

**The matriculation examination:
How can we find out if standards are falling?**

By

Mbithi wa Kivilu

24th November, 2004

Introduction

This paper provides an overview of the current debates about standards of the Senior Certificate in South Africa. It endeavours to answer the question of whether standards are falling by providing both local and international research findings on the performance of South African learners. The paper then provides an explanation of what educational standards are and reviews the major methods used in setting standards. Equivalence of standards across examination bodies could be assessed through comparisons of examination results using a variety of methods. The paper will also provide an overview of some of the methods used in comparing standards across examination bodies. The paper concludes by offering informed opinion on the issues discussed and by asking further questions on this debate of educational standards.

The 2003 Senior Certificate Examination (SCE) results continue the upward trend, which began in 2000 with 73% of the candidates passing. This represents a four percentage point increase on the 69% pass rate achieved in 2002 and a 15 percentage point increase on the 58% pass rate achieved in 2000 (Shindler, 2004). The percentage of candidates who passed with endorsement improved from 15% in 2001 to 19% in 2003. The improvement in pass rate has been perceived by sceptical public as a drop of standards of the Senior Certificate Examination, popularly known as the matric examination. Being the only nationally accepted measure of learners' performance, the quality of teaching and learning, and how well the system is doing, it is critical that the senior certificate examination enjoys public confidence.

Another critical question that we should ask is whether the disparities in performance that existed some years ago among race groups have narrowed or have been eliminated completely. An analysis of education statistics for 1997 and the matriculation results for 1999

and 2000 by Van der Berg (2004) shows that there are large inequalities in results between the different provinces and massive differentials between the poorest schools as measured by amount of school fees, (with an average pass rate of 44 per cent) and the richest schools (97 per cent), and the predominantly African (43 per cent) and predominantly white schools (97 per cent). According to Van Berg while only three of the 179 mainly white schools had pass rates below 80 percent (the lowest pass rate was 68 per cent), most predominantly African schools had pass rates below 60 per cent. The continued disparities in matric performance across income groups and race have serious implications for the country as a whole. For example, the productive capacity of the economy will be seriously constrained, labour market inequalities will persist and pressure will grow for more interventions in the labour market.

The debates about falling standards

Higher education institutions and employers in South Africa complain of the low level of skills of the students graduating from Grade 12. The main complaint is centred on the lack of basic skills in literacy and mathematics that are considered core to further training or employability. The critical question to ask is whether as a country we have well articulated standards (or benchmarks) against which the performance of our learners can be assessed. Can the public and especially the key stakeholders in education claim to have a thorough understanding of such standards? Do the educators and parents understand what their learners are expected to achieve at various levels of the education system?

Educational standards are articulated in the National Curriculum and various Government White papers. But when there exists no common public understanding of what standards our learners are expected to achieve then there will be confusion about whether the standards are

improving or falling. What is not very clear in these debates is the nature of the standards in question. Are these standards the proportion of candidates who achieved a pass, or those who have received endorsement for university entrance? Are they the proportion of candidates who have mastered literacy and numeracy to levels acceptable to tertiary institutions and employers? Until such time that we have consensus on the nature and definition of standards, they will remain in the eye of the beholder as Linda Chisholm states in her article in the *Financial Mail* of 9th February 2004 entitled “Matric pool redefined”.

The debate about standards of the Matriculation Examination in South Africa has reached alarming levels with institutions of higher learning such as universities and technikons starting to administer their own entrance examination to protect themselves against what they consider to be poor quality matriculants and to raise standards. Concerned with the quality of the matriculants entering higher education research studies have also been conducted to test the competency of first year students through the Higher Education Admissions Projects by the South African Universities Vice-Chancellors Association and the Committee of Technikon Principals. One such study conducted in the University of Cape Town to assess basic mathematics competency found that out of the 322 students who took the test, based on content covered in the Grade 11 mathematics syllabus, 30% failed, scoring under 49%, and about 20% scored between 50-59%. The findings of this study reveal a mismatch between the students’ competencies and the high matric marks they obtained to get admission to the university (Naidoo 2004). Critics of the matric examination may argue that the students’ marks were possibly inflated to improve their grades and to push up the country’s overall pass rate.

A similar study conducted by the Tertiary Education Linkages Project (TELP) also produced alarming results from tertiary students. The diagnostic test, which is equivalent to the USA Scholastic Achievement Test, assessed students' performances in English, Mathematics, and Science at several universities and technikons in 2003. The results showed that about 90% of the students failed in mathematical knowledge, scoring less than 50%. It is evident that students graduating from grade 12 are not performing at the expected levels and lack the competencies required to successfully complete a tertiary institution's programme.

Umalusi (2004) conducted a study to investigate the standard of the senior certificate examination by looking at the examination papers, memoranda and marking, statistical moderation process, language compensation and rules and procedures. The results showed that there was evidence of declining levels of conceptual challenge in standard grade papers, and more learners enrolled in 'easy' standard grade rather than higher grade leading to an increase in overall pass rate. The study also raised some concerns about the quality of examination setting, marking and moderation. Based on the findings it was recommended that an investigation be conducted on the cumulative effect of statistical moderation process on pass rates.

With the growth and increasing interdependence of the global economy, policymakers and educationists turn to international comparisons to assess how well national systems of education are performing. These comparisons shed light on a host of policy issues, from access to education and equity of resources to the quality of school outputs. They provide the policymakers with benchmarks to assess their systems' performances, and to identify potential strategies to improve student achievement and system outputs. One such effort by the National Department of Education in 1999 saw the setting up of a commission to

benchmark the South African Senior Certificate with the Scottish Higher Grade examination in order to assess the comparability of our education standards and the quality with international standards. While the results confirmed the high quality, validity and reliability of the Senior Certificate, the content of the South African curriculum and the level of demand on the learners in some content subjects were found to be inadequate compared to the Scottish standards (Umalusi, 2004).

One well-known international study in which South Africa participates in that compares countries worldwide is the Trends in International Mathematics and Science Study (TIMSS). The TIMSS (1999) ranks South Africa last in proficiency levels. About 69% of South African students who took part in the study did not achieve the lower-quarter benchmark. This suggests that a number of under prepared students are entering higher education, resulting in higher failure rates and low throughput rates. Questions have also been raised about the employability of the majority of learners who leave Grade 12 with a Senior Certificate.

The foregoing discussion points to one thing, that South Africa's school system does not prepare matriculating students adequately to cope with the pressures and rigours of higher education studies. There are two possible explanations for this; either the South African education system is not well synchronized at the various levels so that skills mastered at a lower level can enable students to cope with the academic demands at the next level, or standards are simply falling.

Are falling standards a justification for scrapping of senior certificate examination?

There have been voices from the public calling for the complete scrapping of the senior certificate examination. The National Department of Education has already unveiled its plan to phase out matric by the year 2006. But the public poorly understands the reasons for, as well as the character and implications of, these changes. Scrapping the examination would in some way confirm the public fears that it has not been serving any public good. Public discussion of the merits and demerits of the current system, as well as of the system that is to replace it, could serve an important educational purpose (Chisholm, 2004).

The debate about standards in education and calls to scrap public examinations are not unique to South Africa. Most recently, The Times Educational Supplement of 22 October 2004 (<http://www.tes.co.uk/2040906>) in Britain reported that a committee headed by Mr Mike Tomlinson, a former Chief Inspector of Education, has proposed that Britain shelve the 60-year-old examination-based system in favour of a vocational and skills-oriented structure. Should these radical proposals be approved, Britain will also drop the General Certificate of Secondary Education and A levels within 10 years. Reacting to the content of the report the Prime Minister Tony Blair argues that rather than scrapping the GCSE and A-levels, the current system should instead be strengthened with more emphasis on quality rather than on structure. Business leaders and employers in Britain have also raised questions about the proposed system with the head of the Confederation of British Industry saying that for the reforms to win the confidence of employers, they must raise standards and not simply change structures. What matters, according to the Director General of the confederation, is not the examination system Britain has, but what young people are able to achieve.

The debates evoked by the Tomlinson Report has important lessons for South Africa as we try to grapple with the issue of standard of senior certificate examination and its role in the

public education system. The critical question that we should address is how can high standards of education in general and of matric in particular be maintained?

The next section provides an overview of what standards are. The problems of communicating to the public about standards are highlighted especially when there exists no common understanding of such standards.

What are standards?

Assessment in Outcomes-Based Education focuses on the achievement of clearly defined outcomes, making it possible to credit learners' achievement at every level notwithstanding the rate at which they may have acquired the necessary competence (Government Gazette 19640, 1998). These outcomes are provided as statements of achievement in the form of standards or benchmarks.

While the curriculum defines the content based on the broad national goals of education, the educator translates these goals into classroom objectives that learners are expected to achieve by mastering a defined content by the end of a given learning programme. Sometimes, these objectives are so vague and confusing that different educators define them differently and set different standards for their learners. The problem is made worse by educators who lack appropriate teaching skills. With this confusion it is difficult to talk about common standards. The confusion is exacerbated by the administration the senior certificate examination that assumes that learners have gone through the same learning experiences and are therefore targeting common performance standards. To avoid this confusion, accepted level of performance or standard should be communicated in easy to understand manner to all

stakeholders in education so that when we debate about standards we have a common understanding of what we are talking about.

The two main forms of standards are academic content standards and performance standards. Academic content standards are drawn from the national curriculum and reflect the ideas, skills and knowledge in each learning area while performance standards are benchmarks that describe “how good is good enough” (Cizek, 2001). Performance standards define the knowledge a student must demonstrate in order to show that he/she has a certain level of understanding of the required levels of mastery of specified content. To the developers of tests and psychometricians, performance standard usually refers to the point on a test score scale that separates one level of achievement (e.g. pass) from another (e.g. fail), identified through a technically sound process. To educators involved in the development of curriculum and instruction, performance standards often mean a description of what a student knows and can do to demonstrate proficiency in relation to a content standard or cluster of content standards. The National Qualifications Framework (NQF) calls them unit standards, which are statements of the outcomes (knowledge, skills and abilities) that are to be demonstrated by an individual in order to obtain credit for the specific unit standards (HSRC, 1995). In other words, unit standard is outcome-oriented and is expressed in terms of learner capability and level of performance. In a nutshell, standards are benchmarks that specify the levels of performance expected of a learner who has gone through a learning experience in a specified content (both knowledge and skills).

Education standards can be an important tool in improving students’ achievement. Standards inform students and their parents about what society considers essential knowledge for children to learn during a given education cycle. Standards also provide taxpayers with

benchmarks to judge how well the public schools are performing. For their part, politicians and elected officials like standards because they afford them the opportunity to claim that they are doing something concrete to improve the quality of public education. Employers look at standards to get employees with skills that are relevant to the various sectors of the economy. Education standards can be used as a tool for educational change in defining a common core of learning for all students and specifying the common levels of performance and achievement that students are expected to reach.

Even where first-class standards have been developed and approved; the mechanisms developed to implement those standards vary markedly in effectiveness. Also, bad assessment devices and poor performance standards can sabotage good content standards. With such significant potential for missteps in the entire standards process, the education authorities should provide guidelines on education standards that will give lawmakers, education officials, and educators a guide to the crafting of top-notch academic content standards, assessment devices and performance standards, plus effective implementation/accountability strategies.

For standards to have any meaning, Doyle & Pimental (1997) recommend that, education experts need to attach methods of assessment, consequences of attaining or failing to attain the specific standards, and intervention strategies to help students and educators who are not able to attain them. Schools on the other hand must: Supplement their standards with clear educator guidance, be sure their standards and assessments are aligned, establish plans for phasing in incentives and consequences, and provide extra help to students who do not meet the standards. In other words standards alone are not enough.

It is very important to emphasize that good standards are not a panacea for all the ills of today's public education system. They cannot change key obstacles to real learning such as educator training programs that fail to provide the required subject-area competency. They cannot change structural problems such as labour laws that make it next to impossible to fire incompetent educators. They can raise the knowledge level of students but cannot guarantee that they will become more compassionate or moral beings. Bad standards are worse than no standards at all because they cover up shortcomings in classroom instruction and student performance and, therefore end up deceiving parents and the public. All efforts to craft content standards, assessment devices, performance standards, and implementation and accountability programs will be for naught if the details of these systems are not disseminated to parents, educators, school administrators, and the general public.

In most assessment situations, setting of cut-off scores is required before test performance is interpreted. Various levels of performance are granted only if the candidate's score equals or exceeds a specified cut-off score. The practice of setting cut-off scores is commonly called standards setting. If the literature on standard setting is conclusive on any point, it is on the difficulty of setting defensible standards in tests. There is no agreement on a best method, although some procedures are far more popular than others. Jaeger (1989) argues that all standard setting is judgmental. No amount of data collection, data analysis and model building can replace the ultimate judgmental act of deciding which performances are meritorious or acceptable and which are unacceptable or inadequate. All that varies is the proximity of the judgment of a standard to the original performance.

Comparing standards of examination bodies

South Africa has eleven bodies that administer examinations at Grade 12. These are the nine provincial Departments of Education, the Independent Examination Board and Buro Vir Volkseie Christelike Onderwys. Quality assurance across these bodies is the responsibility of Umalusi, the quality assurance body for General Education and Training (GET), Further Education and Training (FET) and ABET Level 4. Although the number of common papers set at the national level and administered by the examination bodies has been increasing steadily, there are still many papers that are set individually by each examination body. Quality of examination papers has often been put on the spotlight during the debate about standards of matriculation examination. The claim has been that some of the papers are poorly set and are often easy leading to high pass rates. Comparison of examination bodies could be an interesting areas of research in the wake of the current momentous changes in education in South Africa. Studies of examination bodies' comparability have basically relied on five different techniques (Tymms & Fitz-Gibbon, 1991) which go beyond the simplistic comparison of pass rates or the proportions awarded certain grades.

A first technique involves comparing pass-rates having adjusted for the type of school or college to which the data refer. An example of such a comparison may be found in Kingdon, Wilmut, Davidson and Atkins (1984) in which the authors refer to "the making of many assumptions in order that useful conclusions may be reached". It is also worth noting that the technique involves the use of data aggregated at the level of the institution, a procedure that has been effectively criticised by Aitkin & Longford (1986). However, the aggregation of data could be avoided by applying such techniques as multi-level analysis techniques that have been highly recommended for hierarchically structured data.

The second technique employs subject pair analysis. In this technique, the distributions of grades for candidates who take two subjects, say Physics and Chemistry, are compared for the different examination bodies, the assumption being that the relative distributions should be the same, within sampling variation, across bodies.

The third technique involves the deliberate collection of cognitive data common to all bodies. This may involve administration of a common test or inclusion of anchor questions in examination papers that are set differently by the examination bodies. The practicalities of producing fair and relevant tests remain a daunting challenge when applying this category of comparison techniques. Equating of scores using the scores on the anchor questions could be done using item response theory. Comparison on a common scale across the examination bodies can then be made.

The fourth technique is the cross-moderation, which has been used in the United Kingdom for a long time and is now commonly used by Umalusi in South Africa, which compares and adjusts for differing severities. It is a technique, which involves examiners from different bodies scrutinizing selected scripts, and making considered judgments concerning their relative merits. To quote Bardell, Forrest & Shoesmith (1978) "Cross-moderation methodology is particularly attractive, for it involves the very people who influence the most critical decisions. It has proved surprisingly difficult to design research studies which will result in conclusions of a quantitative kind capable of being translated into action at grading meetings". Although in South Africa we may not talk about equivalence in the examination bodies especially in the papers that are set separately by each body, the process of moderation and standardization done by Umalusi is an effort to achieve such equivalence.

Lastly, to assess severity of grading, the grades from different bodies are first converted to a numerical scale, for example, a A grade is given 5 points, a B is given 4 and so on. The Analysis of Variance is the most suitable statistical technique to test for difference between the means of the same subject offered by the different bodies. The average grades and the standard error (SE) for each average in each of the subjects offered by each body are computed. It is also important to compute the probability of the obtained pattern of average grades arose by chance if there were in fact no difference between grades awarded by the bodies and if the differences observed were simply due to sampling variation.

Information gathered through the various techniques provide useful evidence on whether the examination results from the different examination bodies provide equivalent measures of learner performance.

Conclusion

While pass rates at the senior certificate examination continue to improve, genuine concerns are being raised about the quality of the graduates from the school system. Disparities in performance at the school level are still evident when schools are compared to average social economic status and racial composition. It would be a waste of resources if tertiary institutions were to start administering their own tests to select students for their programmes, in which case, at least seen as a selection device for third-level education, there would be no justification for the continuation of the senior certificate examination. The research findings from both the local and international studies and those on the senior certificate examination may be symptomatic of serious flaws in the whole education system that requires urgent intervention.

It is imperative to recognize that although standard setting is an important psychometric problem, it is not solely a technical one, but involves value judgments. Involvement of key stakeholders in the process of setting standards should be given serious consideration if such interested parties are to be part of the implementation process. The consequences of appropriate or inappropriate standards for individuals, for institutions and perhaps for society as a whole must be considered.

It is important to use two or more different approaches to standard setting and multiple samples of examiners. Examination results should be used to examine empirical evidence of how a typical sample of examinees performed on the test and this information should be used in evaluating the consequences of setting a particular standard.

Senior certificate examination results alone cannot be used to assess the quality or standards of the education system. National assessment programmes such as the systemic evaluation that target the various levels of the education system offer better and more reliable indicators of quality in the system. Apart from gathering information on the cognitive domain, such a programme would collect information on other issues such as resources, facilities, school management and community involvement in the school activities.

Reference

- Aitkin, M. & Longford, N. (1986). Statistical modelling issues in school effectiveness studies, *Journal of the Royal Statistical Society, Series A*, 149(1), pp. 1-43.
- Bardell, G.S., Forrest, G.M. & Shoesmith, D.J. (1978). Comparability in GCE: a review of the boards' studies 1964-1977 (Manchester, JMB on behalf of the GCE examining boards).
- Chisholm, L. (2004). Matric pool redefined. *Financial Mail* 9th January 2004.
- Cizek, G., (Ed., 2001). *Setting Performance Standards: Concepts, Methods, and Perspectives*. Mahwah, NJ: Erlbaum.
- Doyle, D. P. & Pimental, S. (1997). "Raising the Standard," Coalition for Goals 2000. Thousand Oaks, CA: Corwin Press, Inc./Sage Publications Co.
- Fitz-Gibbon, C.T., Tymms, P.B. & Vincent, L. (May 1994) Comparing Examination Boards and Syllabuses at A-level students' grades, attitudes and perceptions of classroom processes, A-level Information System (University of Newcastle upon Tyne).
- Government of South Africa (1998). *Government gazette 19640, 1998*. Pretoria.
<http://www.ed.gov/pubs/IASA/newsletters/standards/pt2.html>.
- HSRC (1995). *Ways of Seeing the National Qualifications*. Human Sciences Research Council, Pretoria.
- Jaeger, R. (1989). Certification of student competence. In R. L. Linn (Ed.) *Educational Measurement, (3rd Ed.)*. Washington DC: American Council on Education.
- Kingdon, et al. (1984). *Report of the Inter-Board Comparability Study of Grading Standards in Advanced Level* (London, University of London Schools Examination Board).
- Naidoo, S., (2004). "Something not adding up at University". *Business Day* 20th February, 2004. Johannesburg.

Shindler, J. 2004. An analysis of the 2003 Senior Certificate Examination. Edusource No. 43/May 2004- The Education Foundation. Johannesburg.

Texas Education Agency (TEA). (2001).

<http://www.tea.state.tx.us/student.assessment/taks/standards/plan.pdf>

The Times Educational Supplement of 22nd October 2004. (<http://www.tes.co.uk/2040906>).
London.

TIMMS (1999). <http://nces.ed.gov/timms/results.asp>.

Tymms, P.B.; Fitz-Gibbon, C.T (1991). A comparison of examination boards: A levels.
Oxford Review of Education, 1991, Vol. 17 Issue 1, p17.

Umalusi (2004). An investigation into the Senior Certificate Examination. Pretoria.

Van der Berg, S. (2004). School education and transformation.

<http://www.transformationaudit.co.za/articles/Education.pdf>.