

Redefining Assessment Standards: A Framework for Examination Guidelines in South African Basic Education

Mbalenhle H. Ngema^{1*} 

AFFILIATIONS

¹Umalusi, Quality Council for General and Further Education and Training, Pretoria, South Africa.

CORRESPONDENCE

Email: Mbalenhle.Ngema@umalusi.org.za*

Reference Formats (APA 7th Edition)

Ngema, M. H. (2024). Redefining assessment standards: A framework for examination guidelines in South African basic education. In M. L. Mokhele-Makgalwa, M. A. Mohale & T. L. Madise (Eds.), *Proceedings of the 40th AEEA Annual Conference on Reimagining Educational Assessment in the Age of Multiple Dimensions of Learning in a Global Society* (pp. 93–106). ERRCD Forum. <https://doi.org/10.38140/obp2-2024-07>

Copyright:

© The Author(s) 2024.

Published by [ERRCD Forum](#).

This is an open access article distributed under Creative Commons Attribution (CC BY 4.0) licence.



Abstract: Umalusi quality assures assessments for exit qualifications such as the National Senior Certificate (NSC) through various processes, including the evaluation of examination guidelines. The NSC is examined by three assessment bodies, and each body must develop its own examination guidelines, which must be comparable across the assessment bodies. Previous research by Umalusi identified differences in the components contained in the examination guidelines of the three assessment bodies. These differences arose from the absence of a common framework for developing examination guidelines and pose a threat to the maintenance of NSC assessment standards over time, which could undermine the credibility of this qualification. This study aimed to address this gap by developing a framework specifying compulsory components for NSC examination guidelines. Data was collected through qualitative methodology, employing document analysis and systematic literature review. Purposive sampling was used to select six countries and four subjects for evaluation; the sampled subjects were also used to pilot the framework. The find-

ings identified five compulsory components, including general information, subject-specific details, examinable content specifications and weighting, item specifications, and scoring and response specifications. The existence of a common framework is crucial for assessment bodies to produce comparable examination guidelines, ensuring the maintenance of NSC assessment standards. The study recommends that Umalusi adopt the proposed framework and use it as a standard for the development of NSC examination guidelines across assessment bodies. Furthermore, education researchers should consider conducting further research to extend this framework to other qualifications within and outside the Umalusi sub-framework.

Keywords: Assessment standards, quality assurance framework, UMALUSI, national senior certificate, examination guidelines.

1. Introduction

Umalusi, the Quality Council for General and Further Education and Training, is mandated by the *General and Further Education Quality Assurance Act No. 58 of 2001*, as amended, to quality assure assessments at all exit points of all qualifications within the General and Further Education and Training Qualifications Sub-framework (GFETQSF). One such exit point qualification is the National Senior Certificate (NSC), which serves as an access qualification for various post-schooling opportunities. The NSC is assessed by three assessment bodies, namely the Department of Basic Education (DBE), the South African Comprehensive Assessment Institute (SACAI), and the Independent Examinations Board (IEB).

Umalusi fulfils the mandate for quality assurance of assessments by setting assessment standards. In the case of the NSC, the assessment standards are prescribed by the National Curriculum and

Assessment Policy Statement (CAPS), and the examination guidelines provide clarity on the content depth and scope to be assessed.

Various researchers, such as Pawade et al. (2020), Ismail et al. (2020), Raymond and Grande (2019), AlFallay (2017), Mathur et al. (2021), Obilor and Miwari (2022), Osebhojiemen (2019), and Aziz (2021), define examination guidelines differently, but they all emphasise the importance of these guidelines in ensuring that exam question papers align with curriculum and assessment standards.

In South African Basic Education, the purpose of the examination guidelines is to guide examiners on the setting of the NSC examinations and to guide teachers in preparing learners for the examinations. Furthermore, the examination guidelines are used by moderators and Umalusi to establish the compliance of examination papers with the prescribed assessed curricula. Therefore, examination guidelines must provide clear information, and their development should be guided by a standardised framework.

Currently, each of the three assessment bodies is responsible for developing its own examination guidelines. These assessment bodies were inherited by Umalusi and had already established systems in place, each using its own framework for developing examination guidelines. As a result, there was no common framework to guide the development of the NSC examination guidelines across the assessment bodies. This absence of a common framework led to assessment bodies including different components in their examination guidelines. This variation of components raised concerns about the maintenance of NSC assessment standards over time, as inconsistent examination guidelines could undermine the credibility of the NSC. Since examination guidelines are used for the setting and moderation of NSC examination question papers, any potential compromise in their quality could negatively impact the integrity of the NSC.

The absence of a standardized framework necessitated a redefinition of Umalusi's standards, informed by empirical evidence. In its efforts to redefine the quality assurance of assessments, Umalusi has conducted several research studies to evaluate the extent to which the assessment bodies were meeting the required standards and maintaining them over time. In 2017, a study titled "A Comparison of the NSC Subject Assessment Guidelines" examined the comparability of examination guidelines across assessment bodies. Additionally, annual Post Exam Analysis (PEA) projects in 2021 and 2022 evaluated these guidelines further. These studies found inconsistencies in the components of NSC examination guidelines among assessment bodies. This challenge was persistent and long-standing, and addressing it required the undertaking of additional research. This necessitated Umalusi to investigate essential components needed in examination guidelines for the NSC subjects. The research aimed to create a standardized framework providing adequate guidance for developing examination question papers. The focus was on the NSC qualification because it provides access to various post-school opportunities in

the GFETQSF. Additionally, due to public interest in NSC results and examination standards, it is essential to ensure these standards are maintained.

This study aimed to develop a framework potentially applicable to other qualifications within and beyond the Umalusi sub-framework. It also sought to contribute to quality assurance debates and literature on assessed curricula, emphasising the role of standardized frameworks in aligning exams with established standards, particularly in the South African context.

1.1 Problem statement

Examination guidelines are among the documents used to prescribe the assessment standards for the NSC. The role of these guidelines includes guiding examiners in setting NSC examinations and assisting teachers in preparing learners for these examinations. Moreover, they are used by moderators and Umalusi to ensure that examination papers align with the prescribed curriculum and assessment standards. Since the NSC is examined by three assessment bodies, each developing its own guidelines, the examination guidelines must be comparable across these bodies. However, research conducted by Umalusi, including the Umalusi Comparison of the NSC Subject Assessment Guidelines project and the Post Exam Analysis projects in 2021 and 2022, revealed notable differences in the components of the NSC examination guidelines across various subjects and assessment bodies. This is due to the fact that each assessment body currently develops its own examination guidelines based on its own framework, resulting from the absence of a common framework. This challenge has been inherited by Umalusi, which took over already established assessment bodies. The lack of a standardized framework compromises the comparability of assessment standards over time, potentially undermining the credibility of the NSC. To address this challenge, there is a need to develop a standardized framework to guide the development of these guidelines. This framework would ensure that the NSC examination guidelines include all essential components, providing examiners, moderators, and teachers with the necessary clarity and information for effective assessment practices. Furthermore, the framework would ensure comparability of standards across assessment bodies, thereby leading to the maintenance of standards over time and ensuring the credibility of the NSC.

1.2. Research purpose

- To identify components that are essential for the development of a framework to guide the development of the NSC examination guidelines.

1.3. Research question

- What are the essential components necessary for developing a framework to guide the creation of the NSC Examination Guidelines?

2. Literature Review

Examinations aim to accurately measure the acquisition of desired knowledge and skills; they serve as effective assessment tools when thoughtfully designed (Rudolph et al., 2019).

Consequently, developing an examination necessitates a comprehensive consideration of both subject matter content and the targeted behaviour (Osebhohiemen, 2019). Furthermore, the development of an examination includes providing a marking guide with clear scoring guidance to promote consistency in the scoring of items on an examination (American Educational Research Association [AERA], et al., 2014).

Rudolph et al. (2019) underscore the importance of meticulous planning in developing well-structured examinations. This planning phase involves creating an examination question paper guided by a distinct document known by various names worldwide, such as test blueprints, test plans, tables of specifications, or test specifications. In the context of the South African GEFTQSF, the document is commonly referred to as examination guidelines. Throughout this study, the term "examination guidelines" will encompass these diverse aforementioned terms.

Various scholarly sources provide nuanced definitions of examination guidelines, shedding light on their multifaceted nature and indispensable role in ensuring assessment practices aligned with prescribed assessment and curriculum standards. According to Pawade et al. (2020), examination guidelines are intricately designed comprehensive roadmaps that meticulously incorporate and harmonise all dimensions of the curriculum, encompassing diverse domains and their contributions to the examination process.

Likewise, Ismail et al. (2020) liken guidelines to navigational tools, ensuring the inclusion of all curriculum components, thereby guaranteeing the thoroughness of assessment. Raymond and Grande (2019) assert that examination guidelines delineate the essential attributes of a test, offering a clear framework for its design. Expanding on this, AlFallay (2017) highlights that guidelines not only define the overarching structure and orientation of the test but also furnish precise specifications for each task type to be included.

Mathur et al. (2021) conceptualise examination guidelines as the creation of a template that establishes the test's content, specifying the number and types of questions across the course content while assigning relative weight to each topic in accordance with learning objectives. These guidelines furnish a systematic, multi-step approach to assessment by outlining the purpose, scope, content, and method of evaluation. In line with these perspectives, Obilor and Miwari (2022) outline four pivotal steps in test blueprint development: determining instructional objectives, content areas, item types, and preparing a three-way chart or table of specifications.

Concurring with these notions, AlFallay (2017) emphasises that examination guidelines serve the dual purpose of defining the test's overall structure and the specifics of individual task types. This comprehensive nature contributes to the creation of valid, reliable, and fair examinations. This sentiment is further supported by Osebhohiemen (2019), who emphasises that examination guidelines function as empirical tools, ensuring assessments adhere to educational objectives, thus fostering the development of valid examinations.

Aziz (2021) echoes these sentiments, underscoring the role of examination guidelines in aligning objectives with examinations, effectively allocating marks for each question, and appropriately distributing weight across different aspects. According to Abdellatif (2023), Berman et al. (2020), and Fain et al. (2019), mark allocation, time allocation, topic significance, and cognitive complexity are essential in examination guidelines to prevent the over-representation or under-representation of content in examination question papers. In essence, the range of definitions offered by various authors collectively emphasises the pivotal role of examination guidelines as a tool for strategic planning and the development of valid examination question papers. However, their development should be approached with a critical perspective, acknowledging the need for adequate details for those involved in the development of examinations.

3. Conceptual Framework

The NSC examination is a standardized high-stakes examination that is also categorised as an achievement test, as it is used to make a decision on candidates' competency in relation to the NSC qualification. According to Mamolo (2021), high-stakes examinations must prioritise validity. Therefore, this study was framed on the concept of validity, specifically examination validity. Examination validity means the exam is officially accepted and credible (Astuti, 2020). This means that an examination must be valid before trusting its results (Khan, 2019). In other words, an examination should test the skills and content outlined in the curriculum. If it measures unrelated things, it is as ineffective as using a tape measure to measure the weight of a person. Ray et al. (2018) argue that for exams to be valid, they should ensure that the measured learning objectives are a representative measure of the curriculum and are aligned with the stated objectives. Literature identifies varied forms of validity; this study focused on content and construct validity.

Content validity refers to how well an examination aligns with the material taught (Alemayehu et al., 2021). A fundamental requirement is that the examination must accurately reflect the covered content, ensuring thoroughness and inclusivity across the subject area (Alemayehu et al., 2021). Content validity is how well an examination measures its intended target and whether its items represent the entire domain it aims to assess (Adiyaa et al., 2022). It delineates a test's ability to capture a representative sample of subject matter, content, skills, and behavioural transformations under consideration (Astuti, 2020). This conception extends to the degree to which assessment elements align with and embody the intended construct for a specific assessment purpose, embodying relevance and representativeness (Yusoff, 2019). Moreover, content validity encapsulates the alignment of items with the curriculum and objectives, the adequacy of item representation for the intended objectives, the measurement of cognitive levels, and the appropriateness of item difficulty for testing purposes (Sireci & Benítez, 2023). This study drew on the aforementioned scholars' definitions of content validity.

In assessment, the construct is the specific attribute or domain targeted by the examination, excluding unintended attributes (Adiyaa et al., 2022). It is imperative that this construct is measured at an appropriate level of difficulty for the learners (Adiyaa et al., 2022). This understanding aligns with the delineation that construct signifies the knowledge and skill domain assessed in an educational test or a personal attribute gauged in a psychological test (Sireci & Benítez, 2023). Al Lawati (2023) emphasises the importance of outlining item types and providing instructions on how each item type should be assessed to enable accurate assessment of skills. This study considered definitions provided by Adiyaa et al. (2022) and Sireci and Benítez (2023) to define construct validity.

Considering content validity in this study ensured that essential details for the NSC examination guideline development covered all necessary content, guiding the creation of exam papers comprehensively. Emphasising content validity was crucial to prevent exams from deviating from their intended scope and compromising validity.

Similarly, construct validity guided the alignment of question paper development with theoretical constructs of NSC subjects. It ensured that essential details identified would lead to exam guidelines that align with the curriculum in cognitive weighting and skills assessment. Construct validity was crucial here, as neglecting it could compromise exam validity.

4. Materials and Methods

This study adopted a pragmatic research paradigm, following Maarouf's (2019) definition, which identifies three distinct forms of pragmatism. Firstly, functional pragmatism focuses on knowledge for action, wherein the purpose of knowledge is to enhance action and generate practical outcomes. Secondly, referential pragmatism emphasises knowledge about action, viewing the world through an action-oriented lens. Lastly, methodological pragmatism underscores knowledge through action, suggesting that knowledge is constructed or acquired through active engagement with the world.

Although pragmatism is often linked to mixed methods, Kaushik and Walsh (2019) emphasise that for pragmatists, the most appropriate method is the one that most effectively produces the desired outcomes of the inquiry, whether through a single method, multiple methods, or a combination thereof. Adopting a pragmatic approach was, therefore, well-suited to this research, as it informed the research design with a focus on practicality and contextual relevance. Priority was given to practicality and usefulness to optimise the effectiveness and efficiency of examination guidelines. Contextual relevance ensured alignment with the South African Basic Education goals. The principle of continuous improvement supported ongoing reflection and refinement, incorporating feedback from stakeholders, experts, and practical outcomes.

Thus, a qualitative methodological approach was chosen based on the pragmatic paradigm. The qualitative method offered flexibility in data collection and analysis, facilitating rich data

collection by enabling observation and identification of unforeseen issues that were not initially considered or included in the inception stage of the study (Mwita, 2022). Despite the subjectivity inherent in qualitative research, this study mitigated it through an analytical framework developed from academic literature obtained via a systematic literature review approach.

Document analysis served as the primary research method, offering non-obtrusive data collection (Vurayai, 2020). Purposive sampling was employed to select the countries from which examination guidelines would be sourced. Additionally, it was utilised to identify the subjects on which the framework would undergo piloting. Purposive sampling was chosen for its ability to precisely target sources providing pertinent information to address the research question and achieve objectives (Campbell et al., 2020).

Therefore, three international countries outside Africa and three African countries were selected as a sample. The inclusion criteria for countries were the examination of school leaving qualifications and the accessibility of documents. The sampled countries were Swaziland, Botswana, Namibia, Australia, the United Kingdom, and Singapore. The United Kingdom, Australia, and Singapore were chosen for their highly developed education systems, high enrolment rates, and strong academic outcomes. Furthermore, Botswana, Namibia, and Swaziland were selected because their examinations have been quality assured by Cambridge International Examinations. Four subjects were chosen, with the inclusion criteria that the subjects should have been part of the 2023 PEA, with one subject selected per organising field. The chosen subjects were English First Additional Language (FAL), Accounting, Geography, and Mathematics. The DBE NSC examination guidelines were used for piloting, as the DBE is regarded as the minimum standard for assessment bodies examining the NSC.

The data collection technique was document analysis; examination guidelines for the sampled subjects were retrieved directly from the examination council websites of the purposively sampled countries. The steps of document analysis, as described by Morgan (2022), which include determining the sample and type of documents, collecting the documents, establishing authenticity, credibility, and representativeness, and conducting the analysis, were followed to conduct the study.

A literature review on the development of examination guidelines was also used as a data collection method. A literature review was conducted to understand what academic literature considers to be the components of examination guidelines. Various electronic databases were used, and the literature studied represented various educational disciplines in different fields such as education and medical education. The review focused on literature published between 2017 and 2023. The literature review on the development of examination guidelines served as a data collection method and was guided by the research question of this study.

Various electronic databases, such as ERIC, Google Scholar, JSTOR, and SABINET, were used. The keywords employed included "examination guidelines," "assessment guidelines,"

"examination specification," "examination blueprint," "test blueprint," "test specification," "developing question papers," "developing examination guidelines," and "developing test items." These keywords were used separately to ensure the acquisition of as many academic articles as possible. Publications, including grey literature, were selected based on their relevance to the research objectives. The title and abstract of 23,200 publications were screened. The applied filter consisted of articles published from 2017 to 2023 in the field of education to obtain more pertinent information. This resulted in 746 articles, of which 721 full publications were reviewed to remove duplicates and ensure they met the eligibility criteria. The exclusion criteria for academic articles included those not discussing components of examination guidelines in the body of literature. Thus, a total of 24 studies were selected for analysis.

Framework analysis, as described by Goldsmith (2021), was applied to analyse the data obtained from academic literature. The study began with a literature review to identify key themes for an analytical framework, resulting in five main themes and thirty subthemes. These themes were then used to analyse the examination guidelines from selected countries, and new themes that emerged during the analysis were also considered. Finally, common themes and subthemes informed the development of a Framework for the development of examination guidelines for NSC subjects. The Framework for the development of examination guidelines for NSC subjects was piloted using four subjects. A pilot study, as outlined by Lowe (2019), is crucial for testing methods earmarked for a larger study. The purpose of the pilot was to strengthen the framework.

The pilot study was conducted by a panel of four subject experts, comprising a provincial Deputy Chief Education Specialist or a curriculum expert from a tertiary institution, who also served as the team leader, along with one experienced Grade 12 teacher. The panel was provided with the 2023 examination question papers, including their marking memoranda, recent NSC examination guidelines for the evaluated subject, the curriculum policy for the evaluated subject, and Umalusi evaluation instruments.

The instruments for evaluation were framed on the concept of content and construct validity, as espoused in the Framework for the development of the NSC examination guidelines and the Umalusi framework regarding question difficulty. The initial evaluation tool was a Word document containing thematic questions that gathered qualitative data. The second tool, an Excel sheet, collected quantitative data on examination guidelines and question papers to assess alignment. This approach helped gauge the effectiveness of the examination guidelines in the development of valid exam questions.

5. Presentation of Results

The findings from the analysis of the academic literature review and examination guidelines from the selected countries indicate five key categories that are essential for a framework guiding the development of examination guidelines. These categories are further divided into sub-categories for greater specificity.

The first category is general information about the examinations. It includes key details such as background information on the examination guidelines, including the intended purpose of the examination, a description and accommodation of the intended examinee population, the mode of administration, an indication of materials that are needed and will be provided to examinees, guidelines for general instructions to be provided to examinees, an outline of the structure and sections of the examination question paper, an indication of mark allocation for each section of the examination question paper, and the total amount of time allotted for the examination question paper.

The second category is subject-specific information, which refers to details that are directly related to a particular subject area. It encompasses information such as the outline of the learning objectives, the outline of the assessment objectives/specific aims, and the outline of the desired skills to be assessed in the examination guidelines.

The third category is content specification and weighting, which entails a comprehensive breakdown of the subject matter, topics, concepts, or knowledge eligible for assessment in an examination. This process also involves the allocation of marks or points to each of these content areas. Content specification and weighting include an outline of examinable content, an indication of mark allocation for each examinable topic, weighting for each cognitive level per examination question paper, and the weighting of the expected level of difficulty in the examination question paper.

The fourth category is item specification, which provides detailed guidelines or descriptions that outline the specific characteristics and requirements for individual test items (questions) within an assessment. Item specification provides specific information about how each item should be constructed, formatted, and assessed. Item specifications include an indication of item types to be included in the examination, general instructions on how each item type should be assessed, and an indication of skills assessed by each item type.

The fifth category is scoring and response specifications, which refers to the detailed instructions and criteria provided in examination guidelines for evaluating and scoring student responses to different types of items or questions within an examination. These specifications ensure a standardised and consistent approach to scoring across different examiners and examination sessions. Scoring and response specifications include general scoring guidelines for different item types, guidelines for responding to different item types, and sample marking guidelines for each item type.

These five categories, including their sub-categories, are presented as a framework in Figure 1 below.

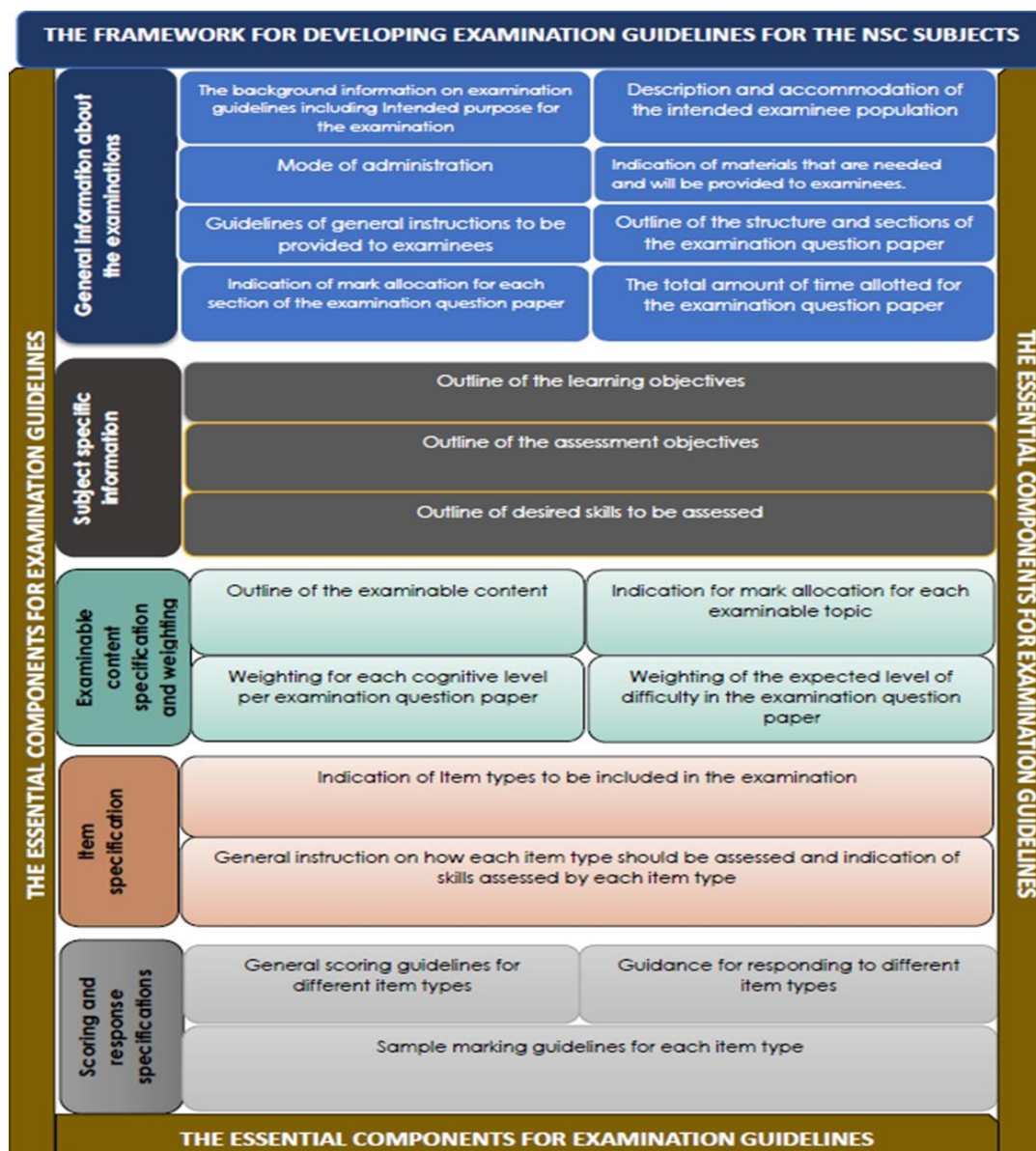


Figure 1: The Umalusi framework for development of examination guidelines for NSC subjects

6. Discussion of Findings

The findings show that examination guidelines must include the five key components listed in Figure 1. Rudolph et al. (2019) highlight the importance of these guidelines for meticulous examination planning. Osebhohiemien (2019) and Mamolo (2021) stress the need for thorough planning that considers content, targeted behaviours, and validity. These components ensure that examinations are both construct-valid and content-valid.

Primarily, the inclusion of general information in examination guidelines is essential and adheres to the principles of construct and content validity. This entails specifying the intended examinee

population, the target grade, and acknowledging potential learning barriers to tailor questions appropriately. The latter is crucial in ensuring the inclusivity of the examination question paper. Moreover, clarifying the examination's purpose ensures the accurate measurement of the intended construct, aligning with educational objectives. Furthermore, specifying the mode of administration, indicating necessary materials, and providing clear instructions ensures consistent delivery and alignment with the curriculum and assessment prescripts. Guidelines for structure and mark distribution prevent content under- or over-representation. Time allocation aligns the examination's difficulty with examinees' abilities. These findings align with the views of Adiyaa et al. (2022) and Alemayehu et al. (2021) regarding the development of valid examination question papers.

In addition, the inclusion of subject-specific information is imperative in examination guidelines and adheres to content and construct validity principles. As elucidated by Obilor and Miwari (2022), clearly defined learning or assessment objectives are vital for identifying measurable behaviours and ensuring appropriate item types. The integration of these objectives enhances alignment with learning outcomes and content. This approach aids in the development and selection of suitable assessment items for accurate assessment of skills.

Furthermore, the incorporation of examinable content specifications and weighting in examination guidelines is indispensable and aligns with content and construct validity principles. Outlining examinable content ensures a representative sample of the entire content area, reducing the risk of repetition or omission. Additionally, considering mark allocation, time allocation, topic significance, and cognitive complexity is essential to prevent the overrepresentation or underrepresentation of significant topics and to focus examinations appropriately, as emphasised by Abdellatif (2023), Berman et al. (2020), and Fain et al. (2019).

Including item specifications and weighting in examination guidelines is vital for adhering to content and construct validity principles. Outlining item types helps align the skills to be examined with the learned content, and allocating marks to different item types prevents biased sampling of assessed skills. Moreover, providing general instructions on assessing each item type clarifies examiner understanding, enabling precise skill assessment, as argued by Al Lawati (2023).

Lastly, incorporating scoring and response specifications in examination guidelines is crucial for maintaining content and construct validity principles. Scoring guidance ensures consistency among markers when evaluating similar items, thus upholding scoring standards over time. Elucidating how specific items should be responded to helps establish clear scoring rules, contributing to the development of marking guidelines that maintain comparable standards over the years. Sample marking guidelines further promote consistency in assessment practices, mitigating subjective biases, and ensuring fairness in the assessment process, as highlighted by AERA et al. (2014).

7. Conclusions and Recommendations

This study addressed the challenge of inconsistencies identified in the NSC examination guidelines across three assessment bodies by developing a standardized framework to ensure alignment with curriculum and assessment standards. The research aimed to identify essential components for examination guidelines and to use them to create a framework that guides the development of comparable NSC examination guidelines.

The study identified five essential components for inclusion in examination guidelines: general information about the examinations, subject-specific details, examinable content specifications and weighting, item specifications, and scoring and response specifications. Piloting the framework demonstrated its effectiveness in fostering consistency and validity, potentially mitigating discrepancies among assessment bodies and enhancing alignment with principles of content and construct validity.

Despite its contributions, the study faced limitations as the scope was confined to the NSC qualification, which may limit the applicability of the findings to other qualifications or educational contexts. Future research is recommended to extend the application of the Umalusi framework to additional subjects and qualifications, both within and beyond the Umalusi sub-framework, to test its robustness across diverse contexts.

For practical implementation, it is essential for Umalusi to adopt the proposed framework as a standard for developing the NSC examination guidelines. Furthermore, regular reviews of the framework are necessary to adapt to evolving educational needs; this will maintain its relevance over time.

In conclusion, this study contributes to redefining the quality assurance of assessments in South African basic education by offering a common framework for developing the NSC examination guidelines. It highlights the importance of standardized frameworks in fostering comparability in assessments, thereby laying the foundation for improved quality assessment practices.

8. Declarations

Funding: This research did not receive any external funding.

Acknowledgements: This research project was conceptualised, executed, and supervised by Ms. Mbalenhle Ngema. The report was reviewed by Dr Agness Matsie Mohale, Dr Stephen Mchunu, and Mr Emmanuel Sibanda from Umalusi, whose invaluable contributions are gratefully acknowledged. The piloting phase, which involved four subjects, was conducted by subject experts contracted by Umalusi, and their contributions are deeply appreciated.

Conflicts of Interest: The author declares no conflict of interest.

References

- Abdellatif, H. (2023). Test results with and without blueprinting: Psychometric analysis using the Rasch model. *Educación Médica*, 24(3), 100802. <https://doi.org/10.1016/j.edumed.2023.100802>
- Adiyaa, O., Appiah, K. E., & Osei-Poku, P. (2022). Achieving content validity of teacher made test in general knowledge in Art: A prerequisite assessment need for SHS teachers in Ghana. *Steadfast Arts and Humanities*, 2(1). <https://orcid.org/0000-0002-5179-2576>
- Alemayehu, W., G., Fufa, S., & Seyoum, Y. (2021). Evaluating the Content Validity of Grade 10 Mathematics Model Examinations in Oromia National Regional State, Ethiopia. *Education Research International*, 2021, 1-11. <http://dx.doi.org/10.1155/2021/5837931>
- Al Lawati, Z. A. (2023). Investigating the characteristics of language test specifications and item writer guidelines, and their effect on item development: a mixed-method case study. *Language Testing in Asia*, 13(1), 21. <https://doi.org/10.1186/s40468-023-00233-5>
- AlFallay, I. S. (2017). Test specifications and blueprints: Reality and expectations. *International journal of instruction*, 11(1), 195-210. <https://doi.org/10.12973/iji.2018.11114a>
- American Educational Research Association. (2014). *Standards for educational and psychological testing*. American Educational Research Association. https://www.testingstandards.net/uploads/7/6/6/4/76643089/standards_2014edition.pdf
- Astuti, P. (2020). Analysis of Content Validity in English Examination Test on Public Health's Students. *Journal of Industrial Engineering & Management Research*, 1(4), 100-113. <https://doi.org/10.7777/jiemr.v1i2>
- Aziz, J. (2021). *Evaluating the role of exam blueprinting as a tool to improve the exam quality and students' achievements* [PowerPoint slides]. Researchbank.
- Berman, A. I., Haertel, E. H., & Pellegrino, J. W. (2020). *Comparability of Large-Scale Educational Assessments: Issues and Recommendations*. National Academy of Education.
- Campbell, S., Greenwood, M., Prior, S., Shearer, T., Walkem, K., Young, S., Bywaters, D., & Walker, K., (2020). Purposive sampling: complex or simple? Research case examples. *Journal of research in Nursing*, 25(8), 652-661. <https://doi.org/10.1177%2F1744987120927206>
- Fain, R., Newton, W. P., & O'Neill, T. R. (2019). Creating a new blueprint for ABFM examinations. *Ann Fam Med*, 17(6), 562-4. <https://doi.org/10.1370%2Fafm.2480>
- Goldsmith, L. J. (2021). Using Framework Analysis in Applied Qualitative Research. *Qualitative Report*, 26(6), 2061-2076. <https://doi.org/10.46743/2160-3715/2021.5011>
- Ismail, M. A. A., Pa, M. N. M., Mohammad, J. A. M., & Yusoff, M. S. B. (2020). seven steps to Construct an Assessment Blueprint: A practical guide. *Education in Medicine Journal*.12(1), 71-80. <https://doi.org/10.21315/eimj2020.12.1.8>
- Kaushik, V., & Walsh, C. A. (2019). Pragmatism as a research paradigm and its implications for social work research. *Social sciences*, 8(9), 255. <http://dx.doi.org/10.3390/socsci8090255>
- Khan, M. A. (2019). Achieving the Validity of Essay Questions in the Subject of English at BA Level Examinations. *Global Language Review (GLR)*, IV (1), 55-59. [http://dx.doi.org/10.31703/glr.2019\(IV-I\).07](http://dx.doi.org/10.31703/glr.2019(IV-I).07)
- Lowe, N. K. (2019). What is a pilot study? *Journal of Obstetric, Gynecologic & Neonatal Nursing*, 48(2), 117-118. <https://doi.org/10.1016/j.jogn.2019.01.005>
- Maarouf, H. (2019). Pragmatism as a supportive paradigm for the mixed research approach: Conceptualizing the ontological, epistemological, and axiological stances of pragmatism. *International Business Research*, 12(9), 1-12. <http://dx.doi.org/10.5539/ibr.v12n9p1>

- Mamolo, L. A. (2021). Development of an Achievement Test to Measure Students' Competency in General Mathematics. *Anatolian Journal of Education*, 6(1), 79-90. <http://dx.doi.org/10.29333/aje.2021.616a>
- Mathur, M., Verma, A., Mathur, N., Kumar, D., Meena, J. K., Nayak, S., ... & Parmar, P. (2023). Blueprint designing and validation for competency-based curriculum for theory assessment in community medicine. *Medical Journal Armed Forces India*, 79, S47-S53. <https://doi.org/10.1016/j.mjafi.2021.10.003>
- Morgan, H. (2022). Conducting a qualitative document analysis. *The Qualitative Report*, 27(1), 64-77. <https://doi.org/10.46743/2160-3715/2022.5044>
- Mwita, K. (2022). Strengths and weaknesses of qualitative research in social science studies. *International Journal of Research in Business and Social Science (2147-4478)*, 11(6), 618-625. <https://doi.org/10.20525/ijrbs.v11i6.1920>
- Obilor, E. I., & Miwari, G. U. (2022). Content Validity in Educational Assessment. *International Journal of Innovative Education Research*, 10(2), 57-69.
- Osebhohiemen, E. (2019). Use of Table of Specification in Construction of Teacher-Made Achievement Test in Mathematics in the Primary and Secondary Schools. *The Melting Pot*, 5(2), 1-15.
- Pawade, Y. R., Mehta, S., Mahajan, A. S., Patil, S. N., Barua, P., Desai, C., & Supe, A. N. (2020). 'Blueprinting in assessment'—an online learning experience. *South-East Asian Journal of Medical Education*, 13(2), 77-85. <https://doi.org/10.4038/seajme.v13i2.214>
- Ray, M. E., Daugherty, K. K., Lebovitz, L., Rudolph, M. J., Shuford, V. P., & DiVall, M. V. (2018). Best practices on examination construction, administration, and feedback. *American journal of pharmaceutical education*, 82(10), 7066. <https://doi.org/10.5688%2Fajpe7066>
- Raymond, M. R., & Grande, J. P. (2019). A practical guide to test blueprinting. *Medical teacher*, 41(8), 854-861. <https://doi.org/10.1080/0142159x.2019.1595556>
- Rudolph, M. J., Daugherty, K. K., Ray, M. E., Shuford, V. P., Lebovitz, L., & DiVall, M. V. (2019). Best practices related to examination item construction and post-hoc review. *American journal of pharmaceutical education*, 83(7), 7204. <https://doi.org/10.5688%2Fajpe7204>
- Sireci, S., & Benítez, I. (2023). Evidence for Test Validation: A Guide for Practitioners. *Psicothema*, 35(3), 217-226. <https://doi.org/10.7334/psicothema2022.477>
- Vurayai, S. (2020). Rurality and exclusion in ordinary level mathematics in Zimbabwe: A document analysis. *International Journal of Learning, Teaching and Educational Research*, 19(6), 370-386. <http://dx.doi.org/10.26803/ijlter.19.6.22>
- Yusoff, M. S. B. (2019). ABC of content validation and content validity index calculation. *Education in Medicine Journal*, 11(2), 49-54. <http://dx.doi.org/10.21315/eimj2019.11.2.6>

Disclaimer: The views, perspectives, information, and data contained within all publications are exclusively those of the respective author(s) and contributor(s) and do not represent or reflect the positions of ERRCD Forum and/or its editor(s). ERRCD Forum and its editor(s) expressly disclaim responsibility for any damages to persons or property arising from any ideas, methods, instructions, or products referenced in the content.