

Standardising Assessment in an Era of Curriculum Reform: The Case of High School Exit-Level Economics Examinations in South Africa

Suriamurthee M. Maistry

University of KwaZulu-Natal, South Africa

KEYWORDS Assessment. Standards. High-stakes Testing. Economics. Benchmarking

ABSTRACT Education policy has undergone radical change in post-apartheid South Africa (SA), and high school exit-level assessment policy has been subject to pressure for reformulation. This article examines recent trends in exit-level high school Economics examinations in SA and reflects on a curriculum reform context in which examining authorities struggle to benchmark standards of assessment. In SA the exercise of standardisation and benchmarking is highly politicised, especially at the crucial, high-stakes school exit level. An account is presented of the contested nature of high-stakes assessment amidst mounting public pressure to show improved pass rates and growing critique of the quality of high school graduates. Data are drawn from a rigorous comparative analysis of standardised high school exit examinations in Economics for a 3-year period (2008-2010). The assessment protocol of the state-controlled Department of Basic Education (DBE) is analysed. The DBE administers the examination of 93% of schools in SA. The findings indicate that the standard of the exit-level Economics examination has varied considerably across the years under review, revealing the fragility and uncertainty that permeates Economics assessment as the examining authority searches for a suitable assessment standard. The article begins with an outline of the existing assessment context in SA, followed by a discussion of the contested nature of high-stakes assessment. An analysis of the selected Economics papers follows, and the article concludes with a discussion of the findings and an exploration of the constructs 'inception-year dilemmas', 'veiled upward shifts in standards' and 'fragile academic vigilance'.

I. INTRODUCTION

Prior to the democratic era (pre-1994), South Africa's fragmented governance arrangements resulted in the creation of 19 education authorities - each with its own curriculum and assessment regimes. The democratic era ushered in a new Government tasked with streamlining this fragmented system into a single system. The State created the National Qualifications Framework (NQF) and the South African Qualifications Authority (SAQA), which were tasked with regulating the educational system in SA. A major part of this restructuring called for streamlining of the examination system, and the implications for high school examinations were significant. Provinces in SA were tasked with setting provincial examinations for all subjects for schools that fell under State control. By 2002 this provincial arrangement made way for all high school exit examinations falling under the direct control

of the National Department of Education, and since then all public schools in SA (93% of all schools in the country) have sat for a common examination set by the national Department of Basic Education (DBE). The remaining 7% are private schools administered by the Independent Examinations Board.

The radical restructuring described above entailed much negotiation and compromise. With regard to high school Economics in particular, there was much contestation about setting the appropriate standard for the matriculation examination. Often flawed arguments that linked standards and competence levels to race surfaced in the corridors of power (Maistry 2001). This was evident when examiners from former White and Indian education departments often made unsubstantiated claims about the superiority of the standard of their ex-department's examinations. Arguments were made for the inclusion of a higher percentage of rote-learning questions in order to cater for students who hailed from disadvantaged backgrounds (see Maistry 2001). In a South African context, these students would be poor African students.

Such assertions, judgemental positions and suspicion were not surprising given that racial

Address for correspondence:
Suriamurthee Moonsamy Maistry
Private Bag X03
Ashwood
3605
South Africa
Email: maistrys@ukzn.ac.za

constituencies that existed as separate, independent entities were now simmering in a new single education melting pot. While there was a proliferation of policies around the governance of education, the material conditions of the poor and working classes remained unchanged. The unevenness of SA's educational landscape is so stark that it is not uncommon to find a highly sophisticated school with First World resources (both physical and human) and a school without sanitation within a radius of 5 km of each other. In this context, common national high school assessments presented a serious challenge.

In 2006 the national Department of Education implemented the new National Senior Certificate (NSC), an outcomes-based curriculum at senior high school level. In SA this comprises Grades 10, 11 and 12, coined the Further Education and Training (FET) band. The first cohort of NSC graduates sat their high school exit examinations at the end of 2008.

The examining authority was under pressure to set a reliable and valid examination for matric Economics that would be fair to all high school Economics students. At the same time, the State was also under pressure to show positive returns on investment in education – which consumes the largest percentage of national budget (17% in 2010/2011). It may not be surprising that taxpayers in SA may become increasingly cynical of the State's attempts at improving student performance.

From the above one can expect that the high school exit assessment context in SA is likely to be fraught with tension and contradictions. However, the purpose of public examinations in SA remains unambiguous, namely that of promotion, selection and certification. This is a high-stakes threshold which South African learners aspire towards, and upon which dreams are fashioned. The distinct feature of high-stakes testing is that the achievement score a student acquires has overwhelming significance and consequences for the student and other stakeholders (Janesick 2001). Examination scores are a determining yardstick for all sorts of educational decision-making as they influence student retention, school ratings and monetary incentives (Jones and Egley 2004).

The burden of the current regime is often not fully comprehended, since test-based accountability systems are the function of agents

who have limited understandings of the institutional realities of accountability (Lederman and Burnstein 2006). In SA the most prominent high-stakes summative assessment is the Grade 12 exit examination (matric); the result of this standardised test has implications for students' progress into higher education and choice of career paths. While the official position from the Department of Education is that one of the purposes of this high-stakes test is to improve the standard of education in the country, it is unclear whether this intention is being realised.

From the perspective of the United States of America, historically standardised tests did not impact in any significant way on curriculum development in schools. They were simply used as a means of reporting to parents and for monitoring trends, because little attention was paid to them by the teachers (Shepard and Dougherty 1991). However, this has changed, and Herman and Golan (1991) assert that standardised testing has assumed a prominent role in efforts to improve the quality of education in American schools, where it is now used for evaluation of the performances of both teachers and schools. Literature on the relationship between high-stakes test results and the curriculum indicates that curriculum narrowing is a likely consequence, especially in schools for which these high-stakes tests are important. A feature of curriculum narrowing is heavy reliance on standardised tests as the principal source of curriculum content (Peterson 2005). When the curriculum is narrowed, content knowledge is more likely to be emphasised over higher-thinking skills, forcing students to memorise what the teacher has identified as crucial knowledge (Bishop 1995). Such tests also determine the nature of both teaching and learning, to the extent that there is an over-reliance on test preparation materials and a regression of pedagogical practice (Moon et al. 2003; Slomp 2008).

Wills and Sandholtz (2009) argue that standardised testing results in what they term 'constrained professionalism', noting that teachers find their flexibility and capacity for independent pedagogic decision-making circumscribed by the demands of such tests. Teachers feel that their ability to do their work as autonomous professionals is seriously compromised (Wills and Sandholtz 2009). Casas (2003) reminds us of the irony of standardised testing at different levels of education, including higher education in the

USA. She highlights the paradox in curriculum policy that advocates for instructional practices that support authentic teaching and learning, yet at the same time subject learners to standardised testing (Casas 2003). A similar scenario is evident in SA. A further irony is that - contrary to popular belief that if students knew that the consequences of test scores were important, this would improve student motivation to want to achieve. High-stakes standardised tests create anxiety, anger and boredom among students.

High-stakes tests must meet a set of strict criteria. They should be able to detect shifts in student achievement associated with improvements in pedagogy. Such assessment, especially when meant to test mastery of content, needs to be jargon-free and specifically related to content standards. Feedback on student performance should signal to teachers aspects of their pedagogy that are working well and those that need attention (Lederman and Burnstein 2006). However, high-stakes tests may create pressure and behavioural change by associating test results with important consequences (Herman and Golan 1991). Such tests may cause teachers to engage in the 'teaching to test' phenomenon, as revealed in a study of Economics teaching in schools in SA (Maistry 2001). The problem becomes compounded when teachers begin to teach to a 'bad' test.

Moon et al. (2003) note that regardless of poverty level, a large percentage of instruction time is consumed in practising for State-mandated tests (Moon et al. 2003). In SA the results from matric examinations and their influence on the curriculum in schools are likely to play out differently in the diverse range of socio-economic contexts that currently exist in schools. Schools range from sophisticated, resource rich, world-class institutions to poverty-stricken institutions that lack electricity, running water and sewage systems. However, one would expect that one of the goals in high schools in SA is to demonstrate improvements in matric pass rates. The extent of the curriculum compromises made to achieve this goal is a moot point.

Jones and Egley (2004) caution that standardised tests provide a powerful advantage for students whose parents are affluent and well-educated, which has significant implications for a country like SA which is plagued by high levels of chronic unemployment and poverty. Arguing from a South African perspective, Beets

and le Grange (2005) draw our attention to how assessment theory and practice are framed within a western paradigm. They signal the need to Africanise assessment and to develop approaches that draw on the principles of Ubuntu, a kind of assessment approach that values a holistic, human approach rather than one that reduces assessment to quantification and numerical indices.

In the section that follows, a description of the methodological approach to the study is presented.

II. METHODOLOGY

This article focuses exclusively on analysis of the public school Economics exit examinations from 2008 to 2010, the years in which assessment of the new FET Economics curriculum in SA became effective. The aim of the study was to establish the relative standard of the examinations by focusing on the level of difficulty and cognitive levels being tested in each question of the examination papers under review. To perform this exercise the analysis instrument employed by the Economics moderation team of Umalusi, the quality assurance body of South Africa was used. This analysis instrument draws on Anderson and Krathwohl's (2004) taxonomy of educational objectives. To distinguish the cognitive demand of each question, three categories were created: 'basic', 'comprehension, application' and 'problem-solving, synthesis'. Each of these cognitive levels was further disaggregated into 'easy', 'moderate', and 'difficult' sub-categories. Table 1 provides a more detailed account of the composition of these analysis categories.

The Economics examination papers were structured so that a candidate selected questions that equalled 300 points (marks) out of a total of 500. The examinations comprised three sections (A, B and C). Section 'A' carried a maximum of 50 points and was compulsory, comprising multiple-choice questions and short, objective-type questions. Section 'B' comprised five 50-point questions, where candidates could choose three out of the five questions. In Section 'C' candidates were expected to answer two out of four questions, each carrying 50 points, and each meant to produce an extended, discursive writing piece. This technical structure remained the same across all three years (2008 - 2010).

Table 1: Types and levels of cognitive demand for Economics (Umalusi 2009 :13)

Type of cognitive demand	Level of difficulty
<i>Basic</i>	Easy
- Recall -	Moderate
- Literal comprehension	
- Making simple evaluative judgements in terms of previously acquired facts	
- etc.	Difficult
<i>Comprehension, Application</i>	
- Understanding, application, analysis of previously acquired information in a familiar context	Moderate
- Making evaluative judgements that require use of a range of previously acquired facts/ information	Difficult
- etc.	
<i>Problem-solving, Synthesis</i>	Easy
- Analysis, interpretation and application of information in a new or unfamiliar context	Moderate
- Synthesis, creation of novel solution or product	
- Evaluation or making judgement in relation to a mixture of old and new material or information	Difficult

Each question within each of the question papers was closely scrutinised in order to apply the categorisation processes accurately. The author and three leading experts in Economics education in SA analysed all papers individually and independently at first. Each examination paper was subsequently subjected to a whole-group comparative analysis. On numerous occasions there were disagreements on categories for different questions. This disagreement was followed by rigorous debate about the cognitive expectations of questions under contention. A thorough scrutiny of the marking memorandum for questions that generated discrepant viewpoints revealed the precise nature of the expected answers for such questions. This helped the categorisation process. In many instances, questions that may have appeared to be cognitively demanding had expected answers in the memorandum that were based largely on recall of factual knowledge. In such cases the team deliberated and reached consensus on the appropriate classification. In essence the categorisation process and subsequent decision for

each question followed a carefully considered and rigorous regime.

Performing such a fine-grained analysis necessarily draws attention to qualitative as well as technical aspects of the examination question papers. The qualitative and technical issues that emerged were sufficiently significant to the overall quality and standard of these national exit-level tests, and are reported in the second part of the discussion in the results section.

III. RESULTS

Comparative analysis of DBE examination papers (2008-2010) in terms of level of cognitive demand and of difficulty.

Table 2 presents a comparison of the percentage distribution of questions across cognitive and difficulty levels for the 2008 - 2010 Economics papers.

Table 2: Percentage distribution of questions across cognitive and difficulty levels for 2008-2010 Economics papers

		Type of cognitive demand		
Paper		Basic	Comprehension/ application	Problem-solving/ synthesis
2008	Economics	47%	43%	10%
2009	Economics	31%	42%	27%
2010	Economics	27%	37%	36%

Figure 1 presents the distribution of questions in the Economics papers across the categories of cognitive demand for the years 2008 to 2010.

In the 2008 Economics examination - the first year in which the new FET curriculum was tested - the largest proportion (47%) of the question paper comprised questions categorised as basic. Of significance is that only 10% of the paper consisted of questions that tested high levels of cognitive ability. Of the 2008 examination, 43% was made up of comprehension and application questions. In 2009, 31% of the examination comprised basic questions, 42% comprehension and application type questions and 27% problem-solving and synthesis questions. In 2010 problem-solving and synthesis questions made up 36% of the paper, comprehension and application 37% and basic questions 27%.

The percentage of questions categorised as basic declined markedly from 2008 to 2010, from

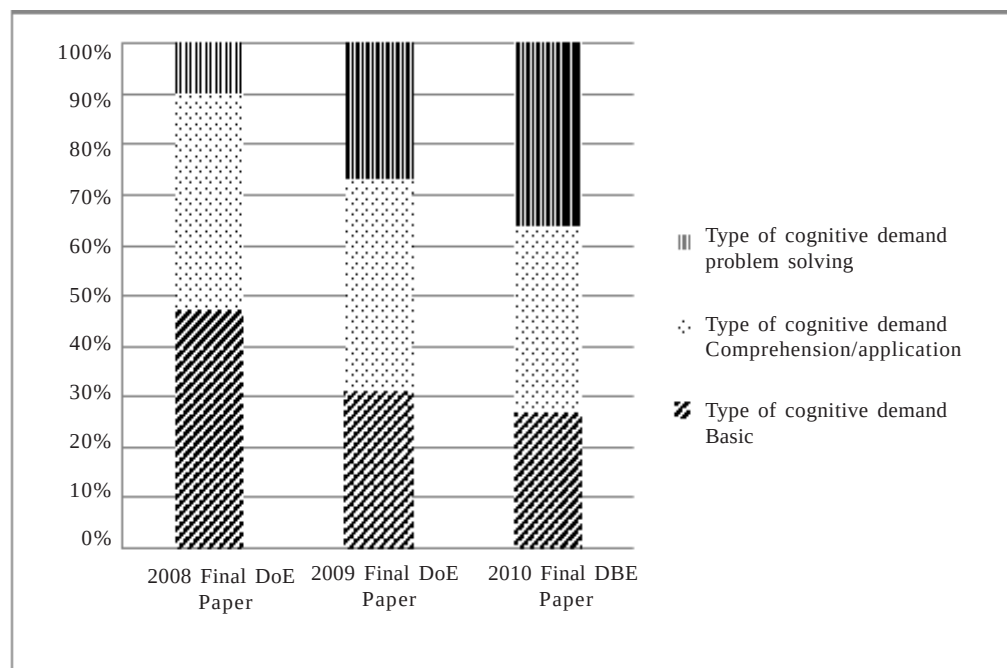


Fig. 1. Distribution of questions in Economics papers across categories of cognitive demand for the years 2008 - 2010

47% in 2008 to 27% in 2010, a decrease of 20% or a percentage reduction of 42%. Questions in the comprehension/application category also declined, from making up 43% of the paper in 2008 to 37% in 2010, a drop of 14%. With regard to questions in the problem-solving/synthesis category, there has been a significant upward trend from 2008 to 2010, as questions in this category rose from 10% of the paper in 2008 to 36% of the paper in 2010, an increase of 260%.

Table 3 presents the distribution of questions in each paper in terms of their level of difficulty:

Table 3: Distribution of questions in Economics papers from 2008 – 2010 in terms of level of difficulty

Paper	Level of difficulty		
	Level 1 (Easy)	Level 2 (Moderate)	Level 3 (Difficult)
2008 Economics	32%	40%	28%
2009 Economics	24%	35%	41%
2010 Economics	20%	52%	28%

Figure 2 gives a graphical representation of the difficulty levels for 2008, 2009 and 2010.

The discussion above on cognitive level distribution has to be understood in relation to the analysis of the distribution of questions across difficulty levels. In 2008 nearly one-third (32%) of the paper was made up questions that were classified in the easy category, 40% were deemed to be of medium difficulty, while questions classified as difficult made up 28% of the examination. In 2009 the percentage of difficult questions escalated to 41% of the total paper, while questions in the medium category made up 35% and 24% were regarded as easy. The 2010 paper reflected a significant increase in the medium category to 52%, with 20% of questions in the easy category and 28% regarded as difficult.

A distinct downward trend was observed in the questions classified as easy, moving from making up 32% of the paper in 2008 to only 20% in 2010, a 38% decrease in questions set at this level. There have been erratic movements in the easy category, from 40% to 35% and spiking at 52% in 2008, 2009 and 2010 respectively. Ques-

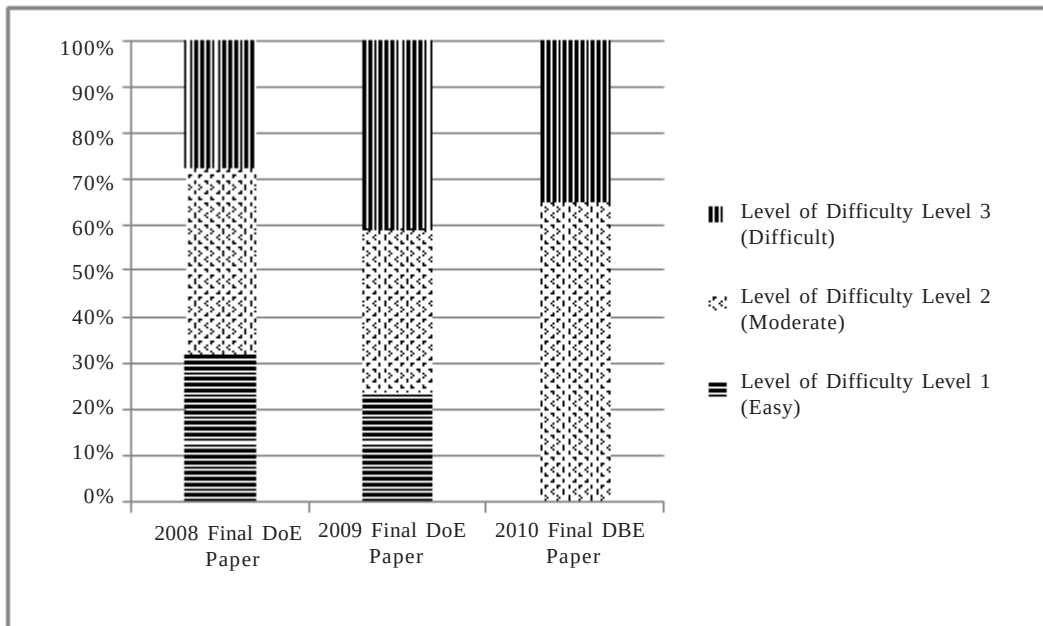


Fig. 2. Difficulty levels in Economics papers for 2008, 2009 and 2010

tions set in the difficult category started at 28% in 2008, increased to 41% in 2009 and then dropped back to 28% in 2010. There has been an increase in the percentage of high-level problem-solving/analytical questions from 2009 to 2010. A close analysis of all problem-solving questions set in 2010 revealed that only 63% could be classified as difficult, 25% as moderate and 12% as easy. In other words, difficult problem-solving questions made up 25% of the total paper; the 2010 paper was more challenging to the weak and below average students since there were more 'problem-solving/analytical questions' and a lower percentage of 'basic' and 'easy' questions.

The Subject Assessment Guidelines (Education 2008) stipulate a 3 : 4 : 3 distribution of questions across the cognitive and difficulty levels. The analysis above indicates a divergence from the prescribed norm on both accounts, namely the level of difficulty as well as cognitive level. A direct comparison of the 2008 final Department of Education examination to the 2009 final Department of Education examination revealed that the ratio of the distribution of difficulty level 1 to difficulty level 2 to difficulty level 3 was as follows: 3 : 4 : 3 for 2008, 2.4 : 3.6 : 4 in 2009 and 2 : 5 : 3 in 2010.

The above analysis indicates that there was a definite upward shift in the standard of the paper from 2008 to 2010, revealing a higher-quality examination paper in 2010 compared to 2009 and 2008.

Qualitative and Technical Challenges

Qualitative and technical phenomena that compromised the standard and quality of the examination question papers were prevalent in all three papers under study. Data from the 2010 paper will be discussed here since the examination set in that year presented peculiar flaws and inconsistencies, being littered with numerous grammatical, syntax as well as conceptual errors. The quality of the Economics 2010 paper in terms of clarity of questions, language accuracy, and explicitness of the instructions at the beginning of each section was poor constructed. More than 50% of the questions in this paper were deemed unsuitable for future use by the panel of Economics education experts.

Section 'A' comprised multiple-choice questions and those that required matching concepts with explanatory statements. More than 30% of this section comprised questions that were poor-

ly phrased (Q1.2 and Q1.3, for example) and questions where more than one answer was acceptable from the choices provided. In certain cases the authors of the examination question papers displayed distinct conceptual misunderstandings of certain economic concepts (Q2.3.2, Q3.5, Q7 and Q8). Section C (extended writing pieces) was conceptually flawed (Q7, Q10); this was evidenced in the memorandum where the expected answers were provided. The memorandum for Q10 in particular did not reflect the major requirement of the question. In some instances, statements were presented as absolute facts when in reality they are ideological statements that could be contested.

The introduction to Section C in the 2008 paper provided a brief guide to learners that contained a statement that points would be allocated for language usage. The issue of whether one assesses for meaning or accuracy of the language in which such meaning is attempted to be presented is hotly contested, especially in a country like SA where the medium of instruction is a second language to a very large percentage of learners. The literature in this area favours the approach to assessment where the examiner searches for meaning in the learner's response rather than becoming sidetracked by their technical language usage (Gibbons 2002; Halliday 1993; Jacobs 2007). Perhaps a more pertinent issue that should be foregrounded in the discursive-type responses expected in Section C is the ability of the candidate to engage with the discourse of the discipline as s/he develops a response. To its credit, the 2009 and 2010 examinations excluded allocation of marks for language usage.

Arguably the most regrettable aspect of all the Economics examinations under scrutiny is the inconsistency in the standard of questions across choices and alternatives. Because the technical structure and format makes provision for students to select questions they wish to answer in Sections B and C, there is then an expectation that each major question is comparable in terms of cognitive level and level of difficulty. This was clearly not the case for all three years, where students could choose either easy or difficult questions for the same points.

A further inherent technical weakness related to the overall structure of the Economics examination question papers. The structure was such that by default candidates were expected

to answer two 50-mark questions that were essentially extended writing pieces (essays), which comprise 33.33% of the total points (marks) achievable in the examination. By nature these extended pieces are intended to be discursive and pitched at a high cognitive and difficulty level, testing argumentation of the phenomenon being assessed: "Questions in this section are of a higher cognitive level" (Economics Examination Guidelines, Grade 12 2009: 3). Questions with a high cognitive demand and level of difficulty were also distributed across other sections in the examination. This means that if the examiners stayed true to the brief of the Examination Guidelines, one could expect a typical examination to comprise difficult and cognitively demanding questions to make up more than 45% of the total examination.

IV. DISCUSSION

The discussion that follows examines three conceptual categories that emerged from the data, namely the 'inception-year dilemmas', 'veiled upward shifts in standards' and 'fragile academic vigilance'.

A key premise from which the discussion below proceeds is that it is problematic to advocate linear, cause and effect or correlation kinds of arguments that link the standard of examinations to pass rates. Pass rates and the distribution of passes on the quality continuum are influenced by a number of variables, one of which is the standard of the examination. Other crucial factors that are likely to impact on pass rates in any particular year in SA could include improved instruction, the duration of industrial action by school teachers (which could mean up to six weeks of no instruction in some years), the hosting of international sporting events like the World Cup Soccer tournament (which shut down schools for five consecutive weeks in 2010), improved or the absence of continuing professional development of teachers, the (non)availability of textbooks, and the competence of teachers who assess examination papers. Of note is that assessment of examinations by selected teachers is an ongoing challenge in SA, where gaps in teacher content knowledge remains a serious impediment. Such an issue has profound implications for the validity and integrity of the examination assessment regime. In the Western Cape region of SA the provincial education de-

partment has begun exploring the idea of subjecting possible examination assessors to a competency test.

For the reasons signalled above, specific reference to pass rates as if they might correlate to the standard of the Economics examination has little currency as it detracts from the main argument this paper wishes to make.

'Inception-Year Dilemmas'

This concept refers to predicaments or the quandary that examining authorities have to manage in the development of novel assessment artefacts. It pertains to somewhat uncertain conditions that may include the lack of meaningful precedents, and social and political influence. Data for the inception year indicate that the standard of the examination set in 2008 was clearly the lowest of the three years under study. There are several plausible reasons as to why the standard may have been pitched in this way. Firstly, this was the first ever exit-level assessment of the new high school curriculum in SA. The outcomes-based curriculum introduced in 1997 had been the subject of much critique from various stakeholders since its implementation (Christie and Jansen 1999; Harley and Wedekind 2004). While there is no overt evidence trail to suggest that the State may have had a hand in determining the 'standard' agenda, it is not unreasonable to infer that the new fragile State, in its attempt to demonstrate the success of its brainchild (the outcomes-based curriculum), as manifested in improved pass rates, would certainly have tacitly welcomed and supported an examination that would help achieve its political imperative. This would also have provided an opportunity to laud the success of the new curriculum. It would also be naive to assume that curriculum and assessment are not politically tainted.

However, low national pass rates and persistently low numbers of high school graduates qualifying for university study continue to plague South African education. To compound the country's education debacle, a new wave of curriculum policy reform is currently under way, with the development of the latest policy statement, namely the Curriculum and Assessment Policy (CAPS) that is scheduled for implementation in 2012. Of significance is that the new policy has been 'sanitised' of all reference to 'out-

comes'. The emphasis has shifted to a much more prescriptive set of mandatory discipline-specific content. Outcomes-based education is no longer the prescribed pedagogical strategy. The implications for how this is likely to impact on the standard of the Grade 12 exit-level examinations will certainly be an area for academic scrutiny through research.

A second possible argument for 2008 having the lowest standard was that the examination had to be designed to accommodate a new, streamlined curriculum. This replaced the old NATED 550 two-tier curriculum, a curriculum streaming that distinguished high achievers from their lower-achieving students. In SA the former would have experienced what was termed a 'higher grade' curriculum with a higher grade examination, and the latter a 'standard grade' curriculum with a standard grade examination. The challenge would then have been to set an examination that factored in the entire spectrum of learners, irrespective of cognitive ability level. This would have been a novel experience for the examination panel. Whereas in the previous examination context the cognitive parameters were more explicit in terms of the cognitive competence of the potential learners in each (higher grade/standard grade) stream, such cognitive parameters were no longer meaningful. Furthermore, the relative distribution of learner numbers has historically been loaded at the standard grade end. The inception year of the assessment protocol was clearly influenced by various factors. While a short three-year period is insufficient to identify long-term trends, the distinct short-term movement from 2008 to 2010 has been in the direction of an improved standard overall. Interestingly, in 2010 the quality assurance agency took a decision not to adjust the raw scores for Economics, which suggests that the distribution of scores and pass rates was within the historically determined norms. This presents as an interesting scenario, given that the shift in the composition of the 2010 Economics examination was towards high difficulty and high cognitive levels with a substantive move away from basic and easy questions.

'Veiled Upward Shift in Standards'

As discussed above, it would be speculative to advance reasons for this phenomenon, except to signal that a number of complexly con-

nected variables are likely to have influenced the achievement levels in 2010. What does emerge is a 'veiled upward shift in standards' - a distinct movement that insidiously plays out without due warning to those who may be directly affected. The effects of such movements are likely to be far-reaching, yet not immediately in the public domain in any official way.

The movement towards higher cognitive demand in the assessment instrument certainly has implications for future classroom instruction. It requires that instructional practices in Economics be responsive to the shifting trend. Failure to raise the level of cognitive demand in daily Economics instructional and assessment practices would compromise pupils who face such expectations for the first time at a high-stakes exit assessment point. The inability to recognise and respond appropriately to such movements would create a disjuncture between Economics teaching and assessment expectations, with dire consequences for pupils who sit this examination.

While the tendency towards an improved standard is praiseworthy, this must be commensurate with practices and learning contexts that are sufficiently nuanced to better prepare learners of Economics for the raised expectations of the exit examinations. The evidence, however, indicates that education systems rarely resemble any kind of desired synchronisation, and are in fact slow to respond to assessment signals - especially where assessment protocols and artefacts are externally and centrally mandated. If anything, externally mandated assessment results in undesirable pedagogical outcomes (Moon et al. 2003; Slomp 2008; Wills and Sandholtz 2009).

The fact that school-based responses of various kinds are generally unsound makes the unevenness in the level of difficulty as it relates to alternative choices presented in the Economics examination a cause for great concern. Apart from such unevenness distorting the overall examination results, there is little doubt that it is likely to fuel ambiguity as to assessment expectations. The nature and extent of the distortion is beyond the scope of this article, but possible future research focusing on a fine-grained analysis of which question choices were most or least popular could provide convincing data as to how this unevenness in cognitive demand across questions plays itself out in candidates'

selection of questions to answer and possible impact on their scores. Of immediate concern though is the need for the examination panel to pay closer attention to devising questions of similar cognitive demand for sections of the examination paper which allow candidates to choose from alternatives.

'Fragile Academic Vigilance'

This phenomenon of paying sufficient attention is captured in the construct 'fragile academic vigilance', which refers to a kind of tenuous tentativeness associated with monitoring and control of the scholarly robustness of any academic artefact. Such fragile academic vigilance is likely to reveal fractures and inconsistencies in the artefact being created. In the current study the assessment artefact revealed, among other fractures, technical faultiness. It is untenable that with the abundance of editorial expertise available in SA that such an oversight is permitted to escape unnoticed by the country's high-stakes assessment machinery. Grammatical, syntax as well as conceptual errors in high-stakes examinations reflect symptoms that the process of screening and selection of high-calibre assessment professionals is not sufficiently robust. Such errata taint the reputation of the examining authority and present a kind of lack of rigour which is not acceptable for a high-level national examination.

Conceptual weaknesses in particular raise issues of the nature of the academic integrity of the assessment panels and the assessment teams they lead when the Economics examinations are finally scored. Persistent conceptual weakness is likely to threaten the credibility of the achievement scores of learners. This is especially crucial if such scores are closely scrutinised by discerning gatekeepers who make decisions on learners' further advancement in study or career. A high-calibre, intellectually astute examination panel is likely to be much more academically vigilant with regard to the issues raised above. Such individuals are also likely to be more perceptive of how inherent structure-related deficiencies in the assessment artefact may conspire to disadvantage or advantage different learners.

In the case under study the structure of the Economics examination question paper by default favoured the loading of over 45% of ques-

tions in the high cognitive demand category, if the examiners applied the official structure to the letter and if the unevenness described above was eliminated. This particular distribution (high-order loading) is not in keeping with the Subject Assessment Guidelines laid out for Economics in the official curriculum (Education 2008). While the assessment guidelines are exactly that - mere guidelines - it would be reasonable to accept that deviations from the declared suggestion of 30% higher cognitive demand questions are in order, and are perhaps a move towards raising the standard of the Economics examination. The official position presented by the Subject Assessment Guidelines and the contradictory *actual* examination paper are likely to send mixed signals to Economics teachers as they prepare their candidates for the examinations. It therefore has direct implications for Economics classroom instruction, since teachers shape the teaching and learning enterprise according to assessment policy directives. The above discussion reflects how the negative effects of fragile academic vigilance, as it relates to the assessment regime, are likely to influence different constituencies.

The debate about mother-tongue instruction and assessment in schools in SA continues. More than a decade after liberation it has become clear that the State has neither the political will nor the resources to back a strategic shift to mother-tongue instruction. The language of instruction and assessment in South African schools remains a vexed problem. The language of assessment in crucial high-stakes tests is certainly a body of knowledge that needs development in SA.

V. CONCLUSION

It becomes clear that in determining standards of assessment in previously divided societies, numerous challenges present themselves. Inception-year dilemmas, veiled shifts in standards and fragile academic vigilance are key issues that present themselves. Contradictory curriculum and assessment policies are likely to complicate and confuse classroom instruction. Technical and conceptual defects in assessment tools are likely to compound such challenges.

Wide swings in standards of assessment from year to year are likely to create uncertainty among both teachers and students. There is cer-

tainly a need for constant critical review of the process of selection of assessment professionals. Similarly, there has to be a critical interrogation of the qualitative processes involved in the development of the ultimate high-stakes exit-level assessment artefact.

International benchmarking exercises and cross-subject comparisons with regard to the assessment of cognitive demand in exit-level examinations could prove to be a useful endeavour as the school Economics assessment regime in SA develops further.

ACKNOWLEDGEMENTS

I extend my gratitude to Michael van Wyk and Linkie Rambuda, who formed part of the expert team that assisted in the analysis of the cognitive levels of the questions contained in the examination papers.

REFERENCES

- Beets P, LeGrange LLE 2005. 'Africanising' assessment practices: Does the notion of Ubuntu hold any promise? *South African Journal of Higher Education*, 19: 1197 -1207.
- Bishop J 1995. The impact of curriculum-based external examinations on school priorities and student learning. *International Journal of Education Research*, 23(8): 653-752.
- Casas M 2003. The use of standardised tests in assessing authentic learning: A contradiction indeed. *Teachers College Record*, 101(4): 1065-1114.
- Christie P, Jansen J 1999. *Changing Curriculum: Studies on Outcomes-based Education in South Africa*. Cape Town: Juta.
- Education 2008. *Subject Assessment Guidelines (Economics)*. Edited by Department of Education: Republic of South Africa. Pretoria: SA Government Printer.
- Gibbons P 2002. *Scaffolding Language, Scaffolding Learning: Teaching Second Language Learners in the Mainstream Classroom*. Portsmouth: Heinemann.
- Halliday MAK 1993. Towards a language-based theory of learning. *Linguistics and Education*, 5(2): 93-116.
- Harley K, Wedekind VR 2004. Political change, curriculum change and social formation. In: L Chisholm (Ed): *Changing Class: Educational and Social Change in Post Apartheid South Africa*. Cape Town: HSRC Press.
- Herman JS, Golan S 1991. *Effects of Standardised Testing on Teachers and Learning - Another Look*. Los Angeles: National Centre for Research on Evaluation, Standards and Student Testing.
- Jacobs C 2007. Towards a critical understanding of the teaching of discipline-specific academic literacies: Making the tacit explicit. *Journal of Education*, 41: 59-82.

- Janesick VJ 2001. *The Assessment Debate: A Review Handbook*. Santa Barbara: ABC Clio Inc.
- Jones BD, Egley RJ 2004. Voices from the frontlines: Teachers perceptions of high-stakes testing. *Education Policy Analysis Archives*, 12(39): 1-34.
- Lederman LM, Burnstein RA 2006. Alternative approaches to high-stakes testing. *Phi Delta Kappa*, 87(6):429-32.
- Maistry SM 2001. The challenges of teaching economics in an under-resourced Black South African school. *Perspectives in Education*, 19(1): 159-71.
- Moon TR CM Callahan, Tomlinson CA 2003. Effects of state testing programmes on elementary schools with high concentration of poverty - good news or bad news. *Current Issues in Education*, 6(8).
- Peterson MS 2005. The ethical dilemmas of high-stakes testing and issues for teacher preparation programs *Journal of College and Character*, 7: 1-11.
- Shepard LA, Dougherty KT 1991. *Effects of High-Stakes Testing on Instruction* in American Educational Research Association. Chicago.
- Slomp DH 2008. Harming not helping: The impact of a Canadian standardised writing assessment on curriculum and pedagogy. *Assessing Writing*, 13: 180-200.
- Umalusi 2009. *Maintaining Standards Report*. Pretoria: Umalusi: Council for Quality Assurance in General and Further Education and Training.
- Wills JS, Sandholtz JH 2009. Contrained professionalism: Dilemmas of teaching in the face of test-based accountability. *Teachers College Record*, 111(4): 1065-1114.