

## Specifying expectancies for final summative assessment: lessons from international exam guideline practice

Dr Celia Booyse

Umalusi (Council for Quality Assurance in General and Further Education and Training),  
South Africa

[celia.booyse@umalusi.org.za](mailto:celia.booyse@umalusi.org.za)

+27 12 030 0960

+27 83 782 2611

### Abstract

In South Africa, numerous formative assessments culminate in a set of examinations (summative assessment) that provide 75% of a learner's results at Grade 12 level. Given the crucial nature of such summative assessments for further study or employment, it is essential to ensure that decisions that are made about learners' competencies are based on the best possible information. Every examination is informed by a guideline about how it needs to be constructed, and what content and skills need to be assessed. As part of its mandated responsibilities to ensure the quality and consistency of school-level assessment in South Africa, Umalusi, the Council for Quality Assurance in General and Further Education and Training undertook research to establish how top-performing school systems in the world specify expectancies in examination guidelines. The study aimed at understanding how current practice in South Africa could be enhanced. The selection of countries surveyed in the study included Finland, Singapore, Canada, Kenya and Ghana. Most examination guidelines surveyed provide a high level of specification, detailing the kinds of claims that can be made about learners' skills, based on their performance in the assessment. The paper provides for specific examples of how expectancies need to be specified to enable better predictions of learner performance. The insights are summarized in a template proposed for a detailed specification of the final assessment. (*Words: 219*)

**Keywords:** examination guidelines; expectancies; specification; performance prediction; assessment; reliability

## **Specifying expectancies for final summative assessment: lessons from international exam guideline practice**

### **1. Background and introduction**

In South Africa, numerous formative assessments culminate in a set of examinations (summative assessment) that provide 75% of a learner's results at the Grade 12 exit level. Umalusi, the Council for Quality Assurance in General and Further Education and Training, is legislated to quality assure assessment at exit points for qualifications registered on the General and Further Education and Training Qualifications Sub-Framework of the South African National Qualifications Framework (NQF). The organisation quality assures assessment offered by all assessment bodies that assess the National Senior Certificate (NSC), the schooling qualification at Grade 12 level. The moderation of question papers is one of the processes undertaken during quality assurance. Moderators are required to determine the quality and standard of examination papers by inferring the level of adherence to curriculum, policy requirements and standards in setting examination questions. Each assessment body develops its own examination or assessment guidelines to determine the subject specifications for examination question papers.

Umalusi has previously undertaken investigations to compare the standard of the assessment guidelines between the three NSC assessment bodies to determine comparability for each subject offered for the National Senior Certificate. Since policy dictates that the Department of Basic Education (DBE) represents the minimum standard, it was important to understand whether the other two assessment bodies do indeed comply with this stipulation, and meet or exceed this standard.

Overall, the levels of comparability were good, but the comparative work revealed certain weaknesses in the overall design of the assessment guidelines that has prompted Umalusi to investigate international practice around the construction of such guidelines. This comparative work resulted in a report that detailed assessment design principles in the following comparator nations:

- Finland
- Singapore
- Ontario State, Canada
- Kenya
- Ghana

The findings of the investigation showed that, while the South African (SA) assessment guidelines are broadly comparable with those of the other nations in the comparison group, the SA guidelines were significantly less detailed than those of the comparators. The lessons learned from the international investigation provided insight on how generic assessment guidelines could provide more detailed and reliable guidance to examiners and moderators.

### **2. Research design and methodology**

The investigation was qualitative in nature and included a literature review, web search and document analysis. Documents on the country systems were analysed to determine curriculum and assessment standards, intended learning and kinds of assessment. The main

features of the examination guidelines were established to determine the requirements of well-formulated subject assessment guidelines.

### 3. Literature review

Public examinations are an essential feature in education systems globally, stemming from the idea that examination performance forms the basis of decisions about the future of candidates (Greaney & Kellaghan, 1995). Livingstone (1988) calls for careful consideration when implementing large-scale examinations. Consideration should be given to -

- available resources;
- what is practically possible within the constraints of the system;
- curriculum coverage in the assessments; and
- adequate exploration of the causes of changes, which are observed from year to year.

When considering an examination system it is important to consider the overall plan of implementation and the manner in which the system will be evaluated. When evaluating the system, considering the technical quality of the system is critical. According to Linn (1994), these sources can be characterised by the following main areas:

- a. *Content analysis*: Examinations need to be consistent with national content standards.
- b. *Fairness*: Fairness towards all population groups is essential in the evaluation of an examination system. Disproportionate failure rates, misinterpretations of between group results and unfair access as a result of these will not go away by applying a different form of assessment. Thus, both screening for bias and the creation of a common standard are required.
- c. *Impact analysis*: Here the consequences of the examination as well as the magnitude and nature of its unintended consequences need to be interrogated. The impact of the examinations on the schools has to be evaluated.
- d. *Ability to generalise*: It is essential to show that the results of the examination are dependable, across markers, time of marking, type of tasks and context of the examination.
- e. *Comparability*: The various assessments need to be comparable to the national standard. The manner in which assessments are compared can be undertaken using statistical equating, statistical moderation and consensus moderation.

Figure 1 below provides an overview of the process required within an examination system.

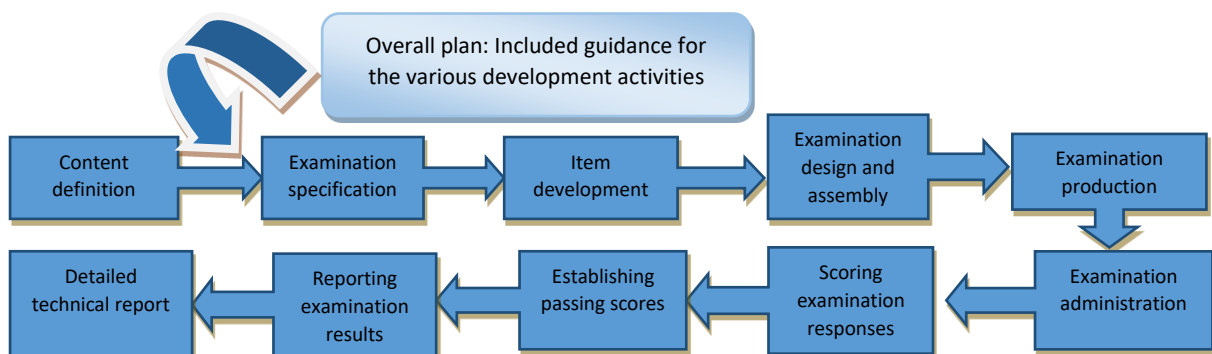


Figure 1: Examination process (adapted from Downing & Haladyna, 2006)

All examination systems across the world follow similar processes. At the very beginning is the identification of content to be examined, and the establishment of examination specifications. Key in this process are the subject assessment guidelines, which are used to guide teachers, examiners and moderators. The first and second steps in the examination are the most important as these speak to the content and examination specification.

The concepts of subject assessment guidelines and test specifications exist on the same continuum. Ultimately, assessment guidelines are always a form of test specification, and the differences relate to the level of detail envisioned in the assessment design document.

*“Test specifications are iterative, generative blueprints of test design. They are written at the item level, and allow test developers or item writers to produce new versions of a test for different test-taking populations. The specs serve as guidelines so that new versions can be compared to previous versions. Specs may detail the properties of individual items or the knowledge, skills, and abilities that are encoded in specific items.”* (Owen, 2018).

Given Owen’s definition (2018) above, it was important for Umalusi to consider the ways in which the South African assessment guidelines differ from test specifications. Taking the definition point-by-point, it became clear where the SA guidelines deviate from the conception of a full test specification.

<b>Feature</b>	<b>Test Specification</b>	<b>Subject Assessment Guideline</b>
<i>Test specifications are iterative, generative blueprints of test design</i>	Yes	Yes
<i>They are written at the item level, and allow test developers or item writers to produce new versions of a test for different test-taking populations</i>	Yes	No
<i>The specs serve as guidelines so that new versions can be compared to previous versions</i>	Yes	No
<i>Specs may detail the properties of individual items or the knowledge, skills, and abilities that are encoded in specific items</i>	Yes	No

As a system, South African schools are orientated toward the Grade 12 exit-level examinations and teaching is often adjusted towards this goal. Though assessment guidance is provided as Section 4 of the Curriculum and Assessment Policy Statement (CAPS) document, the question still stands what *additional* guidelines are needed for assessment. Furthermore, how the guidelines would ultimately adequately prepare candidates for the Grade 12 examinations and how are the guidelines captured in the examination process?

#### **4. Findings from the review of the country systems**

It was evident from the review of the country education systems of Finland, Singapore, Ontario State (Canada), Kenya and Ghana that curriculum standards and assessment play a significant role in establishing new expectations for learning (Binkley et al., 2012). There are several main findings from the review of the literature:

- i. Every country issues subject examination guidelines for key exit points in the system.
- ii. The level of description included in the subject examination guidelines differs substantially.
- iii. The examination guidelines typically focus on the structure of the examination and there is a discussion of what the papers per subject entail.
- iv. Cognitive load and item difficulty are not included in all guidelines reviewed, but the type of items and weighting per section are typically included.
- v. The subject guidelines across the subjects tend to be constructed in the same way. However, some variations are observed in terms of additional information being provided. For example, for mathematics and science, notation and formulas may be provided.

Guidelines from **Finland** provide an overview of the system and the general regulations. The mark allocation is provided as guidelines on how the examinations are marked and processed. The type of questions typically included in the examinations are described. Thus, items such as multiple-choice questions and open-ended questions are explained. Examples of how marks are allocated, also provide information regarding correct, partially correct or incorrect answers. The guidelines focus mainly on how teachers should go about marking the examination. The guidelines are very detailed, and are regularly updated. There is no engagement regarding how the examinations are constructed regarding cognitive load or taxonomy levels. This is perhaps not surprising given the underlying philosophy of the Finnish education system that the role of assessment is to guide the learning process and to support motivation. Feedback is crucial in the assessment process as this is how the learners develop a conception of themselves as learners (Kumpulainen & Lankinen, 2016). It was noted that there *are* detailed guidelines regarding what are appropriate responses and how these should be graded which function as a guide to cognitive load.

The subject assessment guidelines from the **Singaporean education system** typically include the aims of the teaching syllabus. The assessment objectives per subject are elaborated on, and the scheme of the assessment includes the type of question, time and mark allocation as well as weighting. A description of the paper is also typically provided. There are minor differences across subjects, and subject-specific information, for example, formulas and mathematical notation are included in the mathematics guideline.

Having moved away from the use of norm-referenced assessment practices, assessment *for* learning is espoused by the **Ontario Ministry of Education** (2010: 18). The manner in which candidates receive a certificate at the end of study is by accumulating a certain number of credits over 2-3 years. Given the focus on assessment for learning, formative assessment plays an important part and the assessment system used, is based on an assessment framework that identifies four levels of performance.

The **Kenyan National Examinations Council** publishes guidelines and instructions for examinations. The guidelines provide a clinical picture of the examination process. A timetable and the general rules and regulations for the examination are provided. Guidelines are provided to both learners and teachers. Caution regarding when regulations are contravened is also provided. Kenya has recently embarked on substantial curriculum reform and the reform initiatives are still in process of being rolled out, so this information may be outdated soon.

Subject syllabus guidelines are provided in the **Ghanaian education system**, with the examination regulations and rules captured separately. The guidelines offer an outline of topics learners should cover for the exam. The guidelines follow a common structure and focus exclusively on the type of questions that can be expected per paper. Mark allocation is included. The type of cognitive demand or the ratio between easy, medium and difficult questions is not discussed.

The subject assessment guidelines issued by the **South African** Department of Basic Education provide both general and subject-specific guidance on how to design assessments for each subject. The subject assessment guidelines provide clarity regarding the depth and scope of content to be assessed. The guidelines also aim to assist teachers to prepare learners for the exit-level examinations. Currently, the examination guidelines typically include the following elements although these may vary somewhat per subject:

- Time and mark allocation
- The format of the examination questions regarding intentions of the paper should be stipulated, e.g. the scope of content and skills, taxonomy levels and whether basic skills or application are assessed
- Context to help develop learner ability
- Weighting of topics
- Distribution of marks according to taxonomy levels
- Taxonomy levels

The difficulty identified with the subject assessment guidelines in South Africa is that the level of detail required is not always sufficient to provide the guidance required to examiners, moderators, and teachers who use information about the examinations as preparatory tools in class. This paper takes the view that in order to enhance the validity and reliability of the National Senior Certificate (NSC) examinations, it is important to address the lack of detail in the assessment guidelines.

## **5. Recommendations**

What was evident from the education systems reviewed, is that the guidelines ought to be consistent across subjects and include the same level of description. It is also suggested that the subject assessment guidelines are regularly updated.

The difficulty with the subject assessment guidelines in the South African context is that the level of detail offered does not always provide the guidance required. Discrepancies between curriculum documents and the subject assessment guidelines are problematic. Subject assessment guidelines are also not updated as regularly as they are in other systems. Based on the subject guidelines reviewed from the various countries, a generic template has been developed which can be adapted for the various subjects.

The assessment objectives must be clearly stated, per paper, and drawn directly from curriculum policy documents. The time allocation for the paper overall and per section must be included. The number of marks allocated per section and the weighting must be balanced with the time allocated. The taxonomy used in the subject must be described within the context of the subject and not just mentioned. The mark distribution per taxonomy level must be discussed as must the difficulty of items to be included per level. The guidelines should conclude with example questions per taxonomy level in addition to what would constitute a correct response, partially correct response and an incorrect response.

The following structure is suggested for the South African Assessment Guidelines for examinations:

<b>Introduction</b>	Refers to the purpose of the assessment guidelines. Provides clarity on the depth and scope of the content to be assessed.
<b>Assessment objectives</b>	In this section, the major aims and objectives from the curriculum should be included so that it is clear that candidates will be assessed on their ability to demonstrate the knowledge and skills required.
<b>Time allocation, marks and weighting</b>	In this section of the guideline document the time allocation per question, number of marks and weighting has to be specified.
<b>The scope of content and skills</b>	In this section, a description of the taxonomy used should be included. What the various taxonomy levels mean should be adequately described (similar to the example presented in the Canadian system). It is imperative that the descriptions are explicit and included in the guidelines as opposed to assuming that educators will refer to policy documents.
<b>Distribution of marks and difficulty per level</b>	In this section, information is presented regarding what is understood regarding difficulty in terms of content, stimulus, task and expected response. The taxonomy levels, marks, difficulty and weighting per paper should be included.
<b>Example questions per level</b>	In this section of the document, examples of questions per taxonomic level should be included. The taxonomic level is explained more fully, and examples per level and what would constitute a correct response, partially correct response and incorrect response should be included.

## 6. Concluding remarks

The international investigation revealed that no nation surveyed provided full test specifications for national examinations. From the investigation, it became clear that the level of detail required for a standardised psychometric test is not entirely in line with what is required for an examination guideline. Where a test specification will call for measurement of a construct with a singular focus, an examination guideline is concerned with measuring content knowledge as well as an overall construct. There are also practical concerns surrounding a test specification in that the level of detail required in such a document can make it unwieldy for use by content specialists in the absence of assessment specialists.

It was found that the current South African assessment guidelines are too generically stated in certain areas, which could allow for unpredictable variation between one examination and another. The outcome of the investigation suggested something of a middle-ground between the current assessment guidelines in the South African context and full test specifications to provide for the specificity, but also enough agility to produce regular and bespoke assessments year after year in a large variety of content domains.

## 7. References

Adu-Gyamfi, S., Donkoh, W.J., & Addo, A.A. (2016). Educational reforms in Ghana: Past and present. *Journal of Education and Human Development*, 5 (3,) 158-172

Binkley, M., Erstad, O., Herman, J., Raizen, S., Ripley, M., Ricci M.M., & Rumble, M. (2012). Defining 21st century skills. In P Griffin & E Care (Eds), *Assessment and teaching of 21st century skills*. New York: Springer.

Britschgi, V. (2014). The Finnish examination system. Retrieved from [https://www.ylioppilastutkinto.fi/images/sivuston\\_tiedostot/Kehittaminen/YTL\\_presentation\\_English.pdf](https://www.ylioppilastutkinto.fi/images/sivuston_tiedostot/Kehittaminen/YTL_presentation_English.pdf)

Council of Ministers of Education. (2008). Education in Canada. Retrieved from: <https://www.cmec.ca/Publications/Lists/Publications/Attachments/64/Education-in-Canada2008.pdf>

Downing, S.M., & Haladyna, T. M. (2006). *Handbook of test development*. New Jersey: Lawrence Erlbaum Associates.

Earl, L. E. (1995). Assessment and accountability in education in Ontario. *Canadian Journal of Education* 20(1), 45-55.

East African Community. (2014). Draft harmonised curriculum structures and framework for the East African community. Retrieved from: <https://kicd.ac.ke/curriculum-reform/curriculum-structure-and-regulations/#draft5>

EducationCloset. (2016). What is STEAM. Retrieved: <https://educationcloset.com/steam/what-is-steam/>

Eurydice. (2015). Teaching and learning in a single structure education. Retrieved: [https://webgate.ec.europa.eu/fpfis/mwikis/eurydice/index.php/Finland:Teaching\\_and\\_Learning\\_in\\_Single\\_Structure\\_Education](https://webgate.ec.europa.eu/fpfis/mwikis/eurydice/index.php/Finland:Teaching_and_Learning_in_Single_Structure_Education)

Finnish National Board of Education. (2018). Basic education. Retrieved from: [http://www.oph.fi/english/curricula\\_and\\_qualifications/basic\\_education](http://www.oph.fi/english/curricula_and_qualifications/basic_education)

Galloway, A. (2008). The assessment systems of Finland and Queensland. Retrieved from [https://www.sqa.org.uk/files\\_ccc/PNP\\_ResearchReport1\\_TheAssessmentSystemsFinlandQueensland.pdf](https://www.sqa.org.uk/files_ccc/PNP_ResearchReport1_TheAssessmentSystemsFinlandQueensland.pdf)

Gay, L. R., & Airasian, P. (2003). *Educational Research: Competencies for analysis and application* (7th ed.). New Jersey: Merrill Prentice Hall.

Ghana Nursing & Health Center. (2011). Key points of Ghana's new education system. Retrieved from: [https://docs.wixstatic.com/ugd/6753e1\\_2fb231d4f4de81f1d47c455c65f4db1c.pdf](https://docs.wixstatic.com/ugd/6753e1_2fb231d4f4de81f1d47c455c65f4db1c.pdf)

Ghanaweb. (2018). Education Ghana. Retrieved from <https://www.ghanaweb.com/GhanaHomePage/education/>

Giger, M. (2002). A word about Ghanaian schools. Retrieved from: <http://www.gigers.com/matthias/engmala/teachtlk.htm>

Greaney, V., & Kellaghan, T. (1995). *Equity issues in public examinations in developing countries*. Washington: The World Bank.



- Kauppinen, J. (2016). Curriculum in Finland. Retrieved from: [http://dge.mec.pt/sites/default/files/Noticias\\_Imagens/1\\_curriculum\\_in\\_finland.pdf](http://dge.mec.pt/sites/default/files/Noticias_Imagens/1_curriculum_in_finland.pdf)
- Kellaghan, T., & Greaney, V. (1992). *Using examinations to improve education*. Washington: The World Bank.
- Kenyan Institute of Curriculum Development. (2016). Basic Education Framework. Retrieved from <http://www.education.go.ke/index.php/downloads/file/315-basic-education-curriculum-framework>
- Kenyan National Examinations Council. (2018). Examination timetable and instructions. Retrieved from: <https://www.knec.ac.ke/home/index.php?...examination...a-instructions>.
- Killen, R. (2000). *Teaching strategies for outcomes-based education*. Lansdowne: Juta & Co.
- Khoza, S.B. (2015). Student teachers' reflections on their practices of the curriculum and assessment policy statement. *South African Journal of Higher Education*, 29(4), 179–197.
- Kotzé, G. S. (2002). Issues related to adapting assessment practices. *South African Journal of Education*, 22(1), 76-80.
- Kumplainen, K., & Lankinen, T. (2016). Striving for educational equity and excellence: Evaluation and assessment in Finnish basic Education. In Niemi, H, Toom, A & Kallioniemi, A. *Miracle of education: The principles and practices of teaching and learning in Finnish schools*. Rotterdam: Sense Publishers.
- Linn, R. L. (1994). Evaluating the technical quality of proposed national examination systems. *American Journal of Education*, 102 (4), 565-580.
- Livingston, I. D. (1988). Standards, national: Monitoring. In Husen, T & Postlethwaite, T.N. (Eds), *The International Encyclopedia of Education* (pp. 4786-4791). New York: Pergamon Press.
- McMillian, J. H. (2001). *Essential assessment concepts for teachers and administrators*. California: Corwin Press.
- Ministry of Education. (2018). About us. Retrieved from: <http://moe.gov.gh/index.php/about-us/>
- Ministry of Education and Culture. (2017a). Finnish Education in a nutshell. [http://www.oph.fi/download/146428\\_Finnish\\_Education\\_in\\_a\\_Nutshell.pdf](http://www.oph.fi/download/146428_Finnish_Education_in_a_Nutshell.pdf)
- O'Donoghue, R. (2013). *Framing Active Teaching and Learning in CAPS. Fundisa for Change Programme*. Environmental Learning Research Centre, Rhodes University, Grahamstown, and GIZ Expert Net.
- Ontario Ministry of Education (2018). The Ontario Curriculum. Retrieved from: <http://www.edu.gov.on.ca/eng/curriculum/elementary/common.html#display>
- Our Kids. (2014). The Canadian education system. Retrieved from: <http://www.ourkids.net/education-in-canada.php>

Owen, N. (2018). Test Specifications. In *The TESOL Encyclopedia of English Language Teaching* (eds J. I. Lontas, T. International Association and M. DelliCarpini).

Nuffic. (2015). The Kenyan education system described and compared with the Dutch system. Retrieved from <https://www.nuffic.nl/en/publications/find-a-publication/education-system-kenya.pdf>

Nuffic. (2016). The Canadian education system described and compared with the Dutch system. <https://www.nuffic.nl/en/publications/find-a-publication/education-system-canada.pdf>

Nuffic. (2016). The Ghanaian education system described and compared with the Dutch system. Retrieved from: <https://www.nuffic.nl/en/publications/find-a-publication/education-system-ghana.pdf>

Reddy, C. (2004). Assessment principles and approaches. In J. G. Maree & W. J. Fraser (Eds.), *Outcomes-based assessment* (pp. 29-44). Sandown: Heinemann Publishers.

Singapore Assessment and Examinations Board. (2017a). Introduction. Retrieved from: <https://www.seab.gov.sg/pages/about/introduction>

Singapore Assessment and Examinations Board. (2017b). General information. Retrieved from: [https://www.seab.gov.sg/pages/nationalExaminations/PSLE/general\\_information](https://www.seab.gov.sg/pages/nationalExaminations/PSLE/general_information)

Singapore Ministry of Education. (2018a). Education system. Retrieved from: <https://www.moe.gov.sg/education/education-system>

Singapore Ministry of Education. (2018b). Compulsory education. Retrieved from: <https://www.moe.gov.sg/education/education-system/compulsory-education>

Singapore Ministry of Education. (2018c). Desired outcomes of education. Retrieved from: <https://www.moe.gov.sg/education/education-system/desired-outcomes-of-education>

Singapore Ministry of Education. (2018d). 21st Century competencies. Retrieved from: <https://www.moe.gov.sg/education/education-system/21st-century-competencies>

The Chalk Face. (1999). Transformational OBE: What is it? Retrieved from January 23, 2003, from [http://hagar.up.ac.za/catts/learner/1999/moore\\_aj/e-cv/portfolio/projects/tech/newsletters/obe1.doc](http://hagar.up.ac.za/catts/learner/1999/moore_aj/e-cv/portfolio/projects/tech/newsletters/obe1.doc).

Umalusi. (2017). A comparison of NSC subject assessment guidelines. Pretoria: Umalusi.

WAEC. (2015). Regulation, schemes of examination and sample questions. Retrieved from: <https://registration.ghanawaec.org/bece/docs/handbook.pdf>

WAEC. (2018). WAEC Syllabus 2017. Retrieved from <https://waecdirect.blogspot.co.za/2017/02/waec-syllabus-for-all-subjects-phones-ssce-gce-pdf.html>

Wanjohi, A.M. (2017). New education system in Kenya. Retrieved from: <http://schoolsnetkenya.com/new-education-system-in-kenya-an-excerpt-from-basic-education-curriculum-framework/>