

# The Challenge of Cognitive Demand

Johan Muller

## INTRODUCTION

Most discussions to date of the rising pass rates of the Senior Certificate (SC) or matriculation (Matric) examination have tended to focus on what might be called the ‘big picture’—on patterns of student achievement, on outcomes, on the selection and predictive functions of the examination. In this presentation I propose to delve into the ‘black box’ of Matric. What I mean by that is that I wish to examine what it is that the Senior Certificate tries to measure, reading from it what we as a society, through the Senior Certificate examination, expect our school leavers to have learned by the time they leave the formal schooling system.

The question of the moment is: have our scholastic expectations for our school leavers dropped? It is a question that arises rather naturally when we see our Senior Certificate pass rates rising while authoritative educators like Jonathan Jansen tell us that the system has not done enough to reform the teaching and learning process to make a rise in passes credible. Therefore, so public opinion reasons, the standards must have dropped: somewhere along the line; the examination must have lowered its expectations, so that it has become ‘easier’ to pass or to do well.

The Umalusi research was commissioned to answer just this question. Although the validity and generalisability of the research has been questioned in some quarters, I believe its outcomes are more interesting and suggestive than some would suppose. In particular, the subject-based research reports tell a complex and instructive story, and this is reviewed below.

## CAPTURING COGNITIVE COMPLEXITY

Umalusi commissioned six reports from five subject areas: Mathematics Higher and Standard Grade; Physical Science Higher Grade; Biology Higher and Standard Grade; History Higher and Standard Grade, English First Language Higher Grade, and English Second Language Higher Grade. These were chosen to reflect the areas of most concern, but were also the subjects which had been examined by a national examination since 2003, an issue I will return to below. The exception to this was the English First Language paper, which had by 2003 not yet been nationally set.

The Umalusi research forum sharpened the question about standards to the following: have our expectations as expressed by the level of complexity of the examination questions, by their level of cognitive demand, declined? Each research group was asked to operationalise the notion of cognitive demand, which they did in terms of a scale of difficulty, except for the group looking at Biology which used the idea of a skill hierarchy. In addition, each group was asked to compare the cognitive demand level between the papers set in 1992, 1999 and 2003, when most of the papers

considered here were nationally set. Of course, this would not tell us what the pupils actually learned, only what the examination papers set out to assess. The assumption was that teachers would be teaching to what was going to be assessed. The inference was that if it was not being assessed, it was probably not being taught, or learned either. The design is summarised in Table 1.

**Table 1: Umalusi subject-based research**

	1992 or 1993,1999,2003	Difficulty rating On 3 or 4 point scale
Mathematics (SG; HG for 2003)	X	X
Physical Science (HG only)	X	X
Biology (SG & HG)	X	Skill hierarchy
History (SG & HG)	X	-
English (HG)	X	X
English (HG)	X	X

The results are quite unambiguous. In nearly every case the researchers found that the level of cognitive demand had declined, as public opinion had surmised. Here is a sampling of the judgements:

The Senior Certificate examination in Biology since 2001 can be considered to have become easier to pass...  
(Biology, Standard Grade and Higher Grade).

In terms of cognitive and subject specific challenge... a decline in the overall level of challenge of the language papers (1992-2003) and to a lesser extent, in the literature papers  
(English First Language, Higher Grade).

...it is difficult to escape the conclusion that the nationally set paper (Paper 1, 2003) is becoming easier  
(English Second Language Higher Grade).

The obvious decrease in marks allocated for level 3 (highest cognitive level) is dramatic for the national Senior Certificate (2003) with the lowest marks (at level 3) allocated in 2003  
( Maths, Standard Grade).

It is obvious ... that the NSC has a dearth of marks for level 3 items (high cognitive demand) and more than two times the marks for level one items compared to the marks awarded by the IEB.  
(Maths, Higher Grade).

In Physical Science, Physics was found to be getting slightly easier, Chemistry slightly harder. In History the researchers found that marking procedures flatten the demand intentions of the curriculum—‘the system is flawed and inadequate at discriminating between good and bad answers’.

## **IS MATRIC GETTING EASIER?**

What does it mean to say that the exams have declined in terms of cognitive demand, especially if we look at 2003? Does this mean that the nation is on a long slippery

slope to cognitive oblivion? Are our worst fears about to be realised? All those who interpret these results primarily as a progressive trend assume as much. There is a less apocalyptic possibility, however, one that makes fewer assumptions about the irreversibility of unseen forces at work, and that accounts for the results in the following more circumstantial and circumspect way:

In 2001, the exam papers began to be set nationally for the first time. This is an event whose significance we have greatly under-estimated. By 2003, six subjects (the five discussed above, English First Language excepted, plus Accounting) were either wholly nationally set or had national components. This might seem modest, but these subjects account for a very large proportion of candidates registered for the final examination. Consider, in 2003 more candidates sat the English Second Language Higher Grade paper than sat the entire Senior Certificate in 1990. At the end of 2001, the absolute numbers of passes at both Standard Grade and Higher Grade went down, although the proportion of passes went up. More worrying, there were dramatic differences between the provinces, which we could now start to compare, apples with apples, for the first time on performance across national instruments that had been compiled by national teams of examiners. This represented an enormous step forward, educationally speaking: we were for the first time getting not only equity of expectation, but we could also begin to deal with the serious unevenness between provinces, especially with respect to levels of demand, quality of layout, inaccuracies, ambiguous and unfair questions, poor sequencing and so forth—problems we had long been aware of. At the same time, though, a potential political embarrassment arose, since it raised to visibility differential provincial performance that had hitherto been invisible, and this soon became politically explosive.

There are other things to consider, too. Administering the examination nationally meant a vast increase in the complexity of the task. There were far more candidates and markers within the same framework. Partly for technical reasons (getting large numbers of markers operating within the same framework) and partly due to the new ‘skills-based’ methodology adopted, it soon became clear that it was far easier to set shorter questions that were unambiguously right or wrong than it was to set questions which required extended answers. This naturally varied from subject to subject. The Biology research team spell out how this worked out in practice for them: there was a reduction in the weighting of the essay as well as a simplification of language, which led to a considerable increase in the number of questions, each worth fewer marks. Quite inadvertently, this lowered the level of cognitive demand overall: learners had less opportunity to demonstrate knowledge in depth; to articulate in written form; and, as spelled out in the case of Biology, to synthesise across topics.

However, setting more questions means setting more text to process, hence increasing what the English First Language team called ‘language load’. Here we see how an attempt to standardise the papers to respond to the challenge of setting a national paper had paradoxical and unanticipated effects. A similar unexpected train of events occurred in History. Here, setting shorter or ‘resources-based’ questions unexpectedly made it easier to get more marks, or even full marks, than it had been with essay-type questions. As one member of the History team put it, ‘The mediocre tail of candidates is achieving higher symbols than previously and the examination is failing to differentiate between mediocre and exceptional candidates’. Subsequent attempts to

‘contextualise’ the language, to make it more ‘user-friendly’ for English Second Language candidates, added to the problem.

In 2002 and 2003 there was great pressure on the Statistical Moderation team to adjust the marks upwards. This was starkly evident in English Second Language. So paradoxically 2003, the year of six national exams, marked a low point for cognitive demand and it denoted serious political pressure for upward statistical moderation: after all, 2004 was an election year.

This interpretation suggests something like the following:

1. The provinces had been administering the final examination in a very uneven and relatively unaccountable way between 1994 and 2001, with wide disparities and very few quality checks.
2. From 2001 to 2003 national examinations began to be phased in, with the resulting unintended consequences discussed above, including the new visibility of provincial disparities.
3. Reaction from the Department of Education took the form of pressure on examiners to set examinations that would offer greater ‘accessibility’ (English Second Language), on moderators to make generous adjustments, and for soft upward statistical adjustments.
4. Thus, according to this interpretation, the 2002/2003 results were a low point produced in large measure by the requirements of nationalisation, by new forms of skills-based assessment, by the desire to make the examinations more accessible, and by the political pressures of the 2004 election. The largely invisible outcome—invisible to school educators, that is, though not invisible to employers or to university admission officers—was that the schooling system was emitting a cohort or two which had reduced opportunities to demonstrate higher-level cognitive skills, had possibly not even been taught them and, in far too many cases, therefore did not have them.
5. This interpretation also suggests, however, that this is not necessarily either a long-term or a permanent downward trend, but rather more circumspectly a temporary side effect of large-scale national reform. If that is the case, the questions to be asked are:
  - a. Do we ignore it, and run the risk of it hardening into a long-term trend (realisation of the worst-case scenario through political timidity and inaction)?
  - b. Or do we grasp the implications of rapid large-scale reform and move to make adjustments?

If these speculations are correct, then the drop in matriculation marks in 2002/2003 was a compound consequence of:

- the format adjustments of national up-scaling (including new forms of testing);
- statistical adjustments; and
- trying to make the language of papers more accessible.

The last point is, at one level, perhaps the most crucial. Language is the key to all learning. No one disagrees with this. Where they differ is why. For the Department of

Education and their examiners, language—or more properly the language of instruction for second-language speakers—was and is a block to learning principally because it was not sufficiently ‘contextualised’, or insufficiently close to the concerns of disadvantaged learners (insufficiently ‘relevant’). The solution for the Department was to contextualise the questions, to make them more relevant to the lives of learners, which many examining teams did. Thus, lack of contextualisation and relevance are considered to be semantic obstructions to comprehension. However, language is also a cognitive tool that works through the syntactic as well as the semantic structures of language. By acquiring certain forms of language, we acquire the tools for different, higher-level cognitive operations. Without those tools, those forms of cognitive operations are blocked to us; we simply cannot learn them. Language and thought are intertwined in this way, and no amount of semantic tweaking can supply the cognitive cues necessary to perform the cognitive operations when the linguistic tools are not mastered.

This is brought home in shocking fashion in the English Second Language research team’s report. The 2003 English Second Language Higher Grade examination is characterised by them as having a high level of contextual relevance, a dearth of items requiring candidates to use language as a tool for learning and thinking, a lack of genre variety, and a low level of cognitive challenge. Consider—having scored the items for the 2003 English Second Language paper, the team found it offered a lower level of cognitive challenge than its 1992 Department of Education and Training (DET) equivalent.

If the fundamental premise of this group is right, then we should not be surprised that we appear to be making little headway with increasing the attainment levels of candidates in subjects where specialised skills are required, like History, Mathematics and Science. The low level of demand and expectation in the largest language-of-instruction Matric subject, English Second Language, effectively places a cap on the learning levels possible for English Second Language learners in all other subjects. We can pump as many resources into Maths and Science as we like, but if the cognitive resources are not provided in English Second Language, we cannot expect to see learning improvements, either there or in any of the other subjects.

## **CONCLUSION**

Whether advertently or inadvertently, as argued here, low cognitive demand and challenge is a threat to the learning health of the nation. It should be addressed at every stage of the educational cycle:

- in the curricular statements;
- in manuals of exemplars;
- in textbooks and learning materials;
- in examination papers; and
- in marking standards.

Not only do we need regular and systematic checks on the cognitive demand health at the different stages of the process, but we need an overall strategy for checking the alignment of the stages. As the History team showed, a disjuncture between the marking standards and the curricular standards leads to a failure to discriminate good answers from bad, negating a central purpose of the examination.

The task here is both a training one and a monitoring one. Curriculum committee members, examiners and chief examiners, moderators and markers should all be trained to be alert to issues of cognitive demand. This task is probably best delegated to the universities. Monitoring demand and alignment is probably best done by a national agency such as Umalusi. This research is a very good start in that direction.