MODESTUM EURASIA Journal of Mathematics, Science and Technology Education, 2025, 21(3), em2597
ISSN:1305-8223 (online)
OPEN ACCESS Research Paper https://doi.org/10.29333/ejmste/16043

Horizontal alignment between grade 9 mathematics content standards and summative assessment tasks

Agnes Dulu Qhibi [1*], Thelma de Jager [1], Cornelius Johannes White [1]

[1] Tshwane University of Technology, Pretoria, SOUTH AFRICA

Received 30 October 2024 ▪ Accepted 04 February 2025

Abstract

The literature confirms that poor performance in mathematics is a global issue. Studies have established a significant link between the alignment of educational components and high performance. These findings prompted this study, which aims to eliminate misalignment in order to improve learner performance. Specifically, this research examines the degree of alignment between grade 9 mathematics content standards and summative assessment tasks. A qualitative research approach, document analysis, and Webb's (1997) alignment model plus the four sampled summative assessment tasks from four secondary schools were used. The findings revealed extensive misalignment between the two educational components, as many major concepts were not assessed. This study recommends an explicit description of major concepts linked to higher grades in the curriculum. Additionally, artificial intelligence could consider a model that outlines the degree of alignment between two documents through scanning. Furthermore, tertiary institutions should consider establishing modules to educate prospective teachers on aligning educational components.

Keywords: alignment, content standards, mathematics, summative assessment tasks

INTRODUCTION

The trends in international mathematics and science study (TIMSS) revealed that many countries performed below average in grade 9 mathematics (Reddy et al., 2019). A TIMSS report conducted in 2019 revealed that Singapore, Chinese Taipei, Korea, Japan, and Hong Kong learners obtained the highest scores in grade 9 mathematics, while Oman, Kuwait, Saudi Arabia, South Africa, and Morocco achieved the lowest (Reddy et al., 2019). It is evident from the TIMSS report that poor performance in grade 9 mathematics affects not only South African learners, but also those in other countries. Mabena et al. (2021, p. 451) highlights that the teaching and learning of mathematics in grade 9 does not always enable learners to acquire the prescribed learning outcomes. Mabena et al. (2021) further highlights that not acquiring the learning outcomes stipulated in the curriculum and assessment policy statement (CAPS) could impact learners' mathematics performance later in grade 12. This is evident in the grade 12 mathematics performance of South African learners, who achieved below 60% in national mathematics examinations in four consecutive years from 2016 to 2020 (Department of Basic Education [DBE], 2020, p. 182). It is, therefore, significant to enhance mathematics performance in lower grades to enable high performance in higher grades. Kanjee and Mthembu (2015, p. 163) raised the challenge of some teachers being unable to support learners' needs by implementing classroom assessments. In addition, Williamson (2017, p. 303) pointed out that the setting of ineffective school-based assessment tasks, where assessments measure what is not set in the content standards by teachers, tends to negatively impact learners' final examination results. Care et al. (2019, p. 26) claim that not aligning educational components such as content standards, assessment, and pedagogy could lead to potential confusion throughout the system where what should be valued may not be exposed.

In finding a solution to eliminate misalignment and overcome poor performance in mathematics, an investigation was conducted to illuminate the degree of horizontal alignment between grade 9 mathematics content standards and summative assessment tasks.

✉ duluagnes@yahoo.com **(*Correspondence)** ✉ dejagert@tut.ac.za ✉ twhite@global.co.za

Contribution to the literature

- Aligning summative assessment tasks with content standards helps to assess what the curricula intends to achieve. The curriculum should explicitly describe major concepts from the lower grades linked to higher grade concepts to provide a strong foundation for learning future concepts.
- Prospective and current teachers should understand the importance of aligning educational components with the curricula.
- It is impossible to assess all the concepts and skills in the curriculum in one summative assessment task; hence, guidelines to outline major concepts to be assessed are significant.

Webb (1997) emphasizes the significance of horizontal alignment between content standards and assessment to attain the intended learning outcomes. Amiri and Rezvani (2021, p. 52) define horizontal alignment as the essential side-by-side agreement of the facets of education aiming to achieve educational objectives. On the other hand, vertical alignment refers to the up-and-down harmony of educational components from different education levels seeking to accomplish certain objectives (Amiri & Rezvani, 2021, p. 52). The content standards allude to the ideas about what content is most important for learners to acquire in a specific grade (Wixson et al., 2003, p. 69).

As early as 2012, Watermeyer (2012) highlighted that the alignment of educational components is deemed significant as it assists subject disciplines in transforming from fragmented components to integrated, holistic education. Alignment studies between educational components offer valuable insights to teachers on how well the content standards, classroom instruction, and assessment function to help make informed decisions (Abate & Mishore, 2024). Various international and national studies revealed misalignment between content standards and assessment in mathematics (FitzPatrick et al., 2015; Ndlovu & Mji, 2012; Zhao et al., 2023). Thus, investigating the alignment between educational components could improve effective mathematics teaching, learning, and assessment. The literature revealed that few alignment studies have been conducted. Dhlamini (2021, p. 1) confirmed the paucity of horizontal alignment studies; therefore, this study seeks to find possible answers to the following research question:

How does the content structure of grade 9 mathematics summative assessment tasks offered by different secondary schools horizontally align with the grade 9 mathematics content standards in CAPS?

LITERATURE REVIEW

Ananda (2003, p. 1) defines alignment as the degree to which standards, assessments, and other significant educational components, such as classroom instructions, are complementary to ascertain effective learning. The alignment of educational components can be summarized as the connection between what is done in the classroom versus what is intended to be achieved in the curriculum. Alignment of educational components is important, as all elements should support learning in the classroom in order to prevent a detrimental impact on the entire education system, where what should be valued may not be known (Care et al., 2019, p. 26). Lack of alignment between educational components may expose mathematics learners to a curriculum that restricts constructive learning (Smith, 2012, p. 115).

Importance of Conducting Alignment Studies

More than a decade ago, Roach et al. (2008, p. 173) emphasized that the results of alignment studies could help assessment developers and policymakers to refine and align curricula with assessment tasks. Alignment should be used to guide learners towards effective learning (Loughlin et al., 2021). However, teachers experience challenges in turning research findings into practice (Georgiou et al., 2023). Alignment between content standards, assessments, and classroom instructions is associated with increased student learning in developed countries (Atuhurra & Kaffenberger, 2022). This means that conducting alignment studies could help improve instructional practices and assessment in the classroom. Wijngaards-de Meij and Merx (2018) argue that intended learning outcomes can be improved by making it visible through curriculum alignment. Furthermore, alignment between educational components is essential to reach conclusions and enact sound policies (Kazemi, 2022, p. 1). Conversely, Dhlamini (2021, p. 1) highlighted the challenge of limited alignment studies in the education system, which puts it running at risk of operating in disconnected components. Hence, more alignment studies are needed to contribute to improving the education system and mathematics performance in South Africa.

Alignment Studies That Employed Webb's (1997) Alignment Model

Webb's (1997) alignment model is a popular model that has been used to investigate the degree of alignment between content standards and assessment tasks (Gulzar & Mahmood, 2019; Tannenbaum et al., 2015). Using Webb's (1997) adapted alignment model, Tannenbaum et al. (2015) conducted an alignment study between summative assessment tasks and content standards. The

findings highlighted that not all content standards were assessed in the summative assessment tasks, which confirmed misalignment. Tannenbaum et al. (2015, p. 1) further highlights that the summative assessment tasks investigated were developed prior to the content standards, which might negatively impact achieving the intended learning outcomes. It should be noted that assessment tasks should translate what is envisaged to the content standards. Another alignment study was conducted by FitzPatrick et al. (2015), analyzing the objectives and assessment tasks using Webb's (1997) alignment model. The findings revealed that the assessment activities were not well-aligned with the course objectives, as half of the assessment activities did not assess the course objectives. The implication is that mathematics learners may proceed to higher grades without attaining appropriate knowledge and skills on the subject.

A study by Higgins (2013) conducted in Louisville found significant alignment between content standards and assessment using Webb's (1997) alignment model. The assessments were found to be significantly aligned with the content standards. This shows that the assessment measured what is prescribed on the content standards without deviations. Similar findings were obtained by Duke Escobar (2016) between national standards and assessments for elementary mathematics courses in two universities. The findings highlighted that the assessment items included nearly all the national standards for both university programs. These findings show that the assessments were aligned with the national standards. It is critical that assessments should be based on the content standards to achieve the intended learning outcomes. The literature highlights that Webb's (1997) alignment model has been applied in different subjects such as language, arts, mathematics, science, and social studies in more than 10 countries (Blank, 2002, p. 2). This shows how reliable and wide Webb's (1997) alignment model has been used. Therefore, Webb's (1997) alignment model can be used to guide the education system in evaluating the degree of alignment between content standards and assessment across any subject.

Benefits of Aligning Content Standards and Assessment

International and national preliminary alignment studies have investigated the alignment between content standards and assessment (Duke Escobar, 2016; Ndlovu & Mji, 2012). The findings of these alignment studies highlighted both alignment and misalignment between the content standards and assessment. It is, therefore, important to illuminate the benefits of aligning content standards and assessment to work towards eliminating misalignment. Aligning the content standards and assessment tasks helps ensure that learners access the content and skills outlined in those standards (CSAI Update, 2018). Meng (2023) argues that aligning instructional practices with quality standards helps enhance teaching and learning outcomes.

THEORETICAL FRAMEWORK

The theoretical framework guiding this study is based on Webb's (1997) alignment model. Webb's (1997) alignment model consists of five specific criteria: content focus, articulation across grades and ages, equity and fairness, pedagogical implications, and system applicability. These specific criteria logically start with content, followed by learners, then classroom instruction, and finally, the application in an education system. Furthermore, Webb (1997) recommends six content focuses to investigate the alignment between content standards and assessment: categorical concurrence, depth of knowledge consistency, range of knowledge correspondence, structure of comparability, balance of representation, and dispositional consonance.

The categorical concurrence helps to assess whether the assessment exhibits the same content categories as outlined in the content standards (Webb, 2007, p. 10). The depth of knowledge consistency helps to verify if the assessment exhibits the same cognitive complexity as the content standards (Webb, 2007, p. 11). The range of knowledge correspondence assists to verify if the same knowledge span elicited in the assessment corresponds to the knowledge span prescribed in the content standards. The structure of knowledge comparability establishes if the connection of ideas expressed in the performance expectations is consistent between the content standards and assessment. The balance of representation verifies if the weighting of content is consistent between content standards and assessment. The dispositional consonance focuses on learners' attitudes, habits, and beliefs. This study employed only the first three content focuses: categorical concurrence, depth of knowledge consistency, and range of knowledge correspondence. This was done to investigate horizontal alignment between grade 9 mathematics content standards and summative assessment tasks, as guided by the research question and the sample.

MATERIALS AND METHODS

A qualitative research approach was used to better understand the ideas and experiences of the phenomenon under investigation involving non-numerical data (Ugwu & Eze Val, 2023). Again document analysis where written documents are analyzed to avoid researchers' influence on data was also applied (Delice, 2010, p. 3). Purposive sampling was applied as deliberate choice of researchers knowing which qualities participants possess (Etikan et al., 2016, p. 2). Since Mpumalanga Province in South Africa has four districts, four summative assessment tasks for

2.5. A car uses 10 liter of petroleum per 100 km. What is the cost of a journey of 800 km if the petroleum costs R21.00 for 2 liters?

Figure 1. An assessment item on rate extracted from the summative assessment task for school A (Source: Grade 9 mathematics question paper, November 2022, School A)

November 2022 final examination were sampled from the four schools in each district. Permission to sample the summative assessment tasks from the schools was granted by the Mpumalanga Department of Education, while the national Department of Education provided approval to compare these tasks with grade 9 mathematics content standards.

Data Collection and Data Analysis

The researcher and the content analyst, a mathematics specialist, collected data. Training for the content analyst on how to map data was conducted prior to the collection of data. Webb's (1997) three content focuses were used to generate data: categorical concurrence, depth of knowledge consistency, and range of knowledge correspondence. The categorical concurrence is used to identify sub-topics, the depth of knowledge consistency to map content standards and assessment with cognitive levels, and the range of knowledge correspondence to identity knowledge span (Webb, 1997). The categorical concurrence verifies if the assessment addresses the same content categories with the content standards (Webb, 2007, p. 10). The depth of knowledge consistency validates if the assessment is consistent with the content standards in terms of the strength and number of mental connections that learners have to make (Gopal, 2021). The range of knowledge correspondence verifies if the breadth of knowledge required on the content standards is consistent with the assessment (Webb, 2007, p. 13). Zhao et al. (2023) advise that achieving alignment in the classroom is the critical aspect of standard-based curriculum.

Data were analyzed using the three content focuses: categorical concurrence, depth of knowledge consistency, and range of knowledge correspondence where three scales of agreement were used: "full," "acceptable," and "insufficient" (Webb, 1997). The "full" scale of agreement apply where the sub-topics/cognitive levels/knowledge span of the assessment are fully matched with the content standards (Webb, 1997). The "acceptable" scale of agreement apply where nearly all sub-topics/cognitive levels/knowledge span of the assessment correspond with those of the content standards (Webb, 1997). Moreover, the "insufficient" scale of agreement is used where most of the sub-topics/cognitive levels/knowledge span of the assessment are excluded as compared to the content standards (Webb, 1997). It should be noted that cognitive levels recommended in the CAPS document were also used in this study, including knowledge, routine procedures, complex procedures, and problem-solving (DBE, 2011, p. 157).

RESULTS

This section reports on the findings of the study. The findings have been categorized in terms of the broader topics in the CAPS document for grade 9 mathematics content standards. Additionally, the findings highlight the sample of assessment items extracted from the summative assessment tasks to show how they were set.

The Content Structure of the Summative Assessment Task For School A Compared to Grade 9 Mathematics Content Standards

Under whole numbers, only one sub-topic on rate was assessed on the summative assessment task for school A. The properties of numbers, calculations using mathematics operations, estimation, the use of calculators, prime factorization, ratio, direct proportion, and indirect proportion were not assessed. The assessment item on rate extracted from the summative assessment task for school A, which aligns with the grade 9 mathematics content standards, is highlighted in **Figure 1**.

The summative assessment task for school A did not cover sufficient sub-topics under whole numbers, which renders the scale of agreement on categorical concurrence "insufficient." According to grade 9 mathematics content standards, learners were expected to demonstrate knowledge, routine procedures, complex procedures, and problem-solving. However, only problem-solving was covered, which renders the depth of knowledge consistency "insufficient." Only solving problems in context involving rate was covered on the range of knowledge correspondence.

In contrast, knowledge spans, such as describing real number systems involving natural numbers, whole numbers, integers, rational numbers, and irrational numbers, were not assessed. Moreover, calculations involving mathematics operations, estimation, the use of calculators, prime factorization, and solving problems involving ratios, direct proportion, and indirect proportion were also not assessed. This renders the range of knowledge correspondence "insufficient" since many concepts were not assessed.

On integers, the summative assessment task for school A only assessed calculations with integers, omitting properties of integers which highlight the set of basic integers rules. Hence, the categorical concurrence was labelled "insufficient." This omission also affects the depth of knowledge consistency since the expected cognitive levels will not be realized. Hence, the depth of knowledge consistency was also found to be "insufficient." This may negatively impact learning

other continuous topics as prior knowledge is beneficial for study success (Rach & Ufer, 2020, p. 375). The range of knowledge correspondence was also found to be "insufficient" as the calculations of integers only included addition, subtraction, and multiplication, leaving out division.

Furthermore, calculations with integers included squares, cubes, and square roots, omitting cube roots. Additionally, integer properties, including commutative, associative, and distributive properties, as well as additive inverse and multiplicative inverse, were not assessed in the summative assessment task for school A. Therefore, the scale of agreement on the range of knowledge correspondence was "insufficient." Veith et al. (2023) argue that algebra is a crucial component of learning mathematics as it assists learners to handle abstract quantities. It is, therefore, important to cover major concepts on integers as they interconnect with many mathematics concepts, and a good foundation has to be built.

Under exponents, the summative assessment task for school A covered the laws of exponents and calculations with exponents using mathematics operations. Hence, the scale of agreement on categorical concurrence was "acceptable" since most of the major sub-topics were assessed. The depth of knowledge consistency was also "acceptable" as routine procedures were covered, omitting knowledge. The range of knowledge correspondence was also found to be "acceptable" since the first, second, and fourth laws of exponents and integer exponents were assessed. Only the third law of exponents and exponent zero were not assessed. Ulosoy (2019) discovered that most learners do not know the meaning of zero exponent. Therefore, if the third law of exponent and zero exponent is not assessed, the teacher might not know if learners grasped the concepts well before getting to successive grades. The findings on numeric and geometric patterns revealed that the investigation and extension of geometric patterns and the general rule of patterns were covered in the summative assessment task for school A.

In contrast, investigation and extension of numeric patterns were omitted. Hence, the scale of agreement on the categorical concurrence was "insufficient," as numeric patterns are key before learning geometric patterns. However, the scale of agreement on the depth of knowledge consistency was found to be "acceptable" since nearly all the concepts assessed were as cognitively demanding as the concepts on the grade 9 mathematics content standards.

The scale of agreement on the range of knowledge correspondence was "insufficient" as investigation and extension of geometric patterns with constant difference were covered, but numeric patterns and geometric patterns not limited to ratio, learners' own creation, represented in tables, and algebraically were not assessed. An extract from the summative assessment task for school A, which outlines the findings on numeric and geometric patterns, is illustrated in **Figure 2**.

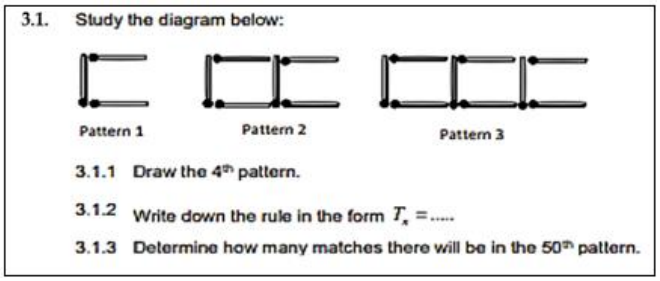


Figure 2. Assessment items on numeric and geometric patterns extracted from the summative assessment task for school A (Source: Grade 9 mathematics question paper, November 2022, School A)

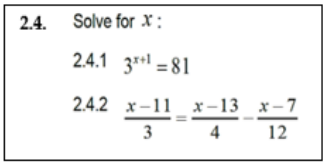


Figure 3. Assessment items on algebraic equations extracted from the summative assessment task for school A (Source: Grade 9 mathematics question paper, November 2022, School A)

Under algebraic expression, the simplification and factorization of algebraic expressions were assessed in the summative assessment task for school A, while algebraic language was omitted. Most of the major concepts were assessed; hence, the scale of agreement on categorical concurrence was "acceptable." The scale of agreement on the depth of knowledge consistency under algebraic expressions was also "acceptable" since routine procedures were assessed on simplification and factorization, leaving out knowledge as one of the expected cognitive levels. Additionally, the scale of agreement on the range of knowledge correspondence was "insufficient," as the knowledge span covered simplification by multiplying monomial by trinomial, factorization by taking out a common factor, difference of squares, and trinomial in the form $ax^2 + bx + c$, where a is a common factor. Multiplying monomial by monomial, monomial by binomial, integer by monomial, integer by binomial and trinomial, dividing integer or monomial by monomial, binomial and trinomial were not covered. It is significant to cover most of the major concepts on assessment to ensure that learners are fully prepared for the next grade. Otherwise, learning new concepts might be difficult for the learners.

On algebraic equations, solving algebraic equations was covered in the summative assessment task for school A; however, setting up an equation to describe a problem situation was not assessed. An extract from the summative assessment task for school A is shown in **Figure 3**.

The scale of agreement on categorical concurrence was "acceptable," while the depth of knowledge consistency was "insufficient" since major concepts of solving equations by factorization and cognitive levels were not assessed. In addition, solving algebraic equations using laws of exponents, additive inverse, and multiplicative inverse was covered, but solving equations by inspection and setting up an equation by describing a situation were excluded. Furthermore, determining the numerical value by substitution and using substitution on equations to generate tables of ordered pairs were also not assessed. Hence, the scale of agreement on the range of knowledge correspondence was "insufficient." Not assessing sufficient content may not give a correct picture of the learners' capability on the content standards. It is, therefore, important to assess concepts and skills effectively to fully prepare learners for future concepts.

Regarding functions and relationships, only input values and output values in the flow diagrams and tables were assessed, determining the rule was not covered for both flow diagrams and tables. Determining input values, output values, and rules of formulae and equations were also not assessed, including justification and interpretation of equivalence or rule presented verbally in flow diagrams, tables, formulae, equations and graphs. Therefore, the scale of agreement on the categorical concurrence and the range of knowledge correspondence was "insufficient." The depth of knowledge consistency was "insufficient" since major concepts were not as cognitively demanding as the major concepts on the grade 9 mathematics content standards.

Under graphs, drawing graphs and determining equations of linear graphs were covered in the summative assessment task for school A, but x-intercept, y-intercept, and gradient as features of linear graphs were not assessed. The scale of agreement on the categorical concurrence was "acceptable," as most of the sub-topics on graphs were covered. The depth of knowledge consistency on the assessment items was not as cognitively demanding as the grade 9 mathematics content standards. Additionally, the scale of agreement on the range of knowledge correspondence was "insufficient" since learners were asked to identify a linear graph when given an equation instead of assessing the skill of drawing a linear graph. Under transformation geometry, translation and reflection were assessed on the summative assessment task for school A, as shown in **Figure 4**.

The assessment items were aligned to the grade 9 mathematics content standard since there was one-to-one correspondence on sub-topics. Rotation was excluded from the 2022 annual teaching plan (ATP) since the content was reduced due to the impact of COVID-19, and the summative assessment task also did not assess it. Thus, the assessment items on the summative assessment task for school A under transformation geometry were well-aligned with the grade 9 mathematics content standards. Therefore, the scale of agreement on the categorical concurrence was "full." The scale of agreement on the depth of knowledge consistency was also "full" as the major concepts were as cognitively demanding as the grade 9 mathematics content standards. It is, therefore, essential to align assessment items to content standards to help achieve the intended learning outcomes. However, the scale of agreement on the range of knowledge correspondence was "acceptable," the assessment items covered translation across quadrants and reflection on the x-axis, but translation within quadrants and reflection on the y-axis were not assessed.

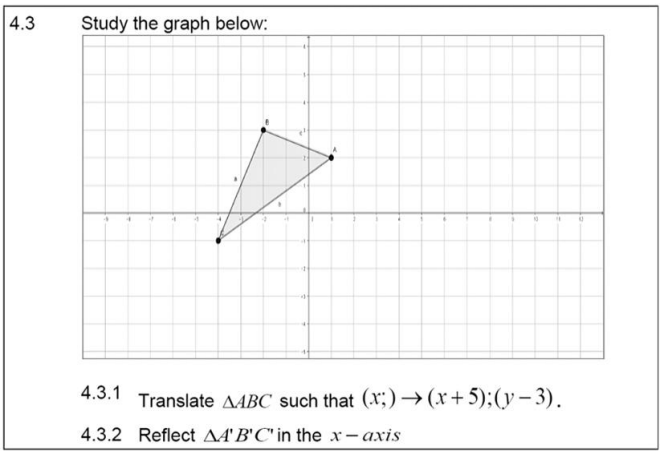


Figure 4. Assessment items on transformation geometry extracted from the summative assessment task for school A (Source: Grade 9 mathematics question paper, November 2022, School A)

Under the geometry of 2D shapes, the properties of triangles and quadrilaterals, as well as the congruency of triangles, were covered in the assessment items for school A, while the similarity of triangles as a major sub-topic was excluded. The scale of agreement on categorical concurrence was "insufficient." The scale of agreement on the depth of knowledge consistency was also "insufficient" as major concepts were not as cognitively demanding as the grade 9 mathematics content standards. Additionally, the scale of agreement on the range of knowledge correspondence was found to be "insufficient." Only the properties of right-angled triangles were covered, but properties of equilateral triangles and isosceles triangles were omitted. Furthermore, definitions of quadrilaterals involving a kite were covered, but parallelogram, rectangle, square, rhombus, and trapezium were not assessed. Additionally, proving the congruency of triangles was covered, while proving the similarity of triangles was not assessed. Conducting alignment studies of educational components becomes critical to support learning in the classroom; otherwise, learning could be compromised and negatively impact the entire education system (Care et al., 2019, p. 26).

Under geometry of straight lines, the intersecting, perpendicular, and parallel lines were covered in the summative assessment task for school A, which renders the scale of agreement of the categorical concurrence "full." The scale of agreement on the depth of knowledge consistency was "acceptable" as nearly all the major concepts were as cognitively demanding as the grade 9 mathematics content standards. The scale of agreement on the range of knowledge correspondence was "acceptable" since most of the knowledge span was assessed on the summative assessment task for school A. Concerning the Pythagoras theorem, the sub-topics assessed on the summative assessment task for school A were in one-to-one correspondence with the grade 9 mathematics content standards; hence, the scale of agreement on the categorical concurrence was labelled "full." However, the scale of agreement on the depth of knowledge correspondence was "acceptable." The cognitive levels envisaged by the grade 9 mathematics content standards were to demonstrate knowledge and routine procedures, but the assessment items only assessed routine procedures. The knowledge span on the assessment items included using Pythagoras' theorem to solve unknown length on right-angled triangles, which was also expected on the grade 9 mathematics content standards. Therefore, the scale of agreement on the range of knowledge correspondence was "full."

Regarding the area and perimeter of 2D shapes, both the area and perimeter of 2D shapes were assessed on the summative assessment task for school A and were as cognitively demanding as the grade 9 mathematics content standards, which caused the scale of agreement on categorical concurrence and depth of knowledge consistency to be "full." In contrast, the knowledge span covered in the assessment items for school A included solving geometric problems involving the perimeter and area of polygons and omitting the area and perimeter of circles. As a result, the scale of agreement on the range of knowledge correspondence was "insufficient." Misalignment of content is a matter that affects many areas. This shows that misalignment between content standards and assessment is a common issue that must be addressed.

Concerning surface area and volume of 3D objects, the surface area and volume of 3D objects were assessed, but capacity was excluded, which put the scale of agreement on the categorical concurrence to be "acceptable." Additionally, the scale of agreement on the depth of knowledge consistency was found to be "full" as major concepts on assessment items were as cognitively demanding as the grade 9 mathematics content standards. The knowledge span covered solving geometric problems involving the surface area of cylinder and rectangular prism, but triangular prism was not assessed. Solving geometric problems involving the volume of a triangular prism was addressed, but the volume and capacity of the rectangular prism and cylinder were not assessed. As a consequence, the scale of agreement on the range of knowledge correspondence was "insufficient."

The Content Structure of the Summative Assessment Task for School B Compared to Grade 9 Mathematics Content Standards

The summative assessment task for school B did not assess any concepts on the following topics: whole numbers, integers, exponents, numeric and geometric patterns, algebraic expressions, algebraic equations, geometry transformation, surface area, and volume of the 3D objects. This differs from the summative assessment task for school A, where all the topics were assessed even though they were not fully assessed. Most of the excluded topics form part of algebra, which means that algebraic skills have been compromised, impacting learning subsequent topics related to algebra.

Regarding functions and relationships, the categorical concurrence covered the flow diagrams, tables, and equations, but formulae, as well as interpreting and justifying the equivalence of different descriptions, were not assessed. Thus, the scale of agreement on the categorical concurrence was "acceptable" since most of the major sub-topics were covered. On the other hand, the scale of agreement on the depth of knowledge consistency was "insufficient" as the cognitive levels of the assessment items were not as demanding as the grade 9 mathematics content standards. The summative assessment task for school B only assessed routine procedures and did not assess complex procedures. This may limit learners' ability to deal with complex mathematics problems. Furthermore, the knowledge span included determining input and output values of flow diagrams, equations, and tables, but omitted the rules for flow diagrams, input values and rules for tables, as well as output values and rules for formulae. This renders the scale of agreement on the range of knowledge correspondence "insufficient."

Under graphs, gradient, drawing linear graphs, x-intercepts, and y-intercepts were assessed in the assessment items for school B, but determining the equation of the linear graph was not covered as a major sub-topic. Hence, the scale of agreement on the categorical concurrence was "insufficient." The summative assessment task for school A assessed knowledge and routine procedures in determining the features of linear graphs, using tables to plot a linear graph, and drawing linear graphs. These were the cognitive levels required by the grade 9 mathematics content standards. However, only knowledge was assessed by using tables to plot the linear graphs, and routine procedures were omitted. Hence, the depth of knowledge consistency was found to be "acceptable," as the assessment items were closely demanding cognitively as the grade 9 mathematics content standards. The scale of agreement on the range of

knowledge correspondence was labelled "insufficient" as the knowledge span covered excluded major concepts such as using tables for ordered pairs to plot the graph and determining the equation of the linear graphs.

Regarding geometry of 2D shapes, the sub-topic covered on the assessment items was properties of quadrilaterals, but properties of triangles, similarity of triangles, and congruency of triangles were not covered, which renders the scale of agreement on the categorical concurrence "insufficient." The scale of agreement on the depth of knowledge consistency was "insufficient" as the major concepts were not assessed, and those assessed were not as cognitively demanding as the grade 9 mathematics content standards. The knowledge span was also compromised as limited major concepts were assessed. Hence, the scale of agreement on the range of knowledge correspondence was "insufficient." On the geometry of straight lines, the assessment items covered parallel lines and intersecting lines, leaving out perpendicular lines. Most major concepts were assessed; therefore, the scale of agreement on the categorical concurrence was "acceptable." The depth of knowledge consistency was also "insufficient," as learners were expected to demonstrate routine procedures and complex procedures in solving geometric problems involving relationships between pairs of angles and lines. However, the summative assessment task for school B required learners to demonstrate routine procedures and problem-solving. Thus, the cognitive levels were not as demanding as the grade 9 mathematics content standards envisaged. The range of knowledge correspondence was found to be "acceptable," as nearly all the knowledge span expected was covered. The knowledge span covered determining a relationship of angles formed by parallel lines and intersecting lines but omitting perpendicular lines. Additionally, in solving geometric problems, parallel lines and intersecting lines were assessed, but perpendicular lines were omitted. Therefore, the range of knowledge correspondence was "acceptable" since major concepts were assessed.

Regarding the Pythagoras theorem, both the scales of agreement on the categorical concurrence and the range of knowledge correspondence were "full." The sub-topics and knowledge span aligned well with the grade 9 mathematics content standards. In contrast, the depth of knowledge consistency was "insufficient" as the cognitive levels were not as demanding as the grade 9 mathematics content standards. The grade 9 mathematics content standards expected learners to demonstrate knowledge and routine procedures, but the summative assessment task for school B required learners to demonstrate knowledge and problem-solving. The assessment items extracted from the summative assessment task for school B are reflected in **Figure 5** and **Figure 6**.

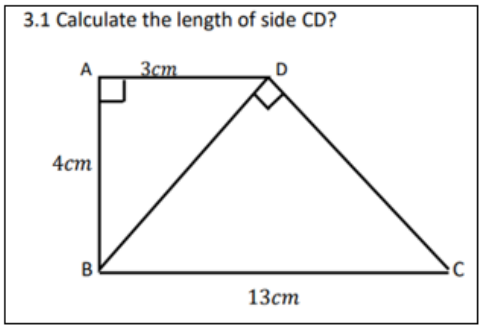


Figure 5. An assessment item on the Pythagoras theorem extracted from the summative assessment task for school B (Source: Grade 9 mathematics question paper, November 2022, School B)

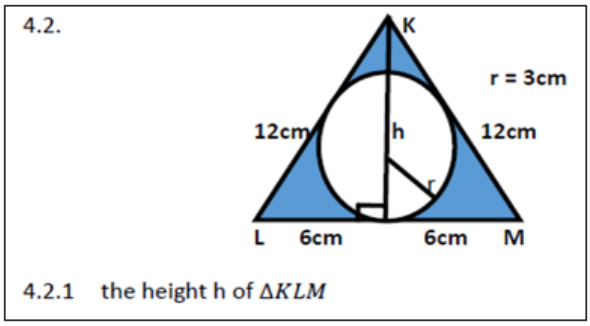


Figure 6. An assessment item on the Pythagoras theorem extracted from the summative assessment task for school B (Source: Grade 9 mathematics question paper, November 2022, School B)

These assessment items show alignment between summative assessment tasks and grade 9 mathematics content standards under the Pythagoras theorem. Moreover, the scale of agreement on the range of knowledge correspondence was "full" since all the major concepts were assessed on the Pythagoras theorem. Under area and perimeter of 2D shapes, the scale of agreement on the categorical concurrence was "full" since there were assessment items on area and perimeter, which were envisaged by the grade 9 mathematics content standards. The scale of agreement on the depth of knowledge consistency was "insufficient." The grade 9 mathematics content standards expect learners to showcase skills in knowledge, routine procedures, and problem-solving. However, the summative assessment task for school B required learners to demonstrate knowledge, routine procedures, and complex procedures. Thus, the cognitive levels were less demanding than the grade 9 mathematics content standards. The knowledge span included solving problems and calculating the area of polygons, circles, and the perimeter of polygons, but the perimeter of a circle was omitted. Therefore, the scale of agreement on the range of knowledge correspondence was "acceptable."

1.4. A recipe needs ¾ cups of sugar, 1½ cups of flour, and a 1/3 cup of milk. Determine the ratio of the ingredients in a simplest form.

Figure 7. An assessment item on whole numbers extracted from the summative assessment task for school C (Source: Grade 9 mathematics question paper, November 2022, School C)

The Content Structure of the Summative Assessment Task For School C Compared to Grade 9 Mathematics Content Standards

Only one sub-topic, the ratio, was assessed on whole numbers, omitting properties of numbers, calculations using mathematics operations, estimation, the use of calculators, prime factorization, ratio, direct proportion, and indirect proportion. This renders the scale of agreement on the categorical concurrence "insufficient." The scale of agreement on the depth of knowledge consistency was "insufficient" since many cognitive levels on whole numbers were not assessed. The scale of agreement on the range of knowledge correspondence was "insufficient." This was concluded since assessment items only covered ratio in its simplest form and did not cover solving problems involving ratio, rate, and direct and indirect proportion. Furthermore, the properties of numbers, including natural numbers, whole numbers, integers, rational numbers, and irrational numbers, were also not part of the assessment. Moreover, calculations involving mathematics operations, estimation, use of calculators, and prime factorization to find the lowest common multiples and highest common factors were also not covered. A sample of an assessment item from the summative assessment task under whole numbers is reflected in **Figure 7**.

The summative assessment task for school C did not assess integers, transformation geometry, geometry of straight lines, functions, and relationships. This deprives learners of an opportunity to showcase their skills on these topics. Under exponents, all sub-topics, which include laws of exponents, integer exponents, and calculations of exponents using mathematics operations, were assessed on a summative assessment task for school C. Hence, the scale of agreement on the categorical concurrence was "full." Yet, the scale of agreement on the depth of knowledge consistency was "insufficient" since the cognitive levels on the assessment items were not as demanding as the cognitive levels on the grade 9 mathematics content standards. The knowledge span on exponents covered calculations using the second, third, and fourth laws of exponents and exponent zero, but the first law of exponents was not assessed. Additionally, calculations using subtraction and multiplication involving numbers in exponential form were covered, but operations such as division and addition were omitted. Hence, the scale of agreement on the range of knowledge correspondence was "acceptable."

Under numeric and geometric patterns, investigating and extending geometric patterns and the general rule of patterns were covered, but investigating and extending numeric patterns were not. A sample assessment of items on numeric and geometric patterns from the summative assessment task for school C is shown in **Figure 8**.

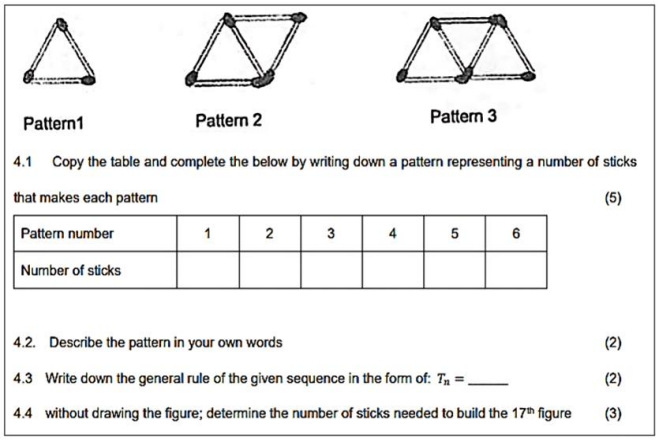


Figure 8. Assessment items on numeric and geometric patterns extracted from the summative assessment task for school C (Source: Grade 9 mathematics question paper, November 2022, School C)

The scale of agreement on the categorical concurrence was found to be "insufficient," and many sub-topics on numeric and geometric patterns were not assessed. The scale of agreement on the depth of knowledge consistency was "full," the assessment items assessed the expected cognitive levels from the grade 9 mathematics content standards. Both knowledge and routine procedures were assessed from the assessment items. The knowledge span assessed under numeric and geometric patterns includes investigating and extending geometric patterns not limited to sequences involving constant differences and represented in tables, describing a general rule of patterns in learners' own words and algebraic form. The patterns involving learners' own creation and represented algebraically were not covered, and numeric patterns were not assessed. As a result, the range of knowledge correspondence was "insufficient."

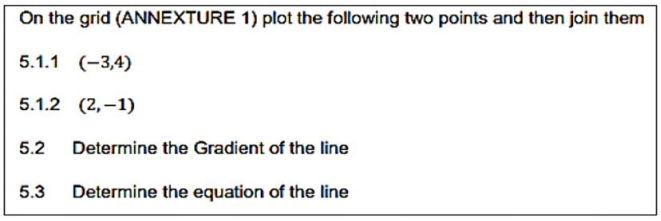
Regarding algebraic expressions, simplification and factorization were assessed, while algebraic language was not covered. This renders the scale of agreement to be "acceptable." The fact that algebraic language was not assessed is a concern as Marpa (2019) discovered that pre-service teachers could not translate mathematics statements into algebraic symbols. The scale of agreement on the depth of knowledge consistency was "insufficient," as the cognitive levels were not as demanding as the grade 9 mathematics content

standards. The knowledge span covered simplification, using associative laws, distributive laws, and laws of exponents to add and subtract algebraic terms, multiplying integers by binomials, while dividing the integers or monomials by monomials, binomials and trinomials were not covered. Additionally, factorization that involves common factors, differences of two squares, and trinomials of the form $x^2 + bx + c$ were addressed, but trinomials in the form $ax^2 + bx + c$, where a is a common factor were not assessed. Moreover, determining the squares, cubes, square roots, and cube roots of single algebraic-like terms was also not part of the assessment items. Thus, the scale of agreement on the range of knowledge correspondence was "acceptable."

Solving algebraic equations was assessed, but setting up an equation to describe problem situations was omitted. This renders the scale of agreement on the categorical concurrence "acceptable." The scale of agreement on the depth of knowledge consistency was "insufficient," as the assessment items were not as cognitively demanding as the grade 9 mathematics content standards. This shows that aligning assessment items with content standards in terms of cognitive levels is a challenge. The knowledge span on algebraic equations covered solving equations using additive inverse and multiplicative inverses; however, setting up an equation to describe problem situations, using inspection, using laws of exponents, and determining the numerical value by substitution were not covered. Hence, the scale of agreement on the range of knowledge correspondence was "insufficient."

Under graphs, all three content focuses, categorical concurrence, depth of knowledge consistency, and range of knowledge correspondence, were found to be "insufficient." The sub-topics covered were gradient, determining equations of linear graphs, and plotting points on the Cartesian plane, which were assessed instead of drawing linear graphs, while x-intercept and y-intercepts were not addressed. This makes the scale of agreement on the categorical concurrence "insufficient." It is significant for assessments not to deviate from what is expected by the content standards in order to be well-aligned. This is highlighted since the assessment items expected learners to plot points on the Cartesian plan rather than draw a linear graph, confirming misalignment. Considering the depth of knowledge consistency, learners were expected to demonstrate knowledge and routine procedures, but only knowledge was assessed. Therefore, the scale of agreement on the depth of knowledge consistency was "insufficient." The knowledge span on graphs covered gradient, omitting x-intercept and y-intercepts.

Additionally, determining the equation of the linear graph and using ordered pairs to plot the points on the Cartesian plane were assessed, but using tables of ordered pairs to draw the graph was not covered. Hence, the scale of agreement on the range of knowledge correspondence was "insufficient" as the skill of drawing linear graphs is key. The assessment items on graphs extracted from the summative assessment task for school C are outlined in **Figure 9**.


On the grid (ANNEXTURE 1) plot the following two points and then join them

5.1.1 $(-3,4)$

5.1.2 $(2,-1)$

5.2 Determine the Gradient of the line

5.3 Determine the equation of the line

Figure 9. Assessment items on numeric and geometric patterns extracted from the summative assessment task for school C (Source: Grade 9 mathematics question paper, November 2022, School C)

Under geometry of 2D shapes, the scale of agreement on the three content focuses: categorical concurrence, depth of knowledge consistency, and range of knowledge correspondence, was "insufficient." The sub-topics covered were congruency of triangles and similarity of triangles, while properties of triangles and properties of quadrilaterals were not covered. This renders the scale of agreement on the categorical concurrence "insufficient." Additionally, the scale of agreement on the depth of knowledge consistency was "insufficient" since the assessment items were not as cognitively demanding as the grade 9 mathematics content standards. The range of knowledge correspondence on the geometry of 2D shapes covered proof for congruency and similarity of triangles and solving geometric problems involving unknown sides on triangles using similar triangles. Conversely, properties of quadrilaterals and properties of triangles were not covered; hence, the scale of agreement on the range of knowledge consistency was "insufficient."

On the Pythagoras theorem, the scale of agreement on the categorical concurrence and range of knowledge correspondence was found to be "full." The sub-topics and knowledge span covered in the summative assessment task for school C under the Pythagoras theorem were well-aligned with the grade 9 mathematics content standards. The range of knowledge correspondence covered the application of the Pythagoras theorem to solve an unknown length on a right-angled triangle, which covered the expected grade 9 mathematics content standard. However, the scale of agreement on the depth of knowledge consistency was "acceptable," as only routine procedures were assessed, and knowledge was not covered. Concerning the area and perimeter of 2D shapes, only the area was assessed, and the perimeter was not covered. Hence, the scale of agreement on categorical concurrence was "insufficient." The scale of agreement on the depth of knowledge consistency was found to be "insufficient," as the cognitive levels were not as demanding as the

1.1.1 Which of the following number is a rational number?

A. $\sqrt{3}$

B. $\sqrt{16}$

C. $\sqrt{-9}$

D. $\sqrt{13}$

Figure 10. Assessment items on whole numbers extracted from the summative assessment task for school D (Source: Grade 9 mathematics question paper, November 2022, School D)

grade 9 mathematics content standards. The summative assessment task for school C assessed knowledge and routine procedures, while grade 9 mathematics content standards expected learners to demonstrate knowledge, routine procedures, complex procedures and problem-solving. Therefore, the cognitive levels of the summative assessment task for school C were not as expected based on the grade 9 mathematics content standards. The scale of agreement on the range of knowledge correspondence was "insufficient" since the knowledge span only assessed solving geometric problems and calculations involving the area of polygons and circles, and nothing was assessed on the perimeter.

Under the surface area and volume of 3D objects, the scale of agreement on the categorical concurrence, depth of knowledge consistency, and range of knowledge correspondence were "insufficient." Only volume was assessed, surface area and capacity were not assessed, and the cognitive levels assessed were not as demanding as the cognitive levels on the grade 9 mathematics content standards. This confirms that the summative assessment task for school C was not well-aligned with the grade 9 mathematics content standards.

The Content Structure of the Summative Assessment Task For School D Compared to Grade 9 Mathematics Content Standards

The content structure of the summative assessment task for school D highlighted that almost all the topics outlined in grade 9 mathematics content standards were assessed, except for exponents, surface area, and volume of 3D objects. Under whole numbers, only real number systems were covered, but estimation, calculations of whole numbers using mathematics operations, the use of calculators, prime factorization, ratio, rate, and direct and indirect proportions were not covered. The scale of agreement on categorical concurrence was "insufficient." The sample of assessment items are reflected in **Figure 10**.

The "insufficient" scale of agreement on categorical concurrence applies where there is no one-to-one correspondence of sub-topics (Webb, 1997). The depth of knowledge correspondence was also found to be "insufficient" since the cognitive levels on the assessment items were not as demanding as those on the

1.1.3 $100 \times 10 - 100 \div 10 =$

A. 990

B. 900

C. 910

D. 901

Figure 11. Assessment items on integers extracted from the summative assessment task for school D (Source: Grade 9 mathematics question paper, November 2022, School D)

4.2. Given the sequence of numbers: ***6; 10; 14; 18; ...***

4.2.1. Write down the next two numbers of the sequence.

4.2.2. Write down the general term (Tn) of the given sequence.

4.2.3. Which term in the sequence is equal to 202?

Figure 12. Assessment items on numeric and geometric patterns extracted from the summative assessment task for school D (Source: Grade 9 mathematics question paper, November 2022, School D)

grade 9 mathematics content standards. The knowledge span assessed included a real number system that covered rational numbers, leaving out natural numbers, whole numbers, integers, and irrational numbers. Moreover, calculations on whole numbers using mathematics operations, estimation and calculators were not assessed. Similarly, prime factorization and solving problems in contexts involving ratio, rate, and direct and indirect proportion were also not assessed. As a result, the scale of agreement on the range of knowledge correspondence was "insufficient." Under integers, only calculation with integers was covered, but properties were not assessed. A sample of assessment items is shown in **Figure 11**.

The scale of agreement on categorical concurrence, depth of knowledge consistency, and range of knowledge correspondence were "insufficient." The cognitive levels on integers were not as cognitively demanding as the cognitive levels envisaged by grade 9 mathematics content standards. Furthermore, most of the knowledge span and sub-topics expected by grade 9 mathematics content standards were not assessed. The topic of numeric and geometric patterns was also assessed on the summative assessment task for school D, but only numeric patterns were covered, omitting geometric patterns. Hence, the scale of agreement on the categorical concurrence was "insufficient." A sample extracted from the summative assessment task for school D is given in **Figure 12**.

The grade 9 mathematics content standards expect learners to demonstrate knowledge and routine procedures for rule generalization, investigation, and extension of numeric and geometric patterns. However, the assessment items only required learners to apply knowledge and routine procedures to investigate and extend numeric patterns. Therefore, the scale of agreement on the depth of knowledge consistency was

"acceptable." The knowledge span covered the extension of number patterns involving common differences, leaving out geometric patterns, constant ratios, and patterns of learners' own creation, represented in tables and algebraically. Additionally, describing the general rule of patterns in algebraic language was covered, but describing the general rule of patterns in own words was not covered. Hence, the scale of agreement on the range of knowledge correspondence was "insufficient."

On algebraic expressions, the scale of agreement on the categorical concurrence was found to be "acceptable." In contrast, the scale of agreement on the depth of knowledge consistency and the range of knowledge correspondence was "insufficient." The categorical concurrence covered most sub-topics, while the cognitive levels were not as cognitively demanding as the grade 9 mathematics content standards. Hence, the scale of agreement on the range of knowledge correspondence was found to be "insufficient," as most of the knowledge span was not assessed. On algebraic equations, solving equations was assessed, but setting up an equation to describe problem situations was not assessed. Therefore, the scale of agreement on the categorical concurrence was "insufficient."

Moreover, the agreement scale on the depth of knowledge consistency was "insufficient" because the cognitive levels were not as demanding as the grade 9 mathematics content standards. The assessment items only included solving equations using additive inverse and multiplicative inverse, but they left out solving equations by inspection, substitution, and laws of exponents. Additionally, setting up an equation to describe problem situations, analyzing and interpreting equations that describe a given situation were also not assessed. This renders the scale of agreement on the range of knowledge correspondence "insufficient."

Under functions and relationships, only flow diagrams were covered, but tables, formulae, equations, and justifying equivalence of different descriptors of the same relationships were not assessed. The scale of agreement on the categorical concurrence was "insufficient." A sample of an assessment item extracted from the summative assessment task for school D is reflected in **Figure 13**.

The scale of agreement on the depth of knowledge consistency and the range of knowledge correspondence was also "insufficient." The cognitive levels on the assessment items were not as demanding as the grade 9 mathematics content standards. Furthermore, the knowledge span covered on the assessment items did not cover most major concepts. On graphs, features of linear graphs and determining the equation of a linear graph were assessed, but drawing linear graphs was not assessed. Therefore, the scale of agreement on the categorical concurrence was "insufficient" since the major concept of drawing linear graphs was not covered. The issue of not aligning cognitive levels of the assessment items and the content standards seems to be a major concern that needs to be addressed, as seen in many topics.

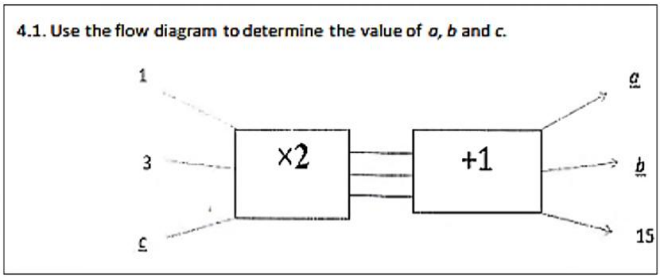


Figure 13. Assessment items on functions and relationships extracted from the summative assessment task for school D (Source: Grade 9 mathematics question paper, November 2022, School D)

Similarly, the cognitive levels on graphs were not as cognitively demanding as the grade 9 mathematics content standards; hence, the scale of agreement on the depth of knowledge consistency was "insufficient." Likewise, the issue of not assessing most of the knowledge span in a topic is also a challenge to many topics. The summative assessment task for school D did not assess most of the major concepts under graphs. Only gradient and determining the equation of the linear graph were assessed, omitting x-intercepts y-intercepts and drawing linear graphs. Hence, the scale of agreement on the range of knowledge correspondence was "insufficient."

All the envisaged sub-topics by the ATP were covered in transformation geometry, translation, and reflection. Rotation was omitted on the ATP due to the implications of COVID-19. The summative assessment task for school D covered both translation and reflection; hence, the scale of agreement on the categorical concurrence was "full." The cognitive levels on the assessment items covered what was envisaged and also exceeded; hence, the scale of agreement on the depth of knowledge was "acceptable." The scale of agreement on the range of knowledge correspondence was also "acceptable," as most of the major concepts on transformation geometry were assessed. The assessment items included translation across quadrants, reflecting an object about the y-axis, writing coordinates of the image and writing the rule used to reflect the object. Reflecting an object on the x-axis and translating an object within and across quadrants were not covered.

On the geometry of 2D shapes, the scale of agreement on both the categorical concurrence, depth of knowledge consistency, and range of knowledge correspondence was "insufficient." The summative assessment task for school D covered properties of triangles, congruency, and similarity of triangles. In contrast, properties of quadrilaterals and solving geometric problems

involving sides and angles in triangles and quadrilaterals were not covered. Additionally, the cognitive levels on the assessment items were not as cognitively demanding as the grade 9 mathematics content standards. Under geometry of straight lines, the scale of agreement on the categorical concurrence and range of knowledge correspondence was "acceptable," while the depth of knowledge consistency was "insufficient." The assessment items covered parallel lines, intersecting lines and solving geometric problems involving relationships between pairs of angles and between the two lines. Perpendicular lines were not assessed; hence, the scale of agreement on the categorical concurrence and the range of knowledge correspondence was "acceptable." However, the cognitive levels covered in the assessment items were less cognitively demanding than those of the grade 9 mathematics content standards. The assessment items only assessed routine procedures, while the grade 9 mathematics content standards expected learners to demonstrate knowledge, routine procedures, and complex procedures.

The scale of agreement on the categorical concurrence and range of knowledge correspondence was "full," while the depth of knowledge was "acceptable." Assessing learners effectively gives a better picture of how far learners have grasped the concept, unlike limiting assessment, which might give a wrong impression in gauging learners' understanding of the topics. Under area and perimeter, the scale of agreement on the categorical concurrence was "full" since area and perimeter were assessed on the summative assessment task for school D. Moreover, the scale of agreement on the depth of knowledge consistency was "insufficient" since the cognitive levels on the assessment items were not as demanding as the grade 9 mathematics content standards. Additionally, the scale of agreement on the range of knowledge correspondence was found to be "acceptable," as most of the major concepts were assessed. The knowledge span included calculating the perimeter of a triangle and the area of a triangle and circle, while the area and perimeter of quadrilaterals were omitted. Investigating alignment by considering the content structure between content standards and assessments provides in-depth information about how the assessment was developed, how often the content standards were assessed, and how content standards overlap with one another (Bae et al., 2019, p. 17). This study illuminated the degree of alignment between summative assessment tasks and grade 9 mathematics content standards by examining the content structure.

DISCUSSION

The challenge of teachers setting classroom assessments which do not measure what it is intended in the curriculum negatively affect learners' performance (Williamson, 2017, p. 303). The study illuminated two significant results of misalignment between grade 9 mathematics content standards and summative assessment tasks. Firstly, the study revealed misalignment between grade 9 mathematics content standards and the four summative assessment tasks. Secondly, disparities identified in implementing grade 9 mathematics summative assessment tasks within one province.

Misalignment Between Grade 9 Mathematics Content Standards and Summative Assessment Tasks

The four summative assessment tasks were found not to be well aligned with the grade 9 mathematics content standards in terms of the categorical concurrence, depth of knowledge consistency and range of knowledge correspondence. Meaning, the sub-topics, cognitive levels and knowledge span assessed were not fully aligned with the requirement of the grade 9 mathematics content standards. Most of the major concepts were not assessed and most assessment items were not cognitively demanding as the content standards. Atuhurra and Kaffenberger (2022, p. 1) also found that teachers tend to cover a wide range of content and cognitive levels that are not well aligned with the curriculum standards. Atuhurra and Kaffenberger (2022, p. 40) further highlight that identifying the areas of misalignment in terms of topics, cognitive demand, and content progression is the first step in bringing classroom instruction and content standards into alignment to promote effective learning. Tannenbaum et al. (2015, p. 9) also found assessment items that were not well aligned with the content standards.

Misalignment of cognitive levels between content standards and assessment was also found by Zhao et al. (2023, p. 1) in China, where cognitive demands of the classroom assessments were not well-aligned with the teaching objectives and curriculum standards. This indicates that misalignment between assessment and content standards is not a new challenge. The implication of misaligned content standards and assessment may affect effective implementation of assessment and subsequently learner performance. Tazkiya (2023) advice that mastering mathematics concepts in elementary level helps to solve mathematics concepts of higher order. It is, therefore, important to extensively teach and assess mathematics concepts in lower grades as they interconnect with mathematics concepts in higher grades, and a good foundation has to be built. Mabena et al. (2021) highlight that not acquiring the learning outcomes stipulated in the lower grades curriculum could impact learners' mathematics performance later in grade 12. Aligning educational components could help achieve the intended learning outcomes as pointed out by Webb (1997, p. 5).

Different findings were obtained earlier by Higgins (2013) and Duke Escobar (2016) where assessments were found to be fully aligned with the content standards.

This shows that the assessments were good translation of the content standards. Porter (2002) advice that assessments should actually translate what is intended in the content standards. Teachers and examiners should be aware that assessments should in fact translate content standards to ensure full alignment of educational components.

Disparities Between Summative Assessment Tasks on Content Standards Assessed

The four summative assessment tasks from the four districts in Mpumalanga Province were inconsistent in terms of the content standards assessed. The summative assessment tasks from the four districts were found to have assessed different content standards for grade 9 mathematics in the 2022 final examination. It appears that teachers used their own discretion to select the content standards to be assessed on summative assessment tasks, as no specific guideline is provided by CAPS. It should be noted that CAPS does not give an explicit description of content standards to be assessed in summative assessment tasks. Instead, it outlines all the content standards to be taught and assessed. Assessing all the content standards in one summative assessment task may not be possible as only fraction is covered. A study by Diko et al. (2015) earlier found that disparities exist in terms of how provinces implement assessment policies. Inconsistency in assessment across schools and provinces is a matter that need to be addressed to ensure that assessment policies are implemented effectively. Areekkuzhiyil (2021) argues that many tests developed in the classroom are not standardized and do not follow systematic test development procedures and principles. Again, Kikwato et al. (2023) found inconsistency in the depth and frequency of monitoring assessments oversight across different schools in Solwezi District of North-Western Province, Zambia by head teachers. This shows that implementation of assessment policies is ineffective and inconsistent. Regulating assessment nationally becomes critical to ensure all schools implement assessment effectively.

CONCLUSION

This study sought to illuminate the degree of alignment between grade 9 mathematics content standards and summative assessment tasks. This study highlighted that grade 9 mathematics summative assessment tasks are not well aligned with the grade 9 mathematics content standards in terms of assessing different content standards and exclusion of most of the major concepts. Therefore, effective implementation of summative assessment tasks which is aligned to the content standards in grade 9 is critical to ensure learners attain the necessary concepts and skills to succeed in higher grades. Stakeholders and the implementers of the content standards should understand that the content prescribed in the CAPS document needs to be translated into classroom instruction and assessed effectively. This kind of alignment will demonstrate if learners are prepared for the next grade and ensure that educational components are well aligned.

Therefore, this study recommends that the major concepts linked to higher grades should be well outlined in the CAPS in order to benchmark and guide the teachers and examiners to ensure consistency in all grade 9 mathematics classes. The study also recommends that alignment modules be established and incorporated at tertiary institutions to fully prepare prospective teachers in the field of teaching and capacitate current teachers on the importance of aligning educational components. Moreover, more alignment studies should be conducted in different subjects to improve the alignment of educational components, which will subsequently help achieve the intended learning outcomes. This study recommends considering artificial intelligence where two documents can be scanned to give a degree of alignment rather than document analysis which consume a lot of time. Finally, this study focused solely on grade 9 mathematics and four schools in the Mpumalanga Province, resulting in a limited scope.

Author contributions: ADQ: conceptualization, methodology, data collection, data analysis, writing original draft, and writing the article & **TDJ & CJW:** reviewing, editing, and supervision. All authors agreed with the results and conclusions.

Funding: No funding source is reported for this study.

Acknowledgments: The authors would like to thank Dr. Dhlamini Zwelithini Bongani for introducing Qhibi Agnes Dulu in alignment studies during her master's degree.

Ethical statement: The authors stated that the study was approved by the Faculty Committee for Research Ethics–Humanities on 14 September 2022 with approval number FCRE/EF/STD/2022/26. Written informed consents were obtained from the participants.

Declaration of interest: No conflict of interest is declared by the authors.

Data sharing statement: Data supporting the findings and conclusions are available upon request from the corresponding author.

REFERENCES

Abate, T., & Mishore, E. (2024). Alignment analysis between teacher-made tests with the learning objectives in a selected school of Central Regional State of Ethiopia. *Heliyon, 10*(11), Article e31869. https://doi.org/10.1016/j.heliyon.2024.e31869

Amiri, A., & Rezvani, R. (2021). A tale of three official English textbooks: An evaluation of their horizontal and vertical alignment. *Language and Related Research, 12*(3), 51-79. https://doi.org/10.29252/LRR.12.3.3

Ananda, S. (2003). Rethinking issues of alignment under no child left behind. *WestEd*. https://api.semanticscholar.org/CorpusID:151624722

Areekkuzhiyil, S. (2021). Issues and concerns in classroom assessment practices. *Edutracks, 20*(8), 20-23.

Atuhurra, J., & Kaffenberger, M. (2022). Measuring education system coherence: Alignment of curriculum standards, examinations, and teacher instruction in Tanzania and Uganda. *International Journal of Educational Development, 92*, 1-16. https://doi.org/10.1016/j.ijedudev.2022.102598

Bae, Y., Tunstall, S. L., Knowles, K. S., & Matz, R. L. (2019). Alignment between learning objectives and assessments in a quantitative literacy course. *Numeracy, 12*(2), 1-25. https://doi.org/10.5038/1936-4660.12.2.10

Blank, R. (2002). Models for alignment analysis and assistance to states. *Council of Chief State School Officers*. https://curriculumanalysis.org/Reference/AlignmentModelsforStateAssist02.pdf

Care, E., Vista, A., & Kim, H. (2019). *Assessment of transversal competencies: Current tools in the Asian region*. UNESCO. https://neqmap.bangkok.unesco.org/wp-content/uploads/2020/09/Assessment-of-TVC-Current-Tools-in-Asian-Region.pdf

CSAI Update. (2018). Standards alignment to curriculum and assessment. *U.S. Department of Education*. https://files.eric.ed.gov/fulltext/ED588503.pdf

DBE. (2011). Curriculum and assessment policy statement: Senior phase mathematics grades 7-9. *Department of Basic Education*. https://www.education.gov.za/Portals/0/CD/National%20Curriculum%20Statements%20and%20Vocational/CAPS%20SP%20%20MATHEMATICS%20GR%207-9.pdf?ver=2015-01-27-160141-373

DBE. (2020). National senior certificate 2020. Diagnostic. Diagnostic report part 1: Content subjects. *Department of Basic Education*. https://www.education.gov.za/Portals/0/Documents/Reports/2021%20NSC%20Reports/Diagnostic%20Report%202020%20-%20Part%201.pdf?ver=2021-03-25-105406-000

Delice, A. 2010. The sampling issues in quantitative research. *Educational Sciences: Theory and Practice, 10*(4), 2001-2018.

Dhlamini, Z. B. (2021). Evaluating the alignment between the grade 9 mathematics annual national assessment and the TIMSS test items. *South African Journal of Education, 41*(3), 1-13. https://doi.org/10.15700/saje.v41n3a1765

Diko, N., Haupt, G., & Molefe, M. R. M. (2011). Reviewing the role of the provincial and district offices in the implementation of assessment policies in the Gauteng and Western Cape provinces. *Human Science Research Council*. https://repository.hsrc.ac.za/handle/20.500.11910/3913

Duke Escobar, V. G. (2016). *An international comparison study of the alignment between standards and assessment of pre-service elementary mathematics teacher preparation courses* [Master's thesis, Boise State University].

Etikan, I., Musa, S. A., & Alkassim, R. S. (2016). Comparison of convenience sampling and purposive sampling. *American Journal of Theoretical and Applied Statistics, 5*(1), 1-4. https://doi.org/10.11648/j.ajtas.20160501.11

FitzPatrick, B., Hawboldt, J., Doyle, D., & Genge, T. (2015). Alignment of learning objectives and assessments in therapeutics courses to foster higher-order thinking. *American Journal of Pharmaceutical Education, 79*(1), 1-10. https://doi.org/10.5688/ajpe79110

Georgiou, D., Diery, A., Mok, S. Y., Fischer, F., & Seidel, T. (2023). Turning research evidence into teaching action: Teacher educators' attitudes toward evidence-based teaching. *International Journal of Educational Research Open, 4*, 1-8. https://doi.org/10.1016/j.ijedro.2023.100240

Gopal, D. (2021). *Application of Webb's depth of knowledge on learners' responses to probability questions* [Doctoral thesis, University of the Witwatersrand]. https://hdl.handle.net/10539/32872

Gulzar, K., & Mahmood, N. (2019). Challenges to maintaining alignment between secondary level mathematics curriculum and assessments in Pakistan. *Journal of Research, 13*(2), 234-246.

Higgins, R. M. (2013). *Examining alignment: National and local assessments and the common core state standards in mathematics* [Doctoral thesis, University of Louisville]. https://doi.org/10.18297/etd/617

Kanjee, A., & Mthembu, J. (2015). Assessment literacy of foundation phase teachers: An exploratory study. *South African Journal of Childhood Education, 5*(1), 142-168. https://doi.org/10.4102/sajce.v5i1.354

Kazemi, J. (2022). *Measuring alignment between components of an educational system: An extension of Porter's alignment index* [Doctoral thesis, The City University of New York].

Kikwato, K.F., Mwanapabu, N.H., & Chanda, C.T. (2023). Exploring the continuous assessment activities and learners' academic performance: A case of selected secondary schools in Solwezi District of North-Western Province, Zambia. *International Journal of Novel Research in Education and Learning, 10*(6), 42-59.

Loughlin, C., Lygo-Baker, S., & Lindberg-Sand, Å. (2021). Reclaiming constructive alignment. *European Journal of Higher Education, 11*(2), 119-136. https://doi.org/10.1080/21568235.2020.1816197

Mabena, N., Mokgosi, P. N., & Ramapela, S.S. (2021). Factors contributing to poor learner performance in

mathematics: A case of selected schools in Mpumalanga Province, South Africa. *Problems of Education in the 21st Century, 79(3)*, 451-466. https://doi.org/10.33225/pec/21.79.451

Marpa, E. P. (2019). Common errors in algebraic expressions: A quantitative-qualitative analysis. *International Journal on Social and Education Sciences, 1*(2), 63-72. https://doi.org/10.46328/ijonses.11

Meng, S. (2023). Enhancing teaching and learning: Aligning instructional practices with education quality standards. *Research and Advances in Education, 2*(7), 17-31. https://doi.org/10.56397/RAE.2023.07.04

Ndlovu, M., & Mji, A. (2012). Alignment between South African mathematics assessment standards and the TIMSS assessment frameworks. *Pythagoras, 33*(3), 1-9. https://doi.org/10.4102/pythagoras.v33i3.182

Porter, A.C. (2002). Measuring the content of instruction: Uses in research and practice. *Educational Researcher, 31*(7), 3-14. https://doi.org/10.3102/0013189X031007003

Rach, S., & Ufer, S. (2020). Which prior mathematical knowledge is necessary for study success in the university study entrance phase? Results on a new model of knowledge levels based on a reanalysis of data from existing studies. *International Journal of Research in Undergraduate Mathematics Education, 6*(3), 375-403. https://doi.org/10.1007/s40753-020-00112-x

Reddy, V., Arends, F., Harvey, J., Winnaar, L., Juan, A., Hannan, S., Isdale, K., & Sekhejane, P. (2022). *The Gauteng Province TIMSS 2019 grade 9 results: Building achievement and bridging achievement gaps*. HSRC Press.

Roach, A. T., Niebling, B. C., & Kurz, A. (2008). Evaluating the alignment among curriculum, instruction, and assessments: Implications and applications for research and practice. *Psychology in the Schools, 45*(2), 158-176. https://doi.org/10.1002/pits.20282

Smith, R. J. (2012). *Alignment of intended learning outcomes, curriculum and assessment in a middle school science program* [Master's dissertation, Edith Cowan University]. https://ro.ecu.edu.au/theses/489

Tannenbaum, R. J., Baron, P. A., & Kannan, P. (2015). Alignment between innovative summative assessment prototypes and the common core state standards: An exploratory investigation. *Educational Testing Service*. https://www.ets.org/Media/Research/pdf/RM-15-07.pdf

Tazkiya, F. (2023). The importance of mastering mathematical concepts in solving mathematical problems in elementary schools. *Ezra Science Bulletin, 1*(1), 11-17. https://doi.org/10.58526/ez-sci-bin.v1i1.3

Ugwu, C. N., & Eze Val, H. U. (2023). Qualitative research. *International Digital Organization for Scientific Research, 8*(1), 20-35.

Ulosoy, F. (2019). Serious obstacles hindering middle school students' understanding of integer exponents. *International Journal of Research in Education and Science, 5*(1), 52-69.

Veith, J. M., Beste, M. L., Kindervater, M., Krause, M., Straulino, M., Greinert, F., & Bitzenbauer, P. (2023). Mathematics education research on algebra over the last two decades: Quo vadis? *Frontiers in Education, 8*. https://doi.org/10.3389/feduc.2023.1211920

Watermeyer, R. (2012). *Curriculum alignment, articulation and the formative development of the learner*. International Baccalaureate.

Webb, N. L. (1997). Criteria for alignment of expectations and assessments in mathematics and science education. *Council of Chief State School Officers*. https://eric.ed.gov/?id=ED414305

Webb, N. L. (2007). Issues related to judging the alignment of curriculum standards and assessments. *Applied Measurement in Education, 20*(1), 7-25. https://doi.org/10.1080/08957340709336728

Wijngaards-de Meij, L., & Merx, S. (2018). Improving curriculum alignment and achieving learning goals by making the curriculum visible. *International Journal for Academic Development, 23*(3), 219-231. https://doi.org/10.1080/1360144X.2018.1462187

Williamson, C. (2017). Teachers' role in school-based assessment as part of public examinations. *US-China Education Review, 7*(6), 301-307. https://doi.org/10.17265/2161-6248/2017.06.005

Wixson, K. K., Dutro, E., & Athan, R. G. (2003). The challenge of developing content standards. *Review of Research in Education, 27*(1), 69-107. https://doi.org/10.3102/0091732X027001069

Zhao, L., Zhao, B., & Li, C. (2023). Alignment analysis of teaching-learning-assessment within the classroom: How teachers implement project-based learning under the curriculum standards. *Disciplinary and Interdisciplinary Science Education Research,* 5(13), 1-23. https://doi.org/10.1186/s43031-023-00078-1