Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 10, No. 4, 50-65
2026
Publisher: Learning Gate
DOI: 10.55214/2576-8484.v10i4.12550

Examining the alignment between mathematics learning outcomes and assessment practices in Moroccan universities: The case of the algebra module

Youssef Ezzahraouy[1*], Mustapha Ourahay[2], Ezzaim Laabid[3]
[1,2,3]Interdisciplinary Research Laboratory in Didactics, Education and Training (LIRDEF), Higher Normal School, Cadi Ayyad University, Marrakech, Morocco; ezzahraouyaz@gmail.com , youssef.ezzahraouy@ced.uca.ma (Y.E.) mu.ourahay@uca.ac.ma (M.O.) ez.laabid@uca.ac.ma (E.L.).

Abstract: This study examines the alignment between mathematics learning outcomes and assessment practices in three Moroccan universities. The purpose is to evaluate whether assessment tests adequately represent the intended learning outcomes and reflect the results of teaching activities. The study focuses on the Algebra 2 module, which serves as a reference framework for the targeted learning outcomes. The research adopts a psychodidactic approach, combining the analysis of assessment questions and the module's specification table. The representativeness of the intended learning outcomes in each test is measured using the chi-square goodness-of-fit technique. The results reveal significant inconsistencies in the representativeness of the content and cognitive levels. The findings indicate a recurrent overrepresentation of algebraic reasoning, as well as variations between years and institutions. The high chi-square values (ranging from 45.65 to 125.32) suggest that several tests do not adequately cover the intended learning outcomes. These findings point to a misalignment between assessment practices and the learning outcomes framework. This study highlights the need to rethink university assessment practices by adopting a systematic approach based on clearly defined learning outcomes and psychometric principles. Such improvements could contribute to more valid, balanced, and equitable evaluation of students' competencies.
Keywords: *Algebra, Alignment, Learning outcomes, Representativeness of an assessment test, University assessment.*

1. Introduction and Problem Statement

The Bachelor's degree program in Mathematical Sciences, Computer Science, and Applications (SMIA) is a national program. It is accredited by the Ministry of Higher Education in accordance with the National Pedagogical Standards (CNPN) for the bachelor's degree. It is part of the Bachelor's, Master's, and Doctorate (BMD) system and is based on both a modular and disciplinary approach. The content of the various mathematics modules in this program is defined by a national commission composed of representatives from mathematics departments of various Moroccan universities.

In line with CNPN requirements, each module must be described by specifying the intended competencies, the knowledge to be acquired, the content to be covered, the pedagogical methods to be adopted, and the assessment modalities to be applied. In practice, competencies are often formulated in general terms and reflect what the instructor should develop in students. The pedagogical methods employed are mostly limited to lectures and tutorials. The assessment is centered on the validation of modules and focuses on the format and weight given to the various tests that comprise it.

Furthermore, the content of the modules is generally presented as chapter and subchapter titles without specifying the exact knowledge to be acquired. This grants university teachers, regarded as disciplinary experts, a high degree of autonomy in both detailing the content to be delivered and assuming responsibility for its instruction.

History: Received: 23 January 2025; Revised: 9 March 2026; Accepted: 12 March 2026; Published: 27 March 2026
* Correspondence: ezzahraouyaz@gmail.com

In this context, it is crucial to emphasize the importance of learning outcomes. A clear definition not only clarifies a module's pedagogical objectives but also guides the implementation of pedagogical activities and underpins the design of assessment tests. Evaluating learning outcomes is, therefore, a strategic tool for both steering the educational system and monitoring student progress throughout their academic journey.

Moreover, the assessment of learning outcomes is essential for managing teaching activities, both within the educational system and in the undergraduate student's academic journey. Focusing on the transmission of knowledge and teaching leads us to question how university teachers design summative assessment tests of their students' learning outcomes: Do the assessment tests target the learning outcomes intended by the program? Are they valid and reliable?

Focusing solely on knowledge transmission and instruction without considering the explicit definition of intended learning outcomes and students' learning development has led us to question how university teachers design summative assessments of their students' learning outcomes.

Many studies have addressed the validity and reliability of assessment tests based on psychometric analysis, which involves internal parameters such as the difficulty index, calculated through the average success rate per item, the discrimination index, Cronbach's alpha, the KR-20 coefficient for multiple-choice questions, and correlations between tests [1, 2].

Other studies focus on the content validity of exam tests, based on the analysis of the alignment between test items and the program's learning objectives. This involves analyzing the content coverage of the curriculum and pedagogical objectives [3-6].

International assessments such as PISA and TIMSS Advanced have also been the subject of analyses using didactic tools. Roditi and Salles [7] and Salles [8] applied Robert [9]'s Level of Knowledge Functioning (LKF) model to analyze the competencies required in PISA and TIMSS Advanced tasks. Chevallard, et al. [10] based their studies on the Anthropological Theory of the Didactic, analyzing assessment tasks through praxeological analysis to examine the correspondence between institutional knowledge, curriculum content, and the mathematical activities required in evaluation tasks. Grapin and Grugeon-Allys [11] developed an item analysis tool used for the CEDRE 2008 assessment. The work of Vantourout and Goasdoué [12] highlighted the relevance of the psychodidactic approach in the construction and analysis of test validity.

Our work combines the analysis of item alignment with the learning outcomes of the curriculum and a psychodidactic analysis of test questions. The aim is to analyze the type of response required by each question's statement.

A complementary approach, illustrated by Jamalzadeh et al. [13], is based on psychometric item analysis and a qualitative study of distractors. The authors employ methods such as Differential Item Functioning (DIF) and Differential Distractor Functioning (DDF) to identify biases and inconsistencies in test items. This approach helps improve the interpretation of results and ensures that tasks genuinely assess the intended competencies. Roditi and Salles [7] and Salles [8] conducted a critical analysis of PISA and TIMSS Advanced assessments using the LKF model proposed by Robert [9]. Their methodological approach consisted of a re-analysis of the knowledge required to answer the items.

Our research examines the validity of summative assessment tests developed by teachers responsible for the Algebra 2 module in the SMIA field at three national faculties of science. It is inspired by the work of Roditi and Salles [7] and Salles [8], adopting Aline Robert [9]'s model. It focuses on the alignment of test items with the learning outcomes relevant to the Algebra 2 module. It departs from the psychometric approach used in the work of Li, et al. [1] and Indriyani, et al. [2], relying instead on a psychodidactic analysis of the test items. This analysis is based on a didactic examination of the tasks and the cognitive processes targeted by the mathematical activity required in each item.

2. Conceptual Framework

2.1. Learning Outcomes and Learning Outcomes Referential

According to Cedefop [14][1], d'Hoop, et al. [15], AUF and CITEF [16][2],[3] the intended learning outcome of a teaching unit describes what a learner should know, understand, and be able to do at the end of a learning process. Each learning outcome is characterized by mathematical content, a cognitive activity or process, a response strategy, a type of mobilization, etc.

Learning outcomes play a key role in the design of curricula, training activities, and assessments. They express the expected results of students' learning activities, explicitly and operationally translating the intended objectives, pedagogical intent, and structuring the teaching-learning activity.

The learning outcomes framework of a module is the set of intended learning outcomes for that module, which should be covered through the various teaching and learning activities associated with it. Its development is based on a methodology that relies on a psycho-didactic analysis of the learning activity underlying the teaching activities.

2.2. Pedagogical Alignment

According to Biggs [17]; Biggs [18] and Biggs [19], pedagogical alignment or constructive alignment expresses the coherence between attended learning outcomes, teaching-learning activities, and assessment strategies. In other words, the teacher should organize teaching-learning activities based on the prescribed learning outcomes. Additionally, they should design the assessment of learning outcomes in alignment with the teaching-learning activities carried out. Through pedagogical alignment, learning outcomes serve as a roadmap guiding the design of both teaching-learning activities and assessments.

2.3. The TIMSS[4] Model for Characterizing Learning Outcomes

The international TIMSS survey [20] assesses learning outcomes in mathematics through two dimensions:

Content Dimension: This dimension describes the set of concepts, knowledge, and procedures targeted by the adopted mathematics curriculum.

Cognitive Dimension: This focuses on mathematical activities and the cognitive processes associated with learning the curriculum. It is structured into three cognitive levels: "Know," "Apply," and "Reason".

Each level describes a set of specific cognitive processes or mathematical activities.

- *Knowing*: This level refers to simple questions where the student responds directly using the methods, techniques, and content acquired in the learning activity.
- *Applying*: This level refers to questions that require adapting knowledge to familiar situations in order to answer a question whose solution is not immediate.
- *Reasoning*: This level refers to complex questions requiring the construction and justification of knowledge. The student proceeds in multiple steps, takes initiatives, chooses strategies, and mobilizes prior knowledge while adjusting their approach.

Table 1.
Synthetic table of the different cognitive processes intended by the program.

Cognitive level	Cognitive Process
Knowing	Recognize, Recall, Verify, Show, Implement, Calculate, Decompose, Solve
Applying	Determine, Search, Find, Investigate, Use, Apply, Transform, Reduce, Simplify, Choose, Show, Invert
Reasoning	Demonstrate, Prove, Justify, Argue, Search, Determine, Analyze and Interpret, Construct, Transfer, Evaluate, Judge, Investigate

[1]European Centre for the Development of Vocational Training.
[2]The Francophone University Agency.
[3]International conference for French-speaking engineers and technicians.
[4]Trends in International Mathematics and Science Study.

Table 1 is inspired by the TIMSS methodology and adapted for the treatment of an algebraic activity.

This cognitive model provides a foundation for assessing learning outcomes in mathematics, emphasizing the importance of each level and its associated processes within the context of algebra.

2.4. Level of Knowledge Functioning Model

The level of knowledge functioning model adopted by Robert [9] is a tool for analyzing the mathematical content involved in a mathematical activity and the level of its mobilization to solve a mathematical task (questions or problems). This model distinguishes between two types of knowledge: knowledge as an object and knowledge as a tool.

According to Douady [21], object-type tasks focus on learning a new concept, subject, or technique. Tool-type tasks, on the other hand, aim to mobilize previously acquired concepts, subjects, or techniques within a problem-solving activity. This type of task is categorized into three levels of mobilization:

- *Technical Level*: This level is characterized by the student's ability to mobilize knowledge (concepts, theorems, properties, procedures, definitions, etc.) in an isolated, direct, and immediate manner without any adaptation. According to Salles [8], tasks at this level require students to perform routine activities by applying a procedure suggested by the problem statement, which is assumed to be automated.
- *Mobilizable Level*: At this level, the knowledge mobilized consists of well-identified and correctly used mathematical concepts. However, they are not applied in a simple and immediate way; instead, they require a transformation of the problem statement before being implemented. According to Salles [8], tasks at this level demand adaptations of the given statement to be successfully completed.
- *Available Level*: At this level, mathematical knowledge is integrated into a mathematical activity without explicit indications guiding the student to use specific knowledge. The student is required to shift perspectives, connect different areas of knowledge, and construct rigorous reasoning autonomously. Salles [8] refers to this level as the "Intermediate Level".

Aline Robert's model complements the TIMSS model by seeking to identify the cognitive processes mobilized in response to a question. Through this identification, we can position the activity within the taxonomic level adopted by TIMSS.

We intend to adopt Aline Robert's model as a tool for analyzing the activity required by a task in order to determine the cognitive level targeted by each task in the assessment test. This model will enable us to conduct a detailed and precise analysis of the mathematical activity targeted by each task in the learning outcomes assessment test.

2.5. The Accuracy of Assessment

We characterize the accuracy of assessment through two dimensions:

- *Pedagogical alignment of assessment* refers to the congruence between what the assessment should evaluate and the learning outcomes targeted by the module's teaching activities. This alignment is reflected in the assessment test by ensuring that each component of the learning outcomes is proportionally represented. The test should adequately cover the different elements of the framework while considering the relative importance of each component [22-24].
- *Reliability and validity of an assessment test* are essential elements to ensure the accuracy of the assessment. They help, on one hand, to precisely determine the student's performance level relative to a predefined threshold and, on the other hand, to measure the alignment between what the test actually evaluates and what it claims to assess [25]. The analysis of this dimension relies on a psycho-didactic approach that examines both the test content and the cognitive activities required to answer each question [26].

It is important to note that didactic variables, as well as the epistemological nature of mathematical knowledge, may influence the difficulty of assessment tasks. According to the Anthropological Theory of the Didactic, mathematical activity can be described through Praxeologies composed of types of tasks, techniques, technologies, and theories. Consequently, analyzing assessment items requires examining the adequacy between the mathematical activity required by the task and the competencies intended to be assessed [16, 27]. Such an analysis helps ensure that the test effectively measures the targeted knowledge while considering the epistemological characteristics of the content and the contextual and cognitive factors that may influence students' performance.

2.6. The Epistemological Nature of Knowledge as a Source of Difficulty

The epistemological nature of university-level mathematical knowledge is considered, according to several studies, an obstacle or a source of difficulty for students in learning advanced mathematics. There is a consensus in previous research on the multiple aspects of this difficulty. Among them, we can cite the abstract nature of mathematical knowledge [28], the transition from numerical thinking to structuralist thinking [29, 30], and the characteristics of formalization, unification, and generalization (FUG) in mathematical knowledge [31].

These elements can be used as didactic characteristics of the items in an assessment test. They can also serve as support for analyzing student performance based on the nature of the tasks required by the test items.

2.6.1. Questions whose answers refer to abstract knowledge

Abstract questions pertain to concepts that lack concrete support and are developed through human thought to understand, examine, or explain problems. Several studies have shown that the abstract nature of mathematical concepts, particularly in abstract algebra, poses an obstacle and a challenge in the learning process [32]. The works of Veith, et al. [33]; Hausberger [34]; Hausberger [35]; Hausberger [36]; Hausberger [37] and Hausberger [38] emphasize that the epistemological nature of algebraic concepts is inherently abstract.

Studies by Di Martino, et al. [39], Gueudet and Thomas [40], and Gueudet and Vandebrouck [41] on the high school-to-university transition indicate that mathematical questions undergo epistemological shifts, with mathematical tasks becoming more abstract and formal at the beginning of university education. This transition is often a source of difficulty for students entering higher education.

2.6.2. Questions Whose Answers Require Formalization

The treatment of this type of question requires the use of symbolic language and a syntax specific to formal mathematics [42]. At university, mathematical activity is supported by formal writing. Several studies have shown that formal mathematical concepts pose difficulties and obstacles for students. These include learning difficulties related to new concepts, such as topological concepts (formal definitions of open and closed sets) [43, 44], difficulties in solving tasks, exercises, or mathematical problems [45, 46], and difficulties related to conceptual understanding [34].

2.6.3. Questions Whose Answers Require Generalizations

Generalized questions focus on notions, methods, phenomena, or generic problems, aiming to identify common characteristics from particular cases and extend the validity of universal laws [47]. This process relies on abstraction and formalization [48] and often involves objects already encountered in other forms [32]. For example, abstract algebra generalizes the properties of structures such as the set of rational numbers, Q, the set of real numbers, R, and the set of complex numbers, C. An element of a field can be a real number, a vector, a function, etc. Linear algebra, which deals with vector spaces, linear systems, and linear applications, relies on abstract algebra, geometry, and analysis [46].

2.6.4. Questions Whose Answers Rely on the Unifying Aspect

A unifying-type question focuses on different mathematical structures that share common properties or analogies, defined by axioms. This involves replacing several notions with a new object, often through a new formalism [31]. For example, the concept of a group is unifying because it applies to various algebraic structures such as integers, permutations, and matrices, unifying and generalizing concepts from different fields.

The work of Harel [49] shows that undergraduate students entering university encounter difficulties in changing perspectives and interpreting complex and abstract objects. For instance, it is difficult for them to "show that a subset of polynomials annihilating at 0 is a vector subspace," as it requires both mastery of abstract definitions and the ability to reason about various conceptual relationships. These difficulties relate to the unifying and conceptual nature of objects in linear algebra, which demands significant cognitive adjustments from students.

Dorier [46] demonstrated that it is sometimes necessary to change the unifying framework, for example, by using affine, analytical, or vectorial methods to solve problems outside of linear algebra.

3. Methodological Framework

Let us recall that the objective of this study is to analyze the alignment of learning outcomes assessment tests, carried out by the teachers responsible for the Algebra 2 module, with the learning outcomes of the teaching activity. We aim to determine whether the assessments are adapted to the intended learning outcomes of the teaching-learning activities, so that students can be correctly assessed on their degree of acquisition of these learning outcomes.

To conduct this study, we analyzed the instructional documents of the teachers responsible for the module, which characterize their teaching activities (lecture notes, tutorial sheets). This analysis allowed us to identify the intended learning outcomes of the teaching activities, which we then organized into a framework for the learning outcomes of the Algebra 2 module. This analysis was the subject of the article titled *"Determining the Learning Outcomes of a University Mathematics Module: The Case of the Algebra 2 Module in the SMIA Fields of Study"*, published in *The International Journal of Science, Mathematics and Technology Learning*.

This framework is inspired by the methodology used by TIMSS for defining learning outcomes to support the design of its standardized tests.

Table 2.
Specification table of the Algebra 2 module.

Content Domain	Content component	Cognitive Levels			Rates (%)
		Knowing (K) (%)	Applying (A) (%)	Reasoning (R) (%)	
Usual algebraic structures	Group (GR) (%)	8.70	9.78	9.78	28,26
	Ring (Rn) (%)	4.35	5.43	7.61	17,39
	Body (Bd) (%)	3.26	0.00	4.35	7,61
	Rates (%)	16.31	15.21	21.74	53,26
One indeterminate polynomial and rational fractions	One indeterminate polynomial (Pol) (%)	18.48	11.95	5.43	35,87
	Rational Fractions (FR) (%)	4.35	5.45	1.09	10,87
	Rates (%)	22.83	17.40	6.52	46,74
Rates (%)		39,13	32.61	28.26	100

This specification table was produced as part of the study [26]. It is characterized by the content dimension and the cognitive dimension, which is based on the taxonomic model of Mullis, et al. [20]. It makes it possible to characterize the framework of learning outcomes intended by the teaching-learning activities.

Beyond its descriptive function, this specification table plays a central role in analyzing the validity and representativeness of assessments. It makes it possible to evaluate the extent to which tests are aligned with the requirements of the curriculum by identifying potential imbalances between the content areas covered, the cognitive levels targeted, and the intended learning objectives. As such, it serves as a key tool for examining the quality of the assessments implemented and their ability to

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 10, No. 4: 50-65, 2026
DOI: 10.55214/2576-8484.v10i4.12550

reliably and meaningfully measure students' learning outcomes.

The creation of this specification table reveals that in the content domain of algebraic structures, the "Reasoning" level is the most represented, followed by the "Knowing" level. In contrast, in the domain of polynomials and rational fractions, the "Knowing" level predominates, followed by the "Applying" level. In the content domain of algebraic structures, reasoning accounts for 21.74% (21.74/53.26) of cognitive activity, while the "Knowing" and "Applying" levels each represent around 15%. Conversely, in the content domain of polynomials and rational fractions, the "Reasoning" level does not exceed 6.52%, whereas the "Knowing" level accounts for 22.83%.

3.1. Methodological approach

This study requires the analysis of course content, tutorials, and assessment tests. We can only consider tests from institutions for which we have both course materials and tutorials. For consistency, comparing the tests also requires us to limit ourselves to those from the same academic year. These constraints have led us to restrict our analysis to assessment tests administered in the following faculties: the Faculty of Sciences Semlalia of Marrakech (FSSM), the Faculty of Sciences Dhar El Mahraz of Fez (FSDM), and the Faculty of Sciences Abdelmalek Essaâdi of Tetouan (FSAME). Furthermore, our analysis focused on tests from the years 2015, 2016, and 2017.

To address the research questions, we will use three analytical tools, namely:

1. The use of the learning outcomes framework for the Algebra 2 module as a tool for processing the coverage of mathematical content in each assessment test. This allows for an analysis of the alignment between the tests and the intended learning outcomes of the program.
2. The use of Robert [9] as a tool for processing the cognitive activity targeted by each test question or as a tool for analyzing the LKF required by each response to a test question.
3. The use of a summary table of the different cognitive processes intended by the program as a tool for classifying the types of cognitive processes and activities targeted by each question.

The last two tools will enable us to categorize the questions according to the cognitive levels adopted in our research.

3.2. Data Handlings

We analyzed and categorized the expected response types for each question in the assessment tests using a multi-step approach:

1. *Identifying the mathematical concepts targeted by the question*: This involves determining the algebraic concepts required in the response, along with their respective content domains and subdomains. This step helps define the "content" dimension of the learning outcomes assessed by the question.
2. *Determining the expected cognitive mathematical activity*: This step identifies the level of knowledge mobilization required in the response by referring to the adopted learning outcomes framework. We used Aline Robert's model to categorize the mathematical content involved in the activity into three levels: Technical, Mobilizable, or Available.
3. *Identifying the major cognitive process targeted by the question and its cognitive level*: This involves determining the cognitive process using the synthetic table of different cognitive processes and the TIMSS-adopted cognitive taxonomy. The "major process" refers to the overarching cognitive process required to answer the question. While a response may contain multiple cognitive processes across different stages, the question itself aims at a process that encompasses these intermediate steps.
4. *Identifying the learning outcomes targeted by the question*: This step consolidates the content domains identified in Step (1) with the cognitive processes and their levels from the other steps, structuring them into the learning outcomes targeted by the question.
5. *Constructing the learning outcomes table for an assessment test*: In this step, we determine the weight of each content component (section and chapter of the module) as well as each cognitive component by calculating the frequency of each component addressed in the assessment test. Each test is

thus represented by a synthetic table of the analysis results.

The data from the assessment tests of each university are processed and presented in a table. Each row represents the learning outcomes of a test, while the different columns correspond to the various dimensions of the learning outcomes targeted by each test.

3.3. Statistical Tool for Assessing the Representativeness of Tests

In this study, we use the Chi-square goodness-of-fit test to estimate the representativeness of each component of the intended learning outcomes, as well as the overall representativeness of each test. The representativeness of each component is calculated using the following formula:

$$\textit{Representativeness of component } \boldsymbol{i} = \frac{(\boldsymbol{TO_i} - \boldsymbol{T_e})^2}{\boldsymbol{T_e}}$$

Where:

$\boldsymbol{TO_i}$: represents the observed rate for each learning outcome component.

$\boldsymbol{T_e}$: represents the theoretical rate given by the specification table.

Furthermore, the representativeness of each test is determined using the Chi-square goodness-of-fit formula:

$$\boldsymbol{X_i}^2 = \sum_{\boldsymbol{i=1}}^{\boldsymbol{8}} \frac{(\boldsymbol{TO_i} - \boldsymbol{T_e})^2}{\boldsymbol{T_e}}.$$

The more the observed rates of content and cognitive components in the test deviate from the theoretical rates defined by the specification table, the higher the Chi-square goodness-of-fit value will be. Conversely, if these rates are closer to the theoretical values, the Chi-square value will be lower.

It is important to note that the analysis of the representativeness of the assessment tests is based on eight categorical variables grouped into five content components and three cognitive components. The degrees of freedom are set at $\boldsymbol{l = 7}$. A chi-square test with a significance level is applied, with the theoretical value being $\boldsymbol{X^2}_{\boldsymbol{théorique}} = \boldsymbol{14,067}$. Therefore, if the calculated chi-square goodness-of-fit value exceeds this threshold, the test sample is not representative of the components of the specification table, whereas a value close to or below this threshold indicates good representativeness.

4. Analysis of Test Results

To analyze the data, we compare the results of our test analysis with the Specification Table (S.T.) of the module. This process allows us to assess the representativeness of the "content" and "cognitive" components of the intended learning outcomes, as well as the reliability of the tests by year and by institution. Finally, we conclude this analysis with an examination of the didactic variables characterizing the algebraic activity targeted by each assessment test.

Table 3.
FSSM, FSDM, and FSAME tests, by content and cognitive component.

Institution	Academic year of the test	Test content representativeness rates (%)					Test cognitive level representativeness rates (%)		
		Usual algebraic structures			Polynomials and Fractions		K	A	R
		Gr	Rn	Bd	Pol	FR			
FS Semlalia-Marrakech	January 2017	47.06	29.41	0	17.64	5.88	29.41	41.18	29.41
	January 2016	35.29	41.18	0	23.53	0	17.65	23.53	58.82
	January 2015	56.25	25	6.25	6.25	6.25	37.5	43.75	18.75
FS-Dhar El Mahrez-Fés	January 2017	47.37	21.05	0	31.58	0	36.84	47.37	15.79
	January 2016	41.18	17.65	5.88	35.29	0	35.29	52.94	11.76
	January 2015	46.67	6.67	0	33.33	13.33	26.67	33.33	40
FS Abd Elmalek Essaadi- Tétouane	January 2017	42.86	14.29	7.14	35.71	0	28.57	14.29	57.14
	January 2016	46.67	46.67	0	0	6.67	46.67	33.33	20
	January 2015	43.37	0	0	50	6.67	37.5	37.5	25
Representativeness according to S.T. (%)		28,26	17.39	7.61	20.65	10.86	10.86	32.61	28.26

This table presents the results of the analysis of assessment tests conducted by various university institutions. Each test is represented by the representativeness rate of the two components: content and cognitive, of the targeted learning outcomes. The last row of the table shows the percentages calculated in the module's specification table.

The data in Table 2 reveal that the component related to field structures is not assessed in six tests, while the rational fractions component is not targeted in four tests. These two sections have the lowest coverage of learning outcomes within their respective domains.

The results presented in this table allow for the identification of learning outcome components that are either underrepresented or overrepresented. However, they do not provide a comprehensive and precise overview of the representativeness of each component within the various assessment tests. A more in-depth analysis is therefore necessary to better understand this aspect and draw relevant conclusions.

4.1. Analysis of Assessment Test Accuracy

To carry out this analysis, we calculated the deviations between the observed test rates and the theoretical rates provided by the specification table for each element of the content and cognitive components. The deviation for each test is calculated using the chi-square goodness-of-fit test.

Table 4.
Calculation of chi-square of goodness-of-fit.

Institutions	A.Y.	Algebraic Structures (A.S.)			Polynomials and Fractions (P.F.)		K	A	R	Chi-squares
		Gr	Rn	Bd	Pol	FR				
FS Semlalia-Marrakech	January 2017	12.51	8.31	7.61	9.26	2.28	2.41	2.25	0.05	46.77
	January 2016	1.75	32.55	7.61	4.25	10.86	11.79	2.53	33.05	125.32
	January 2015	27.72	3.33	0.24	24.46	1.96	0.07	3.81	3.20	66.26
FS-Dhar El Mahrez-Fés	January 2017	12.92	0.77	7.61	0.51	10.86	0.13	6.68	5.50	45.65
	January 2016	5.91	0.00	0.39	0.01	10.86	0.38	12.67	9.63	48.21
	January 2015	11.99	6.61	7.61	0.18	0.56	3.97	0.02	4.88	77.83
Elmalek Essaadi-Tétouane	January 2017	7.54	0.55	0.03	0.00	10.86	2.85	10.29	29.51	82.19
	January 2016	11.99	49.30	7.61	35.87	1.62	1.45	0.02	2.41	110.27
	January 2015	8.08	17.39	7.61	5.57	1.62	0.07	0.73	0.38	50.45
Representativeness according to S.T. (%)		28.26	17.39	7.61	35.87	10.86	10.86	32.61	28.26	

The high Chi-square values highlight a significant discrepancy between the observed percentages and the expected percentages (according to the specification table). This table reveals substantial variations from one year to another and from one university institution to another (see the Chi-square column). The obtained values, ranging between 45.65 and 125.32, far exceed the theoretical reference value (14.067), which is linked to the nature of the assessment tests. These results indicate that the representativeness of each test is significantly higher than the expected theoretical value. It can thus be concluded that none of the tests are aligned with the learning outcomes framework.

For a more in-depth analysis of representativeness, it is essential to calculate the gaps between the representativeness rates of each component of the learning outcomes referential. This involves comparing the desired representativeness, as per the specification table, with the representativeness actually observed in the test. These gaps will reveal the components that are overrepresented at the expense of others.

Table 5.
Deviation between test rates and specification table rates.

Institution	Academic Year of the test	Deviation between acquired rates Test / S.T. Domain A.S. (%)				Deviation between acquired rates Test / S.T. Domain P.F. (%)			Cognitive component rate deviation (%)		
		Gr %	**Rn %**	**Bd %**	**Dom. S.A. %**	**Pol %**	**F.R %**	**Dom.Pol-FR %**	**K**	**A**	**R**
FS Semlalia-Marrakech	January 2017	18.8	12.02	-7.61	23.21	-18.23	-4.98	-23.21	-9.72	8.57	1.15
	January 2016	7.03	23.79	-7.61	23.21	-12.34	-10.86	-23.2	-21.5	-9.08	30.56
	January 2015	28	7.61	-1.36	34.24	-29.62	-4.61	-34.23	-1.63	11.14	-9.51
FS-Dhar El Mahrez-Fés	January 2017	19.1	3.66	-7.61	15.16	-4.29	-10.86	-15.15	-2.29	14.76	-12.47
	January 2016	12.9	0.26	-1.73	11.45	-0.58	-10.86	-11.44	-3.84	20.33	-16.5
	January 2015	18.4	-10.7	-7.61	0.08	-2.54	2.47	-0.07	-12.5	0.72	11.74
FS Abd Elmalek Essaadi-Tétouane	January 2017	15.1	-17.4	-0.47	11.03	-0.16	-10.86	-11.02	-1.63	4.89	-3.26
	January 2016	18.4	29.28	-7.61	40.08	-35.87	-4.19	-40.06	7.54	0.72	-8.26
	January 2015	14.6	-3.1	-7.61	-9.89	14.13	-4.19	9.94	-10.6	-18.3	28.88
Representativeness according to S.T. (%)		28,26	17.39	7.61	53.26	35.87	10.86	46.73	39.13	32.61	28.26

Table 4 indicates that all tests, except for those from January 2017 at Abd Elmalek Essaadi and from January 2015 at Dhar El Mahraz faculties, overrepresented the domain of algebraic structures. The January 2016 test from Abd Elmalek Essaadi faculty emphasized group and ring structures while reducing the level of algebraic reasoning in favor of knowledge.

The January 2015 test from Abd Elmalek Essaadi faculty and the January 2016 test from Semlalia faculty overrepresented algebraic reasoning by more than 28%, at the expense of the other two cognitive levels, thus accounting for more than half of the cognitive activity targeted by the specification table. These tests reduced cognitive activity related to algebraic calculation, particularly in the domains of polynomials and rational fractions, while favoring algebraic reasoning in the domain of usual algebraic structures. It is important to note that, according to Bloom's taxonomy, cognitive activity related to reasoning is more complex than that related to knowledge and application. Consequently, an overrepresentation of reasoning in a test increases its level of difficulty.

The tests from Semlalia Faculty for January 2015 and 2017, as well as those from Dhar El Mahraz in January 2016 and 2017, focused on application, primarily emphasizing algebraic calculation within algebraic structures. In contrast, the January 2016 test from Abd Elmalek Essaadi Faculty increased the "Knowledge" level by 7.54% while reducing the "Reasoning" level by 8.26%, making it easier compared to the other tests.

4.2. Analysis of Didactic Variables

In this section, we analyze the didactic variables that characterize the tests examined in this study. The objective is to assess the degree of epistemological abstraction, as well as the dimensions of formalism, unification, and generalization they involve, which helps estimate the level of cognitive demand associated with these tests. These variables, recognized in the literature as key factors contributing to test difficulty in this module, serve as the analytical framework adopted for this study. The tests from the three faculties of science highlight significant differences in cognitive complexity and the nature of the knowledge mobilized.

The January 2016 test at FSSM, the January 2015 test at FSAME, and the January 2015 test at FSDM emphasize algebraic reasoning as well as content related to usual algebraic structures. These tests rely heavily on knowledge that is both epistemologically abstract and of an FUG (Formal, Unifying, and Generalizing) nature, while demanding a high level of cognitive complexity. Consequently, we can classify these tests as very difficult.

In contrast, the January 2015 and January 2017 tests at FSSM, along with the January 2017 and January 2016 tests at FSDM, and the January 2017 test at FSAME, increase the focus on usual algebraic structures by emphasizing the application of results or usual algebraic properties. These tests focus on the epistemological aspect of abstraction, of an FUG nature, and present a moderate level of cognitive complexity. Thus, we can classify them as difficult.

Whereas the January 2016 test at FSAME, unlike the others, reduces cognitive demands by limiting itself to the "Knowledge" level while significantly increasing content related to usual algebraic structures, as well as the abstract, formal, unifying, and generalizing aspects of the knowledge mobilized. This result leads us to classify this test as less difficult than the others.

5. Synthesis and Discussion

The analysis of assessment tests reveals significant gaps in the representativeness of learning outcome components. Certain elements, such as the structure of bodies (Bd) and rational fractions (FR), are frequently underrepresented. This imbalance highlights an uneven coverage of learning outcomes, which could compromise the overall assessment of students' learning outcomes and the accuracy of the results. The data also underscores significant discrepancies between observed and expected outcomes. These substantial gaps indicate that the tests are not aligned with the learning outcomes. Furthermore, the considerable variations from one year to another and from one institution to another reveal the need to design tests based on evaluation criteria centered on learning outcomes.

Previous studies confirm and enrich the findings of this research by identifying similar issues in the assessment of learning outcomes. Alemayehu Wole et al. [5] demonstrated a weak correlation between tests and the intended learning outcomes of the school curriculum, with an imbalanced coverage of cognitive, psychomotor, and affective domains. Onaiba [6] noted a misalignment between exam objectives and curriculum goals, characterized by an overrepresentation of grammar and reading comprehension and the neglect of listening and speaking skills. Uzun and Kılıçkaya [50] highlighted that some skills are overrepresented, while others, such as oral expression, are neglected. PITUWA [3] found that more than 50% of exam questions are not aligned with course objectives. These findings align with our observations regarding the underrepresentation of key components and the misalignment between observed and expected outcomes. They underscore the need to reform university assessment practices to improve quality and accuracy.

Another major finding is the overrepresentation of algebraic structures, particularly groups and rings, at the expense of other areas such as polynomials and rational fractions. This trend favors algebraic reasoning over algebraic computation, which could increase the difficulty of the tests. Indeed, tests focused on algebraic reasoning are more demanding, as this cognitive activity is more complex than mere knowledge or application. Conversely, those who focus on the application of algebraic properties are perceived as less difficult. This disparity underscores the importance of a balanced test design to effectively assess students' learning outcomes.

The sections involving reasoning, such as groups and rings, are perceived by teachers in this program as fundamental components of the mathematical curriculum for this type of degree. The emphasis placed on these concepts leads us to assume that university mathematics teachers in this program view their discipline as an activity centered primarily on reasoning rather than on the mastery of computational techniques. These sections are characterized more by the presence of algebraic reasoning, which could explain these teachers' preference for assessing reasoning through their evaluation practices.

The works of Rasmussen and Wawro [51], Inglis and Alcock [52] all emphasize the importance that university mathematics instructors place on logical reasoning, conceptual thinking, and mathematical rigor, particularly in fields such as abstract algebra. These studies show that concepts like groups and rings are seen as essential for developing students' ability to reason abstractly and construct proofs. Instructors prioritize the assessment of reasoning and conceptual understanding over mastery of computational techniques, reflecting a disciplinary focus on problem-solving and critical thinking. These findings confirm and contextualize our own observations regarding the emphasis on groups and rings in assessment practices, as well as instructors' preference for evaluating reasoning in university-level mathematics education.

The analysis of tests based on their level of cognitive difficulty reveals significant trends. Tests that draw on epistemologically abstract knowledge, characterized by formal, unifying, and generalizing (FUG) qualities, are classified as very difficult. Conversely, those focusing on the application of algebraic properties are deemed difficult, while those that reduce cognitive demands to the level of simple knowledge are considered less difficult [32, 33, 40, 41, 43-47].

In conclusion, the analysis of assessment tests highlights imbalances in the representativeness of learning outcome components. To improve the quality of assessment, it is essential to better align tests with the learning outcomes framework. This requires a thorough revision of test design to ensure that the content and cognitive level are aligned with the intended learning outcomes. Such an approach would ensure a fairer and more representative evaluation of students' learning achievements.

5.1. Research Implications

This study aims, at the national level, to raise awareness among key stakeholders in higher education in Morocco (teachers, pedagogical advisors, and ANEAQ) about the challenges related to the assessment of learning outcomes at the university level. It highlights, for teachers, the importance of designing assessment tools aligned with the intended learning outcomes, ensuring coherence between assessments of learning and teaching-learning activities.

Furthermore, this research proposes a new methodological approach to analyze the relevance of university assessment tools, drawing on a psycho-didactic perspective. This approach, adaptable to other educational contexts, could promote the harmonization of assessment practices on an international scale. Universities and educational systems in other countries could use the findings of this study as a reference (benchmarking) to compare and improve their own assessment practices.

Moreover, this initiative could facilitate the exchange of experiences between Morocco and other countries, particularly within international networks dedicated to education. By identifying the key elements that ensure the validity of assessment tools, this study contributes to a better understanding of the challenges related to the accuracy of evaluations and the optimization of their design. It thus provides concrete avenues for enhancing the quality and reliability of assessment practices in higher education.

5.2. Research Prospects

This study paves the way for comparisons of assessment practices between Moroccan university teachers and their international counterparts who adopt similar teaching approaches. Analyzing the relevance of test questions, as well as the difficulty and discrimination characteristics of Algebra 2 module questions, will be essential. Finally, future research should focus on constructing and managing a standardized assessment test model for this module within the SMIA field.

6. Conclusion

This research investigates the validity of learning outcome assessment tests for the Algebra 2 module across three faculties of science. The primary objective was to analyze the alignment of these assessments with the intended learning outcomes of the teaching-learning activities, as synthesized in the module's specification table. A psycho-didactic analysis was conducted on the assessment questions and anticipated responses to ascertain the validity and reliability, which are fundamental elements for ensuring the accuracy of the evaluation instruments. Based on this investigation, the following conclusions have been drawn:

The analysis results indicate that the Algebra 2 module assessment tests, in the three science faculties studied, are not adequately aligned with the teaching-learning activities. The content and cognitive components of the learning outcomes are neither valid nor representative. This underscores the imperative for explicit formulation of learning outcomes to ensure the design of high-quality assessment tests. The observed lack of validity, both in terms of content and cognitive demands, across all analyzed tests highlights biased assessment practices by instructors, thereby leading to the adoption of inappropriate learning strategies among students.

These observations can be attributed to the fact that instructors base their pedagogical practices on a content-focused approach, even though university curricula are designed following an objective- or competency-based framework. Consequently, they perceive teaching as merely the transmission of a list of topics [53]. From this perspective, assessment is reduced to selecting a sample of content deemed representative of the taught material, without due consideration for the representativeness of the intended learning outcomes. In essence, instructors' assessment practices are rooted in a pedagogical approach centered on the transmission of academic knowledge.

A clear and precise definition of the expected learning outcomes within the training program is crucial for the design, development, and management of assessment tests. The determination of these learning outcomes supports teaching-learning activities and ensures optimal pedagogical alignment between instructional practices and assessment methods.

6.1. Research Limitations

Our analysis was limited to the tests administered in January 2015, January 2016, and January 2017 in three faculties of science. We were unable to analyze the remaining tests as they were either unavailable or inaccessible. However, these tests could be examined similarly using a psycho-didactic approach, assessing the alignment between assessment tests and teaching-learning activities. To deepen the analysis of the tests, it is essential to consider various internal didactic parameters and external formal elements that may influence their performance, such as linguistic characteristics, implicit and explicit response indicators, as well as intra- or extramathematical contexts, among others.

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.

Acknowledgments:

The authors acknowledge the use of ChatGPT to improve grammar, punctuation, and fluency.

References

[1] Y. Li, S. Xu, and J. Liu, "Development and validation of computational thinking assessment of Chinese elementary school students," *Journal of Pacific Rim Psychology*, vol. 15, p. 18344909211010240, 2021. https://doi.org/10.1177/18344909211010240

[2] V. Indriyani, M. A. Taufiq, R. E. Putri, H. Fendi, S. Syafrilianto, and D. Anggraini, "Validity and reliability of semester tests made by teachers: An evaluation study of English learning," in *Proceedings of the 2nd EAI Bukittinggi International Conference on Education (BICED)*, 2020.

[3] J.-C. K. PITUWA, "Validity of some first semester exams at the Faculty of Psychology and Educational Sciences of the University of Kisangani for the 2016-2017 Academic Year (Case of the Department of Educational Sciences)," *Revue Internationale du Chercheur*, vol. 3, no. 2, pp. 533–551, 2022.

[4] L. Ing, M. M. Borhandden, A. S. Al-Hudawi, T. L. Mohd, and M. K. Naail, "Validity of teacher-made assessment: A table of specification approach," *Asian Social Science*, vol. 11, no. 5, pp. 193–200, 2015. https://doi.org/10.5539/ass.v11n5p193

[5] G. Alemayehu Wole, S. Fufa, and Y. Seyoum, "Evaluating the content validity of grade 10 mathematics model examinations in Oromia National Regional State, Ethiopia," *Education Research International*, vol. 2021, no. 1, p. 5837931, 2021.

[6] A. M. Onaiba, "Do exam aims and content reflect those of the curriculum? An evaluative study," *Asian Journal of Assessment in Teaching and Learning*, vol. 14, no. 1, pp. 55-69, 2024. https://doi.org/10.37134/ajatel.vol14.1.6.2024

[7] E. Roditi and F. Salles, "New analyses of the 2012 PISA survey in mathematics," *Éducation & formations*, vol. 86, no. 11, pp. 235-255, 2015.

[8] F. Salles, "New analyses of the TIMSS advanced 2015 study in mathematics," *Éducation & formations*, no. 94, pp. 41-56, 2017. https://doi.org/10.48464/halshs-01693929

[9] A. Robert, "Tools for analyzing mathematical subjects," *Recherches en didactique des mathématiques*, vol. 8, no. 2, pp. 139-190, 1998.

[10] Y. Chevallard *et al.*, *Advances in the anthropological theory of the didactic*. Cham, Switzerland: Springer International Publishing, 2022.

[11] N. Grapin and B. Grugeon-Allys, "Psychometric and didactic approaches to the validity of an external assessment in mathematics: What complementarities and what divergences?," *Mesure et évaluation en éducation*, vol. 41, no. 2, pp. 37-66, 2018. https://doi.org/10.7202/1059172

[12] M. Vantourout and R. Goasdoué, "Psycho-didactic approaches and validity of assessments," *Éducation & Formations*, no. e-302, 2014.

[13] M. Jamalzadeh, A. R. Lotfi, and M. Rostami, "Assessing the validity of an IAU General English Achievement Test through hybridizing differential item functioning and differential distractor functioning," *Language Testing in Asia*, vol. 11, no. 1, p. 8, 2021. https://doi.org/10.1186/s40468-021-00124-7

[14] Cedefop, *Defining, writing and applying learning outcomes: A European handbook*, 2nd ed. Luxembourg: Publications Office of the European Union, 2022.

[15] E. d'Hoop, D. Lemenu, C. Malhomme, and M. Coupremanne, "Interaction between frameworks, teaching practices, training programs and assessment practices," in *Proceedings of the 24e Colloque ADMÉÉ-Europe*, 2012.

[16] AUF and CITEF, *The competency-based approach: A methodological guide*, 2nd ed. Paris, France: AUF, 2017.

[17] J. Biggs, "Enhancing teaching through constructive alignment," *Higher Education*, vol. 32, no. 3, pp. 347-364, 1996. https://doi.org/10.1007/BF00138871

[18] J. Biggs, "Aligning teaching and assessing to course objectives," *Teaching and Learning in Higher Education: New Trends and Innovations*, vol. 2, no. 4, pp. 13-17, 2003.

[19] J. Biggs, "Constructive alignment in university teaching," *HERDSA Review of Higher Education*, vol. 1, pp. 5–22, 2014.

[20] I. V. S. Mullis, M. O. Martin, and M. von Davier, *TIMSS 2023 assessment frameworks*. Chestnut Hill, MA, USA: TIMSS & PIRLS International Study Center, Boston College, 2021.

[21] R. Douady, "Frame games and the tool-object dialectic," *Recherches en didactique des mathématiques*, vol. 7, no. 2, pp. 5-32, 1987.

[22] R. F. DeVellis and C. T. Thorpe, *Scale development: Theory and applications*, 5th ed. Thousand Oaks, CA, USA: SAGE Publications, 2021.

[23] H. Suen and D. Ary, *Analyzing quantitative behavioral observation data*, 1st ed. New York, USA: Psychology Press, 1989.
[24] A. Anastasi, *Psychological testing*, 6th ed. New York, USA: Macmillan, 1988.
[25] D. Laveault and J. Grégoire, *Introduction to testing theories in education and psychology*, 2nd ed. Bruxelles, Belgium: De Boeck Université, 2002.
[26] Y. Ezzahraouy, M. Ourahay, and E. Laabid, "Determining the learning outcomes of a university mathematics module: The case of the Algebra 2 Module in the SMIA Fields of study," *The International Journal of Science, Mathematics and Technology Learning*, vol. 31, no. 1, pp. 91–112, 2024.
[27] B. Barquero, M. Bosch, and Y. Chevallard, *The anthropological theory of the didactic, in Encyclopedia of Mathematics Education, S. Lerman*, 2nd ed. Cham, Switzerland: Springer, 2020.
[28] K. C. Moore *et al.*, "Using abstraction to analyze instructional tasks and their implementation," *The Journal of Mathematical Behavior*, vol. 74, p. 101153, 2024. https://doi.org/10.1016/j.jmathb.2024.101153
[29] T. Hausberger, *The challenge of structuralist thinking in learning abstract algebra: An epistemological approach, in Enseignement des mathématiques et contrat social, Actes du colloque EMF2012, J.-L. Dorier and S. Coutat, Eds.* Geneva, Switzerland: Université de Genève, 2012.
[30] T. Hausberger, "Teaching and learning abstract algebra at university and first elements of a didactics of algebraic structuralism: Cross-disciplinary studies in mathematics didactics and epistemology," Ph.D. Dissertation, Univ. Montpellier, France, 2016.
[31] A. Robert, *Some specific features of mathematics teaching in post-compulsory education.* Paris, France: IREM de Paris, 1987.
[32] I. R. Mili, "Identifying obstacles and difficulties inherent in learning abstract algebra," Ph.D. Dissertation, Univ. Montréal, Montréal, QC, Canada, 2016.
[33] J. M. Veith, P. Bitzenbauer, and B. Girnat, "Exploring learning difficulties in abstract algebra: The case of group theory," *Education Sciences*, vol. 12, no. 8, p. 516, 2022. https://doi.org/10.3390/educsci12080516
[34] T. Hausberger, "On the concept of (homo) morphism: A key notion in the learning of abstract algebra," *arXiv preprint arXiv:1303.7089*, 2013. https://doi.org/10.48550/arXiv.1303.7089
[35] T. Hausberger, "The object-structure dialectic as a frame of reference for a didactic study of algebraic structuralism," *Éducation et didactique*, vol. 11, no. 11-2, pp. 131-151, 2017.
[36] T. Hausberger, "The (homo) morphism concept: Didactic transposition, meta-discourse and thematisation," *International Journal of Research in Undergraduate Mathematics Education*, vol. 3, no. 3, pp. 417-443, 2017. https://doi.org/10.1007/s40753-017-0052-7
[37] T. Hausberger, "Teaching and learning abstract algebra at university: Towards a paradigm of questioning the world," *Educação Matemática Pesquisa*, vol. 21, no. Especial, pp. 322-337, 2019.
[38] T. Hausberger, "Teaching and learning abstract algebra at university: Elements for a didactics of algebraic structuralism," in *Proceedings of the ARDM National Seminar 2017, Paris, France: IREM of Paris*, 2021.
[39] P. Di Martino, F. Gregorio, and P. Iannone, "The transition from school to university in mathematics education research: New trends and ideas from a systematic literature review," *Educational Studies in Mathematics*, vol. 113, no. 1, pp. 7-34, 2023. https://doi.org/10.1007/s10649-022-10194-w
[40] G. Gueudet and M. O. J. Thomas, *Secondary-tertiary transition in mathematics education," in encyclopedia of mathematics education.* Switzerland: Springer, 2020.
[41] G. Gueudet and F. Vandebrouck, "Secondary-higher education transition in mathematics: What research in didactics teaches us," 2021. https://hal.science/hal-03225490v1
[42] J. Avigad, "Mathematics and the formal turn," *Bulletin of the American Mathematical Society*, vol. 61, no. 2, pp. 225-240, 2024. https://doi.org/10.1090/bull/1832
[43] S. Bridoux, "Teaching the first notions of topology at university: A case study," Ph.D. Dissertation, Univ. Paris Diderot–Paris VII, Paris, France, 2011.
[44] S. Bridoux, "Introducing the concept of convergence with 1980s engineering: Didactic dream or reality for the teacher?," presented at the First Conference of International Network for Didactic Research in University Mathematics, 2016.
[45] R. Najar, "The obstacle of formalism at the beginning of higher education," in *In J.-L. Dorier and S. Coutat, Eds., Teaching mathematics and the social contract: issues and challenges for the 21st century – Proceedings of the EMF 2012 conference, GT7, 2012, pp. 1026–1043*, 2012.
[46] J.-L. Dorier, *The teaching of linear algebra in question.* Grenoble, France: La Pensée Sauvage, 1997.
[47] T. Dreyfus, *Advanced mathematical thinking processes," in advanced mathematical thinking.* Dordrecht: Springer, 2002.
[48] M. Rogalski, "Epistemological and didactic approaches to the activity of formalization in mathematics," in *Teaching Mathematics and the Social Contract: Issues and Challenges for the 21st Century – Proceedings of the EMF2012 Conference GT3*, 2012.
[49] G. Harel, "Promoting linear algebraic reasoning among students: Affordances and challenges," *PRIMUS*, pp. 1-21, 2024. https://doi.org/10.1080/10511970.2024.2327325
[50] A. Uzun and F. Kılıçkaya, "TEOG (TEPSE) English test: Content validity and teachers' views," *Bartın University Journal of Faculty of Education*, vol. 9, no. 3, pp. 680-708, 2020.

[51] C. Rasmussen and M. Wawro, "Post-calculus research in undergraduate mathematics education," *Compendium for Research in Mathematics Education*, vol. 2017, pp. 551-581, 2017.
[52] M. Inglis and L. Alcock, "Expert and novice approaches to reading mathematical proofs," *Journal for Research in Mathematics Education*, vol. 43, no. 4, pp. 358-390, 2012. https://doi.org/10.5951/jresematheduc.43.4.0358
[53] J.-M. De Ketele and F.-M. Gerard, "Validation of assessment tests according to the competency-based approach," *Mesure et évaluation en éducation*, vol. 28, no. 3, pp. 1-26, 2005. https://doi.org/10.7202/1087028ar

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 10, No. 4: 50-65, 2026
DOI: 10.55214/2576-8484.v10i4.12550