing metacognitive thinking skills, reflection, and self-monitoring capabilities, thereby fostering lifelong learning skills and boosting academic motivation. Similarly, Del Gobbo (2016) confirmed that evaluation positively influences students' academic self-efficacy (Amziane, 2026, pp. 408–409).

This necessitates the active collaboration of all stakeholders within the educational enterprise and a tighter integration between the educational system and society. Although the Ministry of National Education initiated educational system reforms that positioned modern educational evaluation as an integral component of the pedagogical process, a sound mechanism to actualize this vision has been lacking. Consequently, practice has remained stagnant, and traditional examinations continue to dominate educational instruments completely. This has led students to neglect consistent studying throughout the academic year, postponing effort until exam periods, which results in sleep deprivation, reliance on stimulants, and severe physical and psychological strain. Therefore, this research focuses on highlighting authentic evaluation and identifying the difficulties that obstruct its implementation by faculty members. Despite its institutional value, secondary education teachers face substantial hurdles in executing these modern strategies (Masmoudi, 2016, p. 88).

Unquestionably, educators must develop evaluation frameworks that correspond directly to their instructional methodologies. During lesson planning, they must envision targeted outcomes and determine which evaluation strategies best measure them, ensuring that what is taught is precisely what is assessed. As the teaching-learning process strives to fulfill diverse curricular learning outcomes, achieving these goals demands the deployment of evaluation strategies that center fundamentally on learning optimization.

In light of these systematic shifts, several structural reforms have emerged within the educational system, requiring swift adaptability from teachers. Educators must remain flexible and keep pace with modern, effective evaluation strategies that gather objective, valid data using diversified instruments. This allows for the generation of quantitative measurements and

descriptive evidence essential for making sound pedagogical judgments and institutional decisions.

Based on this premise, the present study seeks to uncover the reality of the difficulties confronting secondary education teachers regarding new evaluation strategies. Consequently, the following research questions are posed:

1_ What are the difficulties facing teachers in utilizing new evaluation strategies?

2– Do these difficulties vary according to the study variables (gender, educational qualification, experience, and specialization)?

2_ Study Objectives

The present study aims to achieve the following objectives: 1- To uncover differences in the difficulties facing teachers when utilizing new evaluation strategies.

2- To identify the difficulties facing teachers in utilizing new evaluation strategies that may be attributed to the study variables (gender, educational qualification, experience, and specialization).

2_ Study Concepts

1 - Evaluation:

- **Linguistic Definition:** Derived from the verb *qawwama* (to rectify/straighten). *Qawwama al-mu'waj* signifies straightening or removing crookedness, as documented in *Lisan al-Arab* (Ibn Manzur, 1988, p. 192). Similarly, Al-Fayrouzabadi's *Al-Qamus Al-Muhit* states: *qawwamtuhu* means *'addaltuhu* (I adjusted/rectified it), rendering it *qawim* (upright) and *mustaqim* (straight) (Al-Fayrouzabadi, n.d., p. 95).
- **Idiomatic Definition:** Legendre (1988) defined evaluation as "a descriptive or quantitative judgment regarding the value of a person, object, process, or organization, aimed at providing useful data to guide decision-making" (Al-Farabi, 1994, p. 44).

2_ Evaluation Strategy:

An organized process whereby the evaluator endeavors to assure and optimize the quality of the educational process, utilizing all available instruments.

- **Operational Definition:** The structured set of procedures and steps executed to gather objective, reliable, and valid data using diverse instruments in light of predefined goals. This is done to secure quantitative estimates and descriptive indicators that serve as a foundation for making judgments or decisions. The modernized evaluation strategies addressed in this study comprise performance-based evaluation, paper-and-pencil tests, observation, self-evaluation, and communication-based strategies.

Difficulty:

- **Linguistic Definition:** Implies severity, hardship, or adversity. *Al-Mu'jam Al-Wajiz* states: *sa'uba al-amr* means an issue became severe or arduous, and *tasa'aba al-amr* means it grew difficult (Academy of the Arabic Language, 1993, p. 309). *Al-Mu'jam Al-Wasit* defines *al-sa'b* as the arduous, noting that *sa'uba al-amr 'alayhi* means it became prohibitive or difficult (Mustafa et al., 1989, p. 514). Al-Bustani (1996, p. 604) further notes that *al-masa'ib* represents hardships, tribulations, and distress.
- **Idiomatic Definition:** Represents a lack of control over an activity due to an inability to plan or execute it properly, leaving negative consequences on the undertaking (Al-

Bustani, 1996, p. 606). It refers to the distress or hardship faced by an individual during the planning or execution of a task, resulting in operational failure or cessation.

- **Operational Definition:** The various structural, professional, and environmental obstacles and hurdles confronting secondary education teachers during the practical application of evaluation strategies.

5- Secondary Education:

The formal instructional tier that directly precedes higher education (university) and vocational training.

3_ Review of Previous Studies

The First Study:

Conducted by Youssef Khaneche (2006), titled *"Evaluation Difficulties in Middle School Education and Teachers' Strategies to Overcome Them."*

The primary research question was: **What are the evaluation difficulties in the middle school education stage and the strategies employed by teachers to overcome them?** From this, the following sub-questions emerged: 1- Do middle school education teachers experience evaluation-related difficulties?

2– Are there variations in evaluation difficulties based on gender and the source of training?

3– Do middle school teachers utilize the necessary strategies to resolve their evaluation difficulties?

4– Are there differences in strategy implementation based on gender and training source?

5– What are the primary sources providing middle school teachers with coping strategies?

6– Do variations exist regarding these primary sources based on gender and training source?

The researcher developed and administered a dedicated questionnaire to an accidental sample of 88 male and female teachers across selected middle schools in the Province of Sétif. A descriptive methodology was utilized to analyze the difficulties. The statistical techniques deployed included percentages, frequency distributions, and the Chi-square ($\chi 2$) test to determine differences across variables. The study tested the following hypotheses:

1- Evaluation difficulties are expected to be present among middle school teachers at a rate exceeding 60%.

2- There are no statistically significant differences in evaluation difficulties among middle school teachers attributable to gender or training source.

3- Middle school teachers utilize appropriate strategies to resolve evaluation difficulties at a rate exceeding 60%.

4- There are no differences in the deployment of these strategies due to gender or training source.

5- Teaching experience is expected to be the most critical source providing teachers with appropriate strategies to resolve difficulties.

6- There are no differences in the utilized sources of strategies based on gender or training source.

This study aimed to delineate difficulties and strategies while identifying their origins by examining variations between male and female teachers, and between university graduates and graduates of Technological Institutes of Education (ITE).

Results of the Study:

- The prevalence of evaluation difficulties in middle school ranged between 30% and 86%.
- No statistically significant differences were found between male and female teachers, or between university graduates and Technological Institute graduates regarding evaluation difficulties.
- Middle school teachers demonstrated the capacity to implement necessary strategies to resolve difficulties at a rate of 47% to 87%, with no differences found based on gender or educational background.
- Teaching experience played a vital role in providing necessary strategies, as endorsed by 02% to 57% of the sample. No gender-based differences were observed regarding strategy sources, though statistically significant differences emerged between university graduates and Technological Institute graduates.

Utility to the Present Study:

- Identified a primary variable (difficulties in evaluation) and methods for overcoming it.
- Assisted in structuring the research questions and hypotheses.
- Provided guidance on deploying specific statistical methods, including percentages, frequencies, and the $\chi 2$ test.

The Second Study:

Conducted by Badria Al-Mufarrij (2007), titled *"The Reality of Implementing Portfolio-Based Evaluation for Primary School Students."*

The core problem was addressed via the following question: **Does the current pattern of portfolio-based evaluation play an effective role in achieving comprehensive and appropriate student assessment within primary schools in the State of Kuwait?** The sub-questions were: 1– What is the significance of portfolio-based evaluation?

2– What difficulties confront portfolio-based evaluation?

3– What are the advantages and disadvantages of portfolio-based evaluation?

4– Have teachers been enrolled in training courses to master portfolio utilization?

5– Are there statistically significant differences among sample responses regarding portfolio implementation significance based on independent variables (educational zone, gender, academic specialization)?

6– Are there statistically significant differences regarding difficulties faced based on these same independent variables?

7– What proposals can enhance portfolio-based evaluation?

The researcher administered a questionnaire surveying primary school teachers across the six educational zones of Kuwait (boys' and girls' schools). The instrument was applied to a random sample of core subject teachers (N=35), representing 20% of the target population, and primary school supervisors (N=46), representing 15% of their total population. A descriptive methodology was applied to analyze portfolio implementation. The sample selection aligned with the 2006 statistics from the Planning Department of the Ministry of Education. Validity and reliability were verified prior to final administration. The hypotheses were formulated as follows:

- There are no statistically significant differences among sample responses regarding the importance of portfolio-based evaluation due to educational zone, gender, or academic specialization.
- There are no statistically significant differences among sample responses concerning the difficulties faced within the current evaluation pattern.

The study examined the efficacy of the portfolio evaluation model currently implemented at the primary level. The findings revealed that most teachers and supervisors approved of the importance of portfolio evaluation to a moderate degree; however, they indicated that the current educational environment limits its practical utility.

Utility to the Present Study:

- Guided the formulation of research questions and hypotheses.
- Informed the integration of independent variables, specifically gender and specialization.
- Assisted in designing and adapting specific questionnaire items for the current instrument.

General Comment on Previous Studies:

Although research objectives varied across these studies, they collectively converged on the critical importance of educational evaluation. Some investigations explored the impact of gender and training background among middle school teachers in utilizing coping strategies, concluding that no statistically significant differences existed based on these factors.

The second study, focused on portfolio evaluation within primary education, shares a direct relationship with the present research by addressing a subset of our variables. Reviewing it helped clarify the research problem, facilitating the formulation of our questions and hypotheses. While these past studies generally demonstrated positive correlations among their variables, they were conducted within distinct educational environments and across different educational tiers.

4_ Study Hypotheses

1- There are statistically significant differences among the responses of the study sample regarding the difficulties facing teachers in utilizing new evaluation strategies.

2- There are statistically significant differences among the responses of the study sample regarding the difficulties facing teachers in utilizing new evaluation strategies that are attributable to the study variables (gender, educational qualification, experience, and specialization).

Limitations of the Study

The scope of the current study is defined by the following boundaries:

1- **Spatial Limitations:** This study was conducted exclusively at Sheikh Omar Al-Mukhtar Secondary School in M'Sila.

2- **Temporal Limitations:** The empirical field application of this study was carried out during the period extending from January 02, 2026, to January 22, 2026.

3_ Study Method

The current study adopted a descriptive research design.

Study Tools: To fulfill the objectives of the investigation, a questionnaire was utilized, titled: *"Difficulties in Secondary Education Teachers' Implementation of New Evaluation Strategies."*
Study Population: The target population comprised all teachers working at Sheikh Omar Al-Mukhtar Secondary School in M'Sila.
Study Sample: The study was conducted using a non-probability purposive convenience sample consisting of 56 male and female teachers. Responses on the questionnaire were scored according to the criteria detailed in the table below:

Always	Sometimes	Never
03	02	01

Table No. (01): Scoring Key for the New Evaluation Strategies Questionnaire Items.

Average	High	Very High
[1 - 1.66[	[1.66 - 2.32[	[2.32 - 3]

Tool Description: This instrument assesses the difficulties associated with using new evaluation strategies and contains 29 items distributed across three distinct dimensions:

- **Part One:** Difficulties related to the teacher (11 items addressing internal teaching challenges).
- **Part Two:** Difficulties related to students (11 items examining hurdles stemming from student behaviors/abilities).
- **Part Three:** Difficulties related to the curriculum (7 items assessing challenges imposed by curricular frameworks).

Psychometric Properties:

- *Teacher Dimension:* Reliability = 0.68, Validity = 0.82
- *Student Dimension:* Reliability = 0.68, Validity = 0.82
- *Curriculum Dimension:* Reliability = 0.61, Validity = 0.78

  Conclusion: The instrument possesses acceptable levels of reliability and validity for academic research.

2 – Presentation and Analysis of Results

Question One: What are the difficulties facing teachers in using new evaluation strategies?

Sub-Question (01): What are the difficulties facing primary education teachers in using these strategies?

Table No. (02): Mean Scores, Standard Deviations, and Ranks of Evaluation Difficulties.

Statistical Operations / Difficulties	Arithmetic Mean	Standard Deviation	Difficulty Ranking
Teacher	22.975	3.738	01
Student	26.962	3.107	02
Curriculum	17.475	2.322	03
All Difficulties Combined	67.412	6.437	

Statistical Operations / Difficulties	Arithmetic Mean	Standard Deviation	Difficulty Ranking
Total	134.81	15.58	

Commentary: The results indicate that teachers experience notable difficulties when implementing new evaluation strategies. The overall arithmetic mean for combined difficulties was 67.41 (SD=6.43). The intensity of these challenges was high regarding teacher-related factors, moderate regarding student-related factors, and lowest regarding the curriculum.

These varying degrees of difficulty across dimensions can be attributed to dense curricular demands that are poorly aligned with the learners' immediate environment. Furthermore, high class density (overcrowding) inside classrooms exacerbates these challenges, hindering evaluation. This is compounded by limited instructional time per session and a general lack of comprehensive teacher training regarding modern evaluation tools.

Question No. (02): Are there statistically significant differences between the responses of the sample members regarding the difficulties facing primary education teachers in using evaluation strategies due to gender?

Table No. (03): Differences in Evaluation Difficulties Based on Gender (t-test).

Dimension	Gender	N	Mean	SD	Calculated t-value	Significance Level	df
Teacher	Male	12	22.200	3.790	-1.884	Non-Significant	78
	Female	44	23.750	3.564			
Student	Male	12	27.625	2.816	1.940	**Significant**	78
	Female	44	26.300	3.275			
Curriculum	Male	12	17.825	1.985	1.355	Non-Significant	78
	Female	44	17.125	2.593			
All Difficulties	Male	12	67.650	5.846	0.328	Non-Significant	78
	Female	44	67.175	7.045			

Commentary:

- **Teacher Dimension:** No statistically significant differences were observed between male (M=22.200) and female teachers (M=23.750), where t=−1.884 (non-significant).
- **Student Dimension:** Statistically significant differences were present, with male teachers reporting higher difficulty levels (M=27.625) than female teachers (M=26.300), where t=1.940 (statistically significant in favor of males).

- **Curriculum Dimension:** No statistically significant differences were found between males (M=17.825) and females (M=17.125), where t=1.355 (non-significant).
- For **All Difficulties Combined**, the calculated t-value was 0.328. Comparing this against critical value tables, the calculated value is lower than the tabulated t-value (1.64) at the 0.05 significance level, indicating an overall lack of significant difference based on gender.

Conclusion: From the preceding analysis, we conclude that gender does not exert a statistically significant influence on the overall evaluation difficulties faced by secondary education teachers, except within the student dimension, where male teachers experienced significantly higher levels of difficulty ($t>1.64$, $p\leq0.05$).

Question No. (03): Are there statistically significant differences between the responses of the sample members regarding the difficulties facing secondary education teachers in using evaluation strategies due to educational qualification?

Table No. (03 - Repeated): Differences in Evaluation Difficulties Based on Educational Qualification (t-test).

Dimension	Qualification	N	Mean	SD	Calculated t-value	Significance Level	df
Teacher	License (Bachelor's)	32	24.125	3.399	3.619	**Significant**	78
	Higher Schools	14	21.250	3.601			
	Tech Institutes	32	20.330	2.400			
Student	License (Bachelor's)	32	26.958	3.357	-0.015	Non-Significant	78
	Higher Schools	14	26.968	2.741			
	Tech Institutes	32	17.966	2.344			
Curriculum	License (Bachelor's)	32	17.916	2.050	2.130	**Significant**	78
	Higher Schools	14	16.812	2.570			
		32	15.566	2.366			

Dimension	Qualification	N	Mean	SD	Calculated t-value	Significance Level	df
	Tech Institutes						
All Difficulties	License (Bachelor's)	32	69.000	6.094	2.818	**Significant**	78
	Higher Schools	14	65.031	6.286			
	Tech Institutes	32	53.868	5.110			

Commentary:

- **Teacher Dimension:** Statistically significant differences were found, with License (Bachelor's) graduates experiencing higher difficulty levels (M=24.125) compared to Higher School graduates (M=21.250) and Technical Institute graduates (M=20.330), where t=3.619 (significant in favor of License graduates).
- **Student Dimension:** No statistically significant differences were observed (t=−0.015).
- **Curriculum Dimension:** Statistically significant differences emerged, with License graduates exhibiting higher means (M=17.916) compared to Higher School (M=16.812) and Technical Institute graduates (M=15.566), where t=2.130 (significant in favor of License graduates).
- For **All Difficulties Combined**, the calculated t-value was 2.818, which exceeds the tabulated critical value (2.32) at the 0.01 significance level.

Conclusion: The results demonstrate statistically significant variations in response patterns based on educational qualifications, explicitly in favor of License (Bachelor's) degree holders. This finding suggests that while these graduates possess robust subject-matter knowledge, they report greater difficulties due to systemic gaps in ongoing professional training regarding modern curricular frameworks, educational sciences, and applied educational psychology.

Question No. (04): Are there statistically significant differences between the responses of the sample members regarding the difficulties facing secondary education teachers in using evaluation strategies due to experience?

Table No. (04): One-Way Analysis of Variance (ANOVA) for Evaluation Difficulties Based on Years of Experience.

Dimension	Source of Variation	Sum of Squares / Means	df	Mean Squares	Calculated F-value	Significance Level
Teacher	Between Groups	14.813	2	7.407	0.524	Non-Significant
		1089.130	77	14.145		

Dimension	Source of Variation	Sum of Squares / Means	df	Mean Squares	Calculated F-value	Significance Level
	Within Groups	1013.950	79			
	Total					
Student	Between Groups	20.864	2	10.432	1.083	Non-Significant
	Within Groups	742.024	77	9.637		
	Total	762.887	79			
Curriculum	Between Groups	8.266	2	4.133	0.762	Non-Significant
	Within Groups	417.684	77	5.424		
	Total	425.950	79			
All Difficulties	Between Groups	87.475	2	43.737	1.057	Non-Significant
	Within Groups	3185.910	77	41.375		
	Total	3273.380	79			

Commentary: The variance analysis shows that there are no statistically significant differences between or within groups across any of the evaluation dimensions (Teacher F=0.524; Student F=1.083; Curriculum F=0.762). The global calculated F-value for combined difficulties was 1.057, which falls below the critical tabulated value of 4.79.

Conclusion: Professional experience does not significantly mitigate or exacerbate the evaluation difficulties encountered by secondary education teachers within the investigated context.

Question No. (05): Are there statistically significant differences between the responses of the sample members regarding the difficulties facing secondary education teachers in using evaluation strategies due to specialization?

Table No. (05): Differences in Evaluation Difficulties Based on Academic Specialization (t-test).

Dimension	Specialization	N	Mean	SD	Calculated t-value	Significance Level	df
Teacher	Scientific	37	22.216	4.416	-1.704	**Significant**	78
	Literary	19	23.627	2.936			
Student	Scientific	37	27.081	2.772	0.315	Non-Significant	78
	Literary	19	26.860	3.398			
Curriculum	Scientific	37	17.000	2.000	-1.718	**Significant**	78
	Literary	19	2.000	2.518			
All Difficulties	Scientific	37	66.297	6.301	-1.447	Non-Significant	78
	Literary	19	68.372	6.469			

Commentary:

- **Teacher Dimension:** Significant differences exist (t=−1.704), with the literary specialization experiencing higher difficulty.
- **Student Dimension:** No significant variance detected (t=0.315).
- **Curriculum Dimension:** Significant differences observed (t=−1.718), in favor of the literary track.
- For **All Difficulties Combined**, the calculated t-value was -1.447, which is less than the critical tabulated value (2.32) at the 0.01 level.

Conclusion: While overall combined difficulties show no significant difference based on specialization, separate components (teacher and curriculum) reveal that literary track teachers face greater obstacles. This disparity may be explained by the nature of scientific specializations, which emphasize logical, structured, and systematic thinking patterns in evaluation practices compared to literary disciplines.

2 – Discussion of Research Results in Light of the Hypotheses

First Hypothesis:

There are difficulties facing teachers in using new evaluation strategies to varying degrees across teacher, student, and curricular dimensions.

To test this hypothesis, arithmetic means were compared against established benchmarks using the data from Table No. (01). The statistical analysis confirmed the hypothesis, demonstrating that secondary education teachers face distinct, measurable difficulties across all three areas.

2 – 6 – Second Hypothesis:

There are statistically significant differences between the responses of the sample members regarding the difficulties facing teachers in using new evaluation strategies that are attributable to the study variables (gender, educational qualification, experience, and specialization).

To validate or reject this hypothesis, inferential statistical values were systematically analyzed:

- **Gender:** The calculated t-value was 0.328 (Table 14). When compared against the critical tabulated value of 2.32 (df=78), the calculated value was lower, indicating no significant overall difference based on gender.
- **Educational Qualification:** The calculated t-value was 2.818 (Table 15). This exceeds the tabulated critical value of 2.32 (df=78), demonstrating a statistically significant difference in favor of License (Bachelor's) graduates. These educators experience higher levels of difficulty because they report a lack of systemic pedagogical training regarding comprehensive modern evaluation tools.
- **Experience:** The calculated F-value was 1.057, falling short of the tabulated critical value, thereby indicating an absence of significant differences.
- **Specialization:** The calculated t-value was -1.447 (Table 17), which is less than the tabulated critical value of 2.32, showing no significant overall difference.

Consequently, the second hypothesis was only partially validated, demonstrating significance exclusively within the educational qualification variable.

3– Key Findings and Conclusions

- Secondary education teachers encounter varying degrees of difficulty when implementing modern educational evaluation strategies.
- No comprehensive statistical differences in overall evaluation difficulties were found across gender, experience, or academic specialization, a finding that aligns with the previous literature (Khaneche, 2006).

Identified Pedagogical Challenges:

1– Determining how to activate evaluation so it functions as a formative tool serving student learning.

2– High levels of difficulty in evaluating integration scenarios (specifically in establishing objective criteria and performance indicators).

3– Practical user-complexity associated with modern evaluation instruments.

4– Challenges in leveraging evaluation data to advance learning due to dense curricula, time constraints, and strict grading deadlines.

5– Excessive frequency of testing, where students are subjected to formal evaluations more than three times per semester.

6– Structural barriers to implementation, including inadequate time allocations, classroom overcrowding, and external institutional pressures.
7– The persistence of traditional perceptions that limit the scope of evaluation entirely to grading and marking.
8– Wide socio-academic and cognitive disparities among students within individual classrooms.
9– Structural misalignment between the evaluation mechanisms proposed in the curriculum/textbooks and the actual cognitive readiness of students.
10– Teachers remain anchored to legacy pedagogical models (in both organization and execution) due to insufficient opportunities to fully internalize the competency-based approach.
11- High instability within the educational system due to frequent curricular changes and revisions.

List of Sources and References

1– Mustafa, I., Al-Zayyat, A., Abdel-Qader, H., & Al-Najjar, M. (1989). *Al-Mu'jam Al-Wasit* (Vol. 1). Dar Al-Da'wa, Academy of the Arabic Language.
2– Al-Laqani, A. H., & Al-Jamal, A. A. (1999). *Dictionary of educational terms defined in curricula and teaching methods* (2nd ed.). 'Alam Al-Kutub.
3_ Masmoudi, A. (2016). Difficulties in evaluating learners. *Psychological and Educational Studies*, Laboratory for the Development of Psychological and Educational Practices, (17), 88.
4- Al-Mufarrij, B. (2007). *The reality of implementing portfolio-based evaluation for primary school students* (Unpublished master's thesis in psychology). Kuwait.
5– Al-Bustani, A. (1996). *Al-Bustan* (1st ed.). Maktabat Lubnan.
6– Al-Farabi, A., Agdid, M., El-Atrezi, M., Tum, M., & El-Shayb, M. (1994). *Dictionary of educational sciences: Terms of pedagogy and didactics* (1st ed.). Matba'at Al-Najah Al-Jadida.
7_ Amziane, B. (2026). Obstacles to formative assessment among university faculty members. *Journal of Human and Social Sciences*, 15(2), 406–409.
8_ Sabah, S. (2020). *The level of university professors' utilization of continuous evaluation mechanisms from the perspective of education science students* (Unpublished master's thesis). Mohamed Khider University of Biskra.
9_ Al-Mana'sah, H. M. (2020). *The degree of active learning utilization among Arabic language teachers in the upper basic stage in the schools of Qasabat Amman in light of some variables* (Unpublished master's thesis). Faculty of Educational Sciences, Middle East University.
10_ Al-Farabi, A., et al. (1994). *Dictionary of educational sciences: Terms of pedagogy and didactics* (1st ed.). Matba'at Al-Najah Al-Jadida.
11_ Al-Fayrouzabadi, M. d. M. b. Y. (n.d.). *Al-Qamus Al-Muhit* (Vol. 2). Dar Al-Jil.
12_ Khaneche, Y. (2006). *Evaluation difficulties in middle school education and teachers' strategies to overcome them* (Unpublished master's thesis in psychology and education sciences). University of Batna.
13_ Harizi, M. I. (1994). *Evaluating the objectives of examination questions for obtaining the basic education certificate*. Educational Publications.