TOWARDS IMPROVING LEARNER PERFORMANCE IN THE NATIONAL SENIOR CERTIFICATE EXAMINATION – SCRIPT ANALYSIS RESEARCH REPORT

by

Van der Westhuizen, G.J. (Editor)

Co-authors: Maseko, L., Van Lelyveld, K., de Beer, J., Kakoma, L. & Ramnarain, U.

University of Johannesburg

Research conducted for the Examinations and Assessment Board (EAB) of the Gauteng Department of Education (GDE)

SUMMARY REPORT

NOVEMBER 2010

The Examination and Assessment Board/Gauteng Department of Education/University of Johannesburg SCRIPT ANALYSIS PROJECT

–

Research Coordinators: Prof. Gert van der Westhuizen, Dr. Luneta Kakoma, Dr. Umesh Ramnarain, University of Johannesburg

TABLE OF CONTENTS

SUMMARY REPORT

This summary report includes a) a summary of findings across subjects, b) the main findings from each subject area analysis, and c) a background section which describes the purpose, rationale and research approach (see Annexure).

INTRODUCTION

The purpose of this Script Analysis Project is to analyze samples of GDE matric/NSC examination papers of 2008 in order to make recommendations for the improvements of learner performance in various subjects.

The guiding research questions included the following, varied by subject depending on researcher background and expertise:

a) What are the levels of understanding of examination questions/assessment tasks [within and across subjects]? What examination writing approaches do learners use? What differences are there in the language of examination papers, memoranda and examination scripts?

b) What distinguishes high performance from low/failure performance? What are the error patterns/knowledge gaps displayed in examination papers? How do learners respond to cognitive demands of examination papers?

c) What are the various ways in which language co-determine examination performance? How is context and local experiences reflected in examination papers and answers?

Random samples of 1000 scripts per subject were drawn by the GDE examinations administration office.

The actual analyses were mostly qualitative in nature, depending on the analysis frameworks, to include foci on knowledge/content, interpretation and language use, error patterns, textual presentation, and the "cognitive operations" displayed in examination texts. The analyses also included a detailed study of examination papers, memoranda, and examiner reports. Variations in reports reflect the expertise and strengths of research groups involved.

Six sub-reports have been produced, coordinated and written by UJ staff members listed below, and involved post-graduate students in the Faculty of Education at UJ as well as around 36 experienced teacher-subject specialists of the GDE:

A. Computer Applications Technology – Mr. Jerry Maseko

B. English first additional language – Prof. Gert van der Westhuizen

C. History – Mrs. Kristen van Lelyveld

D. Life Science – Prof. Josef de Beer

E. Mathematics and Maths Literacy – Dr. Kakoma Luneta

F. Physical Science – Dr. Umesh Ramnarain

SUMMARY OF SCRIPT ANALYSIS RESEARCH – GENERAL TRENDS IN FINDINGS ACROSS SIX SUBJECTS NSC EXAMINATION OF 2008

1. While we have not calculated actual performance levels of the sample group, the trend is that only a minority of learners performed well. The analyses therefore attempted to identify reasons for poor performance as can be gathered from answers marked as incorrect.

2. Across the six subject areas, the primary reasons for incorrect answers seem to be a combination of reasons which include poor exam writing strategies, limited command of English, lack of content knowledge, and limited background knowledge.

3. Error patterns are subject specific, as can be seen from the subject reports in this study. These are reflected in the analysis grids used, and include criteria such as concept knowledge and understanding, interpretations of graphs/images/symbols, theory knowledge, vocabulary and language use, coherence, application/problem solving, reasoning, etc.

4. Poor exam writing involved misreading of exam questions, not answering questions at all, not answering questions fully, unforced errors, time allocation, etc.

Some sub-reports found the readability of questions problematic and others pointed to the problem of higher order questions not always being understood. Questions requiring extended writing were poorly answered, and this points to the problem of formulation and writing skills.

In quite a number of the scripts evidence was found of planful exam writing, with candidates showing rough notes to plan extended writing, or clarify steps in solving maths problems, for example.

5. Across the subjects, question types and levels distinguished between good and failure performance. Knowledge/memory questions were answered better, with insight/problem solving questions less well, across subjects. The spread of questions across levels seem to be in line with requirements, judging from official examination reports.

6. Limited content knowledge is perhaps the most significant reason for poor performance and incorrect answers across subjects. This reflects on teaching and learning/examination preparation.

In addition, the analyses found a significant number of incorrect answers reflecting limited background/general knowledge on the part of learners. In History, Literature and Life sciences, for example, questions often require of learners to be widely read and

exposed to resources and conversations about curriculum topics. Limited background knowledge and experience partly seem to be a reason for incorrect answers.

7. Limited proficiency in English seems to be a significant reason for incorrect answers which took the form of incomplete sentences, unintelligible formulations, and irrelevant statements.

8. A limited number of marking errors were found. This varied across papers, but not found to be a major problem. Some examples were found of markers not allowing alternative formulations of answers in learners' language not precisely matching that of the memoranda.

9. The recommendations made are subject specific, and include suggestions for improved teaching and learning, examination preparation and writing.

A. SUMMARY OF THE SCRIPT ANALYSIS CAT

COMPUTER APPLICATIONS TECHNOLOGIES (CAT) SCRIPT ANALYSIS PROJECT TEAM

Coordinator: Jerry Maseko	
Paper 1	**Paper 2**
Makgoga Phillicia	Mbele Petrus
Moilwa Elizabeth	Nkwane Florence
Maseko Jerry	Maseko Jerry

EXECUTIVE SUMMARY OF FINDINGS

PAPER 1: MAJOR FINDINGS IN THIS PRACTICAL EXAMINATION

Paper 1 is the practical examination section of the two part examinations. All the work is done on the computer and three software applications are examined. The spread of the applications and associated mark allocation is as File Management and manipulation was 20, Word processing and Spread sheet more than 55 marks each and database work was about 45 marks as well. This spread of marks could have been used in time management and estimates in the examination by the learners.

The learner is working from a compact disc (CD) issued to them containing all the files needed for the examination. The instructions and information section of the paper is very thorough on what each learner must make sure is taken care of by the end of the examination session. The learner applies the theory in doing and solving the presented problems in the question paper in a form of contextualised scenarios.

The grid used included a few categories of error types on both the learner and assessor parts. The categories were - Not answered, following Instructions, understanding Scenario or Analogy used, Language use, expressing themselves through ‘Explain why’, describe, discuss, etc questions, checking if the learner understood a particular Conceptual, actually demonstrating a Practical skill, logic in following an Algorithm / Mis-ordering, and the marker influence in the assessment of the learner’s work under No marking or Incorrect marking. A few key types of errors across the questions are described and discussed.

The first major problem was no work to assess. This issue of no work to assess came in two forms – empty spaces and absolutely no work backup. Referring to Table 3 on the

main document, The most common error for question 1 is Error Code 1 (EC1), with **2 345** errors, and the least common error being marker error, with 8 errors. The last two columns of the table indicate the highest frequencies of errors down the questions and across the codes. It is interesting to note that they all come from EC1. The percentage contribution of EC1 to the tally for each question is as high as from 46% in question 1 to 93% in question 7. This basically means that there were **31 394** counts of no responses to sections of questions, where question 7 was attempted the least of all the questions. An observation can be made that the frequency of questions not done increases from question 1 to the last question of the paper.

About 40 learners had no backup in the form of a CD or printouts of their work presented for the assessment. The whole three hours spent sitting for an examination turned to a zero mark. Some of the presented CDs were empty or had viruses on them, resulting in them not being usable for assessment.

Another reason for EC1 being the most prevalent error is due to the fact that the database sections were the least attempted of the sections of work done by the learners. We noted that questions 6 and 7 accounted for over seventeen thousand EC1 errors. The next application that was poorly done is the spread sheet in the questions that involved it. Over seven thousand counts of no attempts were found in this section.

Our conclusion on the Paper 1 performance is that many learners do not have enough access to a computer for practice. Secondly, it appears that database and spread sheet applications are not taught well enough to the learners to have confidence in using them and attempt them in the examination. The more time the learners have access to computers both at school and away from school, the more confident they will be in exploring and discovering different approaches to solving problems.

It does appear however that word processing is not a major problem for the learners, except when it is integrated with other applications such as the spread sheet and database.

PAPER 2: SOME MAJOR FINDINGS IN THIS THEORETICAL EXAMINATION

Paper 2 expects the learners to have a certain amount of language capacity in both ordinary and common English, and the language or terminology used in computer communication. They need the two sets of languages to understand the question and then respond appropriately by using both sets of languages.

This paper expected learners to express themselves in responding to the questions. Every one of the six questions of Section B have the **'explain why', describe, discuss, etc.** questions where the use of words to describe, discuss or explain a procedure, etc. is expected, with highest count of errors. The use of language to access the meaning

embedded in the questions; and the process of understanding the question to prepare a response plays a major role in what learners submit as their answer to be considered for assessment. Many learners lost points because either they could not express themselves in an acceptable and understandable terminology, or common English, to be awarded points. Alternatively they could not understand the questions clearly enough to make up reasonable responses. Another problem could be that they could not transfer their practical knowledge from working with the actual applications to the theoretical section and translate their knowledge to words.

Referring to Table 4 In the main document, of the **32,166** errors, the most common error was when the learner was expected to use language to answer the question, with **18,931** (58.85%) errors. In the category **'explain why', to describe, to discuss, etc.**, learners need to use words and vocabulary specific to the Computer Applications Technologies and common English. This type of error runs through all the questions and has the highest count in each of them. It was noticeable that the main contributor to this error type count are both questions 3 and 8 dealing with file management and internet banking; and spread sheet use, respectively. From Paper 1 one could predict that learners were not going to do well in a question involving a spread sheet. The questions in the file management included an assumption that learners would be exposed to a variety of versions of the same software. Not all learners have access to internet and the experience of internet banking as an example becomes out of reach and experience.

GENERIC FINDINGS (BOTH PAPERS 1 AND 2) – LANGUAGE SKILLS

It does appear that language skills are critically important for the learners to understand the subject. Since the subject has its own set of vocabulary and terminology, common English can be a bridge to that particular language. Again, the computer applications used in the subject have many affordances a user can benefit from. An example is the explanatory text attached to an icon as one hovers over it. The language used in this example assumes the learner understands Basic English and will be able to explore what the icons claim to be ready to do when used. When this is in place a lot of self-learning becomes possible as opposed to waiting for the teacher to lead every learning opportunity.

Connected to this language efficiency drive are the following recommendations:

- Helping learners improve language use and efficient learning of the subject might result in better performance in the theory paper.
- Expose the learners to the bank of previous question papers and make them write down the steps and procedures of solving a problem from both papers (Papers 1 and 2) to make them practice the use of the appropriate language. In

the process they will be developing and accumulating the correct terminology of the subject and training to follow instructions and procedural steps necessary to accomplish a solution.

B. SUMMARY OF SCRIPT ANALYSIS: ENGLISH FIRST ADDITIONAL LANGUAGE

For the purpose of the UJ/EAB/GDE script analysis project, a sample of 1000 scripts (5%) of candidates in Gauteng Province who wrote Grade 12NSC exams in 2008 were analysed.

FINDINGS AND RECOMMENDATIONS PAPER 1: COMPREHENSION AND LANGUAGE

The analysis grids focused on inferential errors, interpretation of questions, clarity of answers, examination writing errors, and marking errors (see the grid in the report).

The main findings from the analysis of Section A, Reading Comprehension:

a) The error counts show that performance was relatively high, especially in the case of questions which required identification of answers from the readings.

b) Candidates fared poorly with questions requiring interpretation and some creative response, which shows an inability to go beyond surface meaning into deeper meaning and understanding.

c) A limited number of marking errors were identified, mainly due to literal interpretations of memos by markers.

d) Exam writing mistakes are evident, i.e. unforced errors, which are the result of learners not reading questions well; and/or misinterpreting questions.

Main findings: Questions on Visual Literacy:

a) Performance was not good in comparison with Section A – the range in number of incorrect answers across 4 questions is from 235 to 525, with an average of 400.

b) The prominent reasons for incorrect answers include misinterpretations of the visual images and lack of background knowledge.

c) The language used in the presentation of answers contributed to error counts. In addition, candidates lost marks by not fully answering questions.

The **main finding from the language section** is that candidates performed poorly with up to 756 out of 1000 incorrect answers. These related to poor sentence construction, vocabulary use and editing skills.

Recommendations - Paper 1:

a) To enhance reading comprehension performance, emphasis should be placed on extended reading to increase the experience base of learners. This should help widen learners' perspectives; as well as helping them understand different genres. At the same time abilities to make inferences will be enhanced.

b) Learners need to be exposed to examination techniques so as not to commit errors such as rubric infringement errors. It is important for candidates to learn how to follow instructions of the examination paper or question.

c) Marking errors in candidates' scripts call for further clarification of memoranda, and some leeway for markers to interpret answers and give credit where due.

d) To improve visual literacy performance, we suggest more exposure of learners to visual media, extended reading programmes, and some emphasis on the distinction between verbal and non-verbal communication. Emphasis needs also to be placed on the distinction between surface- and deep meanings in order to encourage candidates to go beyond literal interpretations of text.

e) Teaching of language and grammar is an ongoing challenge – poor performance may be improved by studying language more explicitly in the context of readings.

FINDINGS AND RECOMMENDATIONS PAPER 2: LITERATURE

The analysis focused on the poetry and literature sections (contextual and essay), looking at trends in incorrect answers by question type and level; questions drawing strongly on general knowledge, poetry theory knowledge, deeper levels of understanding and insight; and those requiring some kind of problem solving challenge.

The analysis grid worked with the error types of language errors (such as grammatical errors) in the presentation of answers, lack of insight into text, lack of background and content knowledge relevant to the prescribed text, and 'rubric infringement' (not answering the question in full).

Findings:

a) The Poetry section was chosen by 228 candidates in the sample, and performance was generally poor given the high percentage of errors across questions.

b) The most prominent reasons for incorrect answers are the lack of background- and poetry knowledge.

c) Misinterpretations of the poem, or parts of the poem, seem also to be an important reason why answers were marked incorrect.

d) Marking errors seem relatively high, indicating that memoranda were applied rigidly and variations of formulations not credited.

e) In contextual questions the reasons for poor performance revolved around a lack of literary theory knowledge, lack of insight into the text, and a lack of content knowledge of the prescribed books.

f) Rubric infringement errors were found across all contextual question levels. Although few, these errors are unnecessary as they indicate candidates not being well prepared for reading exam paper instructions, as well as question instructions.

g) Essay questions on prescribed texts - very few essays were rated good/excellent in all five criteria: knowledge of the book, knowledge of literary theory, question interpretation, presentation of answer content, and presentation of answer form.

h) The reasons for poor performance in the section essays on the prescribed texts include poor content knowledge of the prescribed work and of required literary theory.

i) Inaccurate interpretations of essay questions added notably to poor performance.

j) The way candidates presented their essays in terms of content and form seem to also add to poor performance.

Recommendations:

a) Teaching of literary theory and devices and functions needs to be emphasised.

b) Learners seem not well read – answers display poor background and general knowledge. Extension of reading programmes cannot be overemphasised.

c) Historical background of texts needs to be taught explicitly.

d) Practice of reading poems and writing about them needs to be prioritised.

e) The development of exam writing skills is an ongoing need, and should include practice in reading, analysing, and planning for different types of exam questions. Self-assessment formats need to also be extended.

f) Memoranda for marking seem limited by not allowing sufficiently for variation in answers. The consequence is that answers do not always get the credit they

deserve. The recommendation is that alternative formulations need to be included/extended.

FINDINGS AND RECOMMENDATIONS PAPER 3: ESSAY WRITING

In the case of Paper 3, the different genres of written work were analysed in terms of a set of criteria which focussed on background knowledge, topic understanding, genre match, understanding of exam questions, language proficiency, subject engagement, and writing style and coherence.

a) By far the majority of essays, 71.2% have been judged to display limited to very limited background to the topic of writing. The topic for letters, about a class trip, excludes learners from schools where such trips can not always be afforded, who clearly do not know the geography of South Africa, e.g. they don't distinguish between Durban and Mpumalanga. This should be addressed by the GDE examiners.

b) The percentage of written pieces where there was a complete mismatch between topic and genre was 26.8%. This means that candidates were not always able to comprehend the topic and genre and write accordingly.

c) Most learners scored below average symbols for their synthesis of personal writing style and topics chosen to write on, which means that they did not understand what type of writing they were capable of.

d) Almost three quarters 72.3% of essays were judged to display poor analysis and thinking about the topic of essay. This figure reflects poorly on the candidates' preparation for the exam and their ability to voice their thinking.

e) When added up, the essays written by 25.3% of candidates seem to display a fair to good attitude towards English as a subject.

f) A quarter of essays were judged to have shown average to good/excellent coherence while the majority not.

g) The spread of actual scores show that 35% of candidates got less than 50% for their essays, 39% between 51–60% group, and 26% higher than 60%. Most of the marks have been judged as fairly allocated - 76.7% were seen as being marked fairly, with 15.5% being given marks that were viewed as too high, and 7.8% as too low.

h) As far as exam writing conventions are concerned (focus 10) the finding is that around 45.3% of essays display some average to good effort in planning/rough work appropriate to essay writing exams.

RECOMMENDATIONS:

Performance in Paper 3 seems to be strongly determined by a combination of factors – these include background knowledge, language used, essay writing skills, exam question interpretation, genre knowledge, etc.

Recommendations for the improvement of learner performance include:

a) Background knowledge - helping candidates expand their background knowledge, e.g. by means of educational tours/visits, extensive reading, reading newspapers in class, etc.

b) Exam writing - help learners to understand what is required, how to read questions properly, analyse topics, choose topics that they are familiar with, and plan their writing.

c) Essay writing capabilities – develop genre knowledge, skills to write coherently, and how to integrate and use what they know from other subjects.

d) Love of the subject – encourage learners to be motivated, caring for the subject, and their own literate capacity in English.

C. SUMMARY OF THE SCRIPT ANALYSIS OF THE SUBJECT HISTORY

For the purpose of the UJ/EAB/GDE script analysis project, a sample of 2000 history scripts (the history component consists of two papers, so 1000 of each history paper), amounting to 5% of candidates in Gauteng Province who wrote Grade 12 National Senior Certificate (NSC) exams in 2008, were analysed.

FINDINGS AND RECOMMENDATIONS OF PAPER 1

Question 1 was most answered question of the four possible questions in Paper 1, signalling learners were most confident with this course work on how the Berlin Wall intensified Cold War tensions in Europe. The learner results in each question varied, however it emerged that the higher skill questions according to Bloom's Taxonomy of Learning were not well answered on a consistent basis, indicative of a lack of practice in those skills. The question raising the most concern is that of extended writing. There are very few overwhelmingly well answered questions; with less distinction shown between the results of the correct, partially correct and incorrect answers.

As question 2 was the least answered question of Paper 1, it appears the learners have less confidence with this learning matter on how Patrice Lumumba's rule undermined the struggle for freedom in the Congo. The results here point to the skill, in this case application, being present, however in question 2.1.4 not achieving at all and in 2.6.2 achieving well. This inconsistency can only be remedied through practice of the skill throughout the years of preparation to the final examination. As in question 1, the extended writing question is of great concern with exceptionally few learners with sufficient skill to achieve the requirements of the question. Essay writing is a skill that needs to be taught and practiced, with the learners receiving feedback and recommendations from their teachers. The time constraints placed on the learners in the exams needs to be recognised, and teachers should facilitate sessions where the learners are placed in circumstances similar to that of the examinations so that the learners are in the best possible position in the final examination.

FINDINGS AND RECOMMENDATIONS OF PAPER TWO

Of the four possible questions in Paper 2, question 1 was the best answered question. The subject knowledge on how the Soviet Union's collapse influenced ending Apartheid in South Africa is clearly more understandable to the learners. The most well answered question in this paper was a knowledge based question according to Bloom's Taxonomy of Learning, requiring less skill. The questions answered least well are to be found higher on Blooms pyramid thus requiring more skill, and reflecting the lack of skill which needs to be rectified through practice prior to the examination.

Question 2 was the least answered question in Paper 2. The topic of this question was based on how the USSR collapse caused Benin to re-imagine itself. It appears that the learners were not comfortable with this topic and for this reason so few chose to answer it. The best answered question was based on the most basic skill of knowledge on Bloom's pyramid of Learning and the worst answered question was based on the second highest step on the pyramid, demonstrating the skill of analysis. The extended questions were particularly badly answered in question 2, denoting an extreme lack of ability to synthesise and evaluate. This skill is only achieved through practice and guidance.

D. SUMMARY OF THE SCRIPT ANALYSIS LIFE SCIENCES

FET Life Sciences is composed of four knowledge areas: (a) tissues, cells and molecular studies; (b) structure, control and processes in basic life systems; (c) environmental studies; and (d) diversity, change and continuity. Examination papers are structured around these knowledge areas, with Paper 1 assessing the first two knowledge areas, and Paper 2 assessing the second two knowledge areas. Four researchers were assigned to analyse Paper 1 (two Life Sciences teachers, and two GDE officials); and four researchers analysed Paper 2 (three Life Sciences teachers, and one GDE official).

This script analysis research focused on the following aspects:

1. Learners' knowledge (including Life Sciences terminology) and comprehension
2. Learners' application skills
3. Misconceptions that learners hold
4. Problem solving skills
5. The Nature of Science, and scientific investigations (e.g. hypothesizing skills)
6. Language (the language of the papers, as well as learners' language skills)
7. The incorporation of indigenous knowledge in the curriculum, as stated by the NCS.

The Rubric used for the analysis looked at aspects such as whether incorrect answers stemmed from a lack of knowledge, a lack of application- and problem solving skills, graphing or calculation errors, whether learners struggle to hypothesize, misconceptions (that were classified), and also more generic problems that could result in underperformance, such as language proficiency, poor examination technique, and poor time management (Rubric provided on page 6). However, we also triangulate, by using data obtained from a large GDE Research Project, where the responses of 250 Life Sciences teachers were obtained on the new curriculum.

[Previous number is '1. Summary of findings' – check consistency with below numbering]

OUR MAJOR FINDINGS REGARDING PAPER 1

Many learners do not know the **terminology,** especially in the genetics section. A common problem that was identified was that many learners do not know, for example, the difference between chromosomes, daughter chromosomes, chromatids, homologous chromosomes, etc. Another problematic aspect is the **application of knowledge**, especially where diagrams are provided and the learners have to apply their knowledge in a specific context. The same applies to questions like question 1.4, where learners have to apply knowledge based on a graph that is provided. Many

learners find it difficult to **interpret graphs**. Another major problem is learners' **lack of scientific reasoning** (for example, in question 3.2.2). Aspects such as genetic modification were answered poorly, and it seemed as if teachers did not spend adequate time on this topic.

Our conclusion is that teachers focus too much on **transmission of knowledge**, and do not adequately address higher-order thinking skills, or the syntactical nature of the subject. Many students find it difficult "to think like scientists" – they have a problem in formulating hypotheses, and how to apply knowledge in a Science-Technology-Society context.

OUR MAJOR FINDINGS REGARDING PAPER 2

The evolution section in Paper 2 is a big concern. **Students do extremely poorly in many of the questions**. Question 2.1.1 and 2.1.2, dealing with natural selection, had average scores of respectively 27% and 20.2% (2.16 out of a possible 8, and 0.404 out of a possible 2, respectively). What is worrying is the low score of learners in questions requiring definitions (for example, "species" in 2.4.2); the average score was 0.542 out of a possible 3 marks - or 18.06%. Questions dealing with **human evolution** (notably questions 3.1.1 and 3.1.2) were answered very poorly – the average mark for question 3.1.1 was 14.1%, and for question 3.1.2 was an average mark of 21.2%. It seems as if many teachers did not teach this section at all. In our analysis of our GDE Research Project data, it was found that many teachers have religious objections to teaching evolution, and many teachers do not spend adequate time on this theme. This will be discussed later in this report. We realized, from the learners' answers, that often a religious register is used instead of a scientific register. This was confirmed through the interviews we had with teachers. Here follows but one example:

"*I teach evolution exactly the way that it is in the syllabus. Obviously I bring religious viewpoints into it when I teach. I was very apprehensive with evolution because I was concerned by the influence it would have on the children. But we were given assurance that they would not put any questions in the paper that would challenge us, religiously and I think that alleviate some of my fears. So I approach it as, it`s something that has to be taught and obviously we try to* ***ensure that we are not going to let our children buy into the theory of evolution***." (Interview with teacher; our emphasis).

"*Learners are very negative, very bored, very disinterested. I would say definite, over three years that I have taught evolution. They believe it is absolute nonsense, it doesn`t make any sense, they really just laugh most of the time*." (Our comment: not surprising, if one keeps in mind that teachers 'ensure that children do not buy into the theory of evolution').

GENERIC FINDINGS (BOTH PAPERS 1 & 2)

Learners' language skills

An examination paper's sentence complexity plays a big role in learner performance (Dempster, et al., 2008). The sentence complexity in the November 2008 papers was determined by Dempster, et al. (2008). Firstly, the number of words per sentence was counted. Secondly, the number of long words (words with more than seven letters) were counted. The data is provided in Table 1.

Table 1: Readability of the Nov 2008 Life Sciences examination scripts

(Dempster, Celliers, Wiese, and Preethall, 2008)

	Paper 1	**Paper 2**
Sentence complexity	11.8 ± 4.2	12.5 ± 4.6
Nr of long words	3.9 ± 4.6	4.0 ± 4.2

There seems to be an average about four long words per question. We concur with Dempster, et al., that the complexity of language in the Life Sciences papers is a cause for concern. Dempster, et al. (2008) indicated that 38% of questions had a sentence complexity above the average of 12 words per sentence.

In general, many learners lack the skill to **write a scientific essay**, and answers show a lack of planning and little alignment between paragraphs. Our recommendation is that more reading and writing activities are needed in the Life Sciences classroom.

A lack of scientific reasoning: Finding a balance between the substantive and syntactical nature of Life Sciences

Many learners find it difficult to follow the reasoning of the scientific method, where they have to formulate hypotheses, design experiments to test the hypotheses, make critical observations, interpret the data, and make valid conclusions. This should be emphasized in the classroom. From the script analysis, it is clear that not sufficient attention is given to the syntactical nature of the subject in the classroom. The so-called "scientific method" is not understood by many learners. They find it difficult to make inferences, to critically apply knowledge to data sets such as graphs, to hypothesize, or to make conclusions. The focus seems to be mainly on "preparation for the examination". Themes in the curriculum that seems to be particularly challenging to learners are genetics and natural selection (evolution).

RECOMMENDATION

Teachers should emphasize the scientific method and critical thinking in the classroom. Terminology should receive more attention, as well as higher-order thinking skills. Teachers' Pedagogical Content Knowledge (PCK) needs to be a focus in professional development programmes.

E. SUMMARY OF THE SCRIPT ANALYSIS MATHEMATICS AND MATHS LITERACY

INTRODUCTION

Although sharing a similar characteristic of deviation from accepted mathematical thinking; learner misconceptions, errors and mistakes in mathematics have different meanings (Smith, DiSessa, and Roschelle, 1993) that are now pointed out.

Mistakes or slips result from misreads or incorrect working out, for example. Learners make mistakes unintentionally and can readily correct them by themselves once they become aware of them. Mistakes can be easily corrected because there is no deep conceptual understanding linked with them. So the resolution of a mistake is much easier and uneventful. But unseen mistakes interfere with performance in mathematics because if learners cannot correct them in time, they can lose the leads when working out long mathematics problems that require earlier results to be inputted in the later stages of the question. Mistakes therefore are random, non-systematic and non-recurring and are due to performance lapses rather than planning.

It is a different story with errors as learners who commit errors may continue to commit them, even when they are pointed out to them. This is because the errors are sensible to their constructors. These systematic errors are usually exhibited in learners' artifacts such as their writings or speech (Riccomini, 2005). They are due to learners' alternative understandings of focal mathematical concepts. These alternative understandings that learners have conceived in their minds make sense to them but are at variance with accepted canonical knowledge held by the mathematical community or in mathematics textbooks. The alternative conceptions, or misconceptions, are usually generalizations of specific mathematical knowledge to unfamiliar contexts in which learners fail to negotiate the limits of specific knowledge. For example, having noted that $3x + 5x$ gives the same answer $8x$, as $5x + 3x$ (the addition operation), learners would go on to think that the subtraction operation is also commutative. They would incorrectly deduce that $5x - 8x$ give the same answer $3x$, just as $8x – 5x$. (But in another context the *sum* or *difference* of $5x$ and $8x$*;* or $8x$ and $5x$ give the same respective answers!). The limited success of mathematical results is puzzling to learners. The mis-generalisation of valid mathematical knowledge to novel situations represents a misconception. The misconceptions can be due to faulty learner inferences or the nature of mathematics itself as the above discussed example on addition and subtraction shows.

The main point as far as this research is concerned is that learners often have defensible though incorrect reasons for their errors, hence it is important to note that learners have integrity in the answers they provide to mathematics questions (Green, Piel, and Flowers, 2008). It is up to teachers and researchers to interpret learners' lines of thinking drawn upon when they answer mathematics questions. Even in learners'

apparently correct responses may recline seeds of a misconception. This is because some learners may have correct mathematics answers for the wrong reasons. It is important then to firstly identify and classify the errors the learners are making in answering particular questions and then secondly, examine *why* learners make the errors they make. This helps to find out and reveal their reasons for the answers they give. This grounded knowledge helps teachers to formulate appropriate interventions and interceptions that engage the inadequate or deficient learners' reasoning to assist learners to see for themselves why their understanding falls short in answering questions.

Sometimes errors are due to learners' intuition rather than misapplication of prior acquired knowledge. Errors then signify an underlying misconception responsible for their occurrence. It is important then, that for teaching to be effective, error analysis must always accompany it, so that teaching strategies engage the thinking patterns that make learners to commit the mathematical errors they commit. When the underlying notions that learners resort to are not addressed in pedagogy, teaching becomes a hit and miss affair hinging on mathematics logical structure. This failure to expend learners' psychological standpoint squanders the potential to improve learners' learning and performance.

It is for this reason that Script Error Analysis (SEA) is so critical, as researchers diagnose what learning difficulties examinees face for the purpose of re-directing that expertise to schools for later generations of learners to profit from the findings.

RESEARCH METHODOLOGY

At first the mathematics paper examination questions/items were analysed for their mathematical demand using Stein, Smith, Henningson, and Silver's (1993) mathematics task demand framework. According to Stein et. al., mathematics tasks lie in a spectrum from low levels of mathematical demand to high levels mathematical demand. The four levels in order of increasing demand are memorization, procedures without connection to mathematical contexts or understanding tasks, procedures with connection to mathematical contexts or understanding tasks and doing mathematics, or problem solving tasks. This question analysis was important for the purposes of eventually mapping the students' errors and misconceptions to the items' levels of mathematical demand. The examination questions were analysed using a revised Bloom's taxonomy as well as Hodes and Nolting (1998) error descriptions according to error types. Four errors types were considered:

- **Careless or Unintended errors:** These are errors which occur even when learners have the required knowledge to perform a task correctly and they usually occur when a learner is distracted when doing mathematics.

- **Procedural errors/Execution errors:** These are errors that are due to the misuse of a rule or formula or carrying out a wrong calculation.

- **Application errors:** This occurs when a learner knows the concept but cannot apply it to a specific situation or question.
- **Conceptual errors or Structural errors:** These are made when a learner does not understand the properties or principles required to successfully answer the mathematics item.

It is important to note that these errors are not mutually exclusive and it is possible that an error could fall into several of these categories.

The sample consisted of 4000 scripts randomly selected for Script Analysis Project (SAP) research. The unit of analysis was the students' errors and misconceptions in answering examination items. A protocol was developed for each paper to assist with the analysis. The frequency of the errors was quantified according to the error type. The episodic errors were selected and described in detail with exemplars from the learners work. This exercise was followed by further qualitative and a quantitative analysis of errors.

The 4000 scripts in this sample were divided into three learner ability groups prior to analysis. These groups were based on the learners' overall performance in each paper.

- Group 1 (learners who achieved a minimum of 60%) - 556 learners
- Group 2 (learners achieving a minimum of 30% and less than 60%) -1429
- Group 3 (learners achieving less than 30 %) -1898

The analysis of the group 1 learner group could be of interest to tertiary institutions since it is probably these learners that would pursue a mathematically-oriented course. The analysis of the group 1 and 2 learners would provide valuable feedback to educators regarding common conceptual errors as these learners have exhibited some mathematical ability. The group 3 learners represent a significant proportion of the sample (63%). These learners show a limited understanding of the National Curriculum Statement outcomes. There are several reasons why these learners are in this group and among them are a combination of poor work ethic, ineffective learning and inadequate/ineffective teaching during their school careers. It was difficult to analyse conceptual errors for this group in all the three papers. In many cases, the questions were not attempted or the answer was totally inappropriate and unrelated to the question. It was evident that many of these group 3 learners were not prepared for the examination and lacked fundamental and basic mathematical skills that they should have learnt in earlier grades.

The analysis was presented in tables which displayed a breakdown of the overall learners' performance from each group for each question.

The taxonomy levels reflected in these tables for each question were obtained from the "Internal Moderator's Report". A brief overview of these is given below for quick reference (obtained from the NSC Assessment Guidelines):

- **Knowledge** (algorithms, recall, knowledge and use of formulae, identifying from data sheet, theorems)
- **Routine procedures** (familiar problems which could integrate Learning Outcomes, perform well-known procedures, identifying and manipulating formulae, simple applications and calculations that may require interpretation)
- **Complex procedures** (mainly unfamiliar and involves integration of LOs, higher level calculation and reasoning, can be abstract without real world context)
- **Problem Solving** (non-routine unseen problems, extrapolation from solutions, high level reasoning, identification of course to follow when solving a problem)

These levels indicate the cognitive demand for each question. It was interesting to compare the results for each group in the various questions and reflect on their performance with respect to the taxonomy levels.

SUMMARY OF FINDINGS IN MATHEMATICS SCRIPTS

Every question in each paper was analysed for the above stated error and using the protocol that was developed. The 4000 scripts were divided into three categories of learners.

High achievers 65 – 100%

Mid achievers 50 – 64%

Low achievers 0 – 49%

- In general the performance was low, 26% pass (above 50%) out of the 4000 scripts.
- The most prominent errors were associated with lack of conceptual knowledge; these were conceptual errors that averaged 59% across the four papers (4000 scripts). These errors were predominantly common among the low achievers.
- There were very few unintended errors, 10.75%, and these were predominantly common among the high achievers

- The application errors were 11.5% of the 4000 scripts. These errors were more common among the high achievers and the mid achievers.
- Procedural errors were 18.75% of the 4000 scripts and these were predominantly among the mid achievers and the high achievers.

	Careless or Unintended errors **% of 1000 scripts of each paper**	**Procedural errors/Execution errors** **% of the 1000 scripts of each paper**	**Application errors-** **% of the 1000 scripts of each paper**	**Conceptual errors or Structural errors** **% of the 1000 scripts of each paper**
Math Paper 1	**7**	**11**	**18**	**64**
Math Paper 2	**14**	**24**	**9**	**53**
Math Lit Paper 1	**13**	**18**	**11**	**58**
Math Lit Paper 2	**9**	**22**	**8**	**61**

Figure 1 Percentage of errors per Mathematics paper

From the table above it can be noted that most of the errors across all the papers were conceptual. They indicated that learners' lack of conceptual understanding and therefore conceptual knowledge. It also shows that because the learners lacked conceptual knowledge they made less careless or unintended errors, because such errors are associated with one who knows the concept or the procedure to solve the problem but in the process makes a careless error.

DISCUSSIONS AND RECOMMENDATIONS

In analysing these scripts it was disturbing to note that in certain questions the standard of marking was questionable. Many learners received marks for incorrect work, whilst others were not awarded marks for partially correct answers. Overall students' performance is bimodal. There are distinctly two groups of performances. The large majority of students performed poorly while a small minority performed extremely well. It

is clear that poor performers are debilitated by many errors and misconceptions they hold of mathematical ideas. Among them are:-

- Students seem not to have the necessary conceptual grasp of mathematical concepts, as such they depend on memorised procedures that are mis-applied to situations they do not apply to.

- Students' minimal grasp of algebra; the research revealed that students' algebraic knowledge was weak and this led them to create their own rules of working out algebraic expressions. These rules were inevitably wrong, even though they were sensible to the students.

- The scripts showed that the low achievers needed a lot of time to engage with mathematical ideas at school in order for them to develop conceptual knowledge and procedural knowledge of mathematics necessary to perform well.

- The students' lack of conceptual knowledge can be linked to teachers' lack of conceptual and instructional knowledge necessary to effectively disseminate mathematical knowledge. Given that mathematics is a one of culture's main stored knowledge, learners can only get to know this cultural and historical knowledge through mediation by more knowledgeable others (MKO), (Vygotsky, 1978). It is crucial that teaching of mathematics be led by teachers knowledgeable in mathematics themselves.

- A strong recommendation is that the teachers of mathematics must be highly knowledgeable about mathematics themselves before the learners can learn from them. Teachers must undergo good and effective professional development programmes that are informed by their instructional as well as mathematical content needs. A thorough needs analysis must be conducted for all mathematics teachers in order to develop programs that are aligned to the teachers needs.

- This research shows unquestionable evidence that learners have persistent and widespread errors and misconceptions on algebra and that they mix mathematical ideas; for example, financial mathematics ideas are linked to sequences and that calculus is evoked each time ideas on functions, polynomials or graphs are raised. This indicates that learners struggle to make sense of mathematics for which little guidance seems to be forthcoming. In the same way it is clear on calculus questions that more than 70% of students showed complete unawareness of calculus in answering calculus questions.

- Above all, this report points to some typical errors common to many learners of mathematics. This report recommends that teachers strive to establish whether these errors are occurring in their classes, and as a matter of fact target them and

other errors not pointed out in this research as a basis for improving teaching and learning of mathematics, on the basis of the identified mathematical needs that the learners have.

- Language is seen as one of the major determinants of learning mathematics; as such it is important to research how the use of home languages could be used to improve the teaching and learning of mathematics. The research found out that questions that were linguistically pregnant were not well answered and a majority of learners performed dismally on such questions.

- This report urges teachers to research how much their teaching can improve if they target learners' misconceptions in mathematics, and stresses the value of diagnosing learners' needs and difficulties in mathematics as a cornerstone for teaching mathematics.

- This research urges policy makers to consider error analysis and diagnosis as a compulsory course for pre- and in-service teachers as it has potential to revolutionise education. This research recommends constant exposure and discussion of learners' misconceptions as fundamental to good mathematics teaching.

- It is suggested that more time be spent in training the markers before they start marking scripts. Markers should attempt the paper that they will be marking in order to be familiar with the questions (this must be done prior to the first day of marking). Markers should be asked to answer the questions themselves before receiving a memorandum. The most competent senior markers should lead discussions on each question. The discussion must outline possible errors and alternative mark allocations, as well as how to deal with the different conceptual and procedural errors that may be encountered. Relevant excerpts from scripts should be provided as examples to illustrate the strategies to all markers. It is suggested that the markers of each question discuss the feasibility of different versions and make sure that they fully understand the question before marking commences.

F. SUMMARY OF THE SCRIPT ANALYSIS PHYSICAL SCIENCES

Physical Sciences Examination Script Analysis project team:

Coordinator: Umesh Ramnarain	
Researchers:	
Celeste van Niekerk	Aleyamma Joseph
Xolani Mhlungu	Morgen Musimbo
Dories Thantsha	Martin Mudindo

A total of 1000 scripts that were randomly selected were analysed. This constituted approximately 5% of the candidates who wrote the 2008 Physical Sciences examination.

The analysis of the 2008 national examination scripts focussed on the following aspects:

1. The readability of the examination items.

2. The misconceptions in learner responses to the physics items.

3. The performance of learners in chemical representations.

FINDINGS AND RECOMMENDATIONS

Readability of physics examination items

An analysis of the questions in the examination paper identified text which had linguistic features synonymous with scientific writing. These features included complex sentences, density of information, subordinate clauses, unfamiliar words, and ambiguous phrases. The study revealed that where a question displayed a linguistic feature of scientific writing contributed to students doing an incorrect calculation, focusing on the wrong aspect of the question, repeating a segment of a question, and misinterpreting a word or a phrase. In view of this finding, the validity of assessing for scientific knowledge and skills in a written examination becomes questionable, and we recommend that greater attention needs to be paid to the language of science in the classroom.

Misconceptions in physics

The study revealed that students have a naïve, superficial and fragmentary understanding of scientific phenomena that is based upon their experiential knowledge. These conceptions were not theoretically grounded as students appeared to string together pieces of knowledge in response to contextual feature of the problem situation. We believe the method of analyzing the explanations does offer researchers and teachers a reliable and efficient way by which written student explanations can be probed for misconceptions.

The performance of learners in chemical representations

There was no significant difference in the performance of learners on tasks involving macroscopic, symbolic and submicroscopic representations. However, learners did perform significantly worse on tasks that demanded a transition from one form of representation to another. In particular, learners experienced difficulty on tasks where they needed to shift from submicroscopic and symbolic levels of representation, and on tasks involving macroscopic, symbolic and microscopic transitions.

ANNEXURE

BACKGROUND TO THE EAB/GDE/UJ SCRIPT ANALYSIS PROJECT

1. Focus and rationale

Poor performance in the Matriculation examinations is an ongoing concern for students, parents, policy makers as well as other stakeholders. Government actions to improve learner performance include a range of responses, such as improving infrastructure, revising curricula, and improving teacher education. These actions depend on a deep understanding of why learners perform poorly, and what the factors are which contribute to poor performance. While there are obviously a wide range of factors involved, one of the least researched areas is examination scripts.

Our belief is that a detailed analysis of scripts will provide insight into the nature of the difficulties learners experience in the learning process, and enable the development of interventions which would improve learner performance. This research therefore involves a detailed analysis of senior certificate examination scripts, to see what we can learn through an analysis of the trends and patterns of written responses and errors.

Research of this nature will assist the GDE EAB group and the University of Johannesburg, as educational researchers and trainers of teachers, in providing feedback to the system on what the key problems are that learners are experiencing, whether these are conceptual problems or content that has not been covered. It is also hoped that the research may assist Umalusi in improving the quality of examination papers as well as of the marking and moderations systems. This project should be seen as a first step in building a more systematic and rigorous analysis of our examinations and systems of feedback to our schools.

The analysis of samples of scripts from six subjects involved the development and use of analysis grids, focusing on knowledge levels, gaps in content knowledge, patterns of reasoning, understanding of questions, examination writing approaches, common errors, etc.

2. Project purpose and aims

The purpose of this Script Analysis Project is to analyze samples of GDE matric/NSC examination papers of 2008 in order to make recommendations for the improvements of learner performance in various subjects.

3. Scope of the research

The SAP went through different phases:

1. a preparatory phase where the project was set up and analysis frameworks developed [December 2008 to April 2009].
2. a phase two where examination scripts written at the end of 2008 were analyzed systematically [May 2009 to December 2009].
3. a phase three where analyses were validated and refined [2010].

Based on expertise available at UJ, the project focused on Mathematics, Maths Literacy, Physical Sciences, Life Sciences, History, English first additional language, and Computer Applications Technology.

4. Research questions

The guiding research questions included the following, varied by subject depending on researcher background and expertise:

a) What are the levels of understanding of examination questions/assessment tasks [within and across subjects]? What examination writing approaches do learners use? What differences are there in the language of examination papers, memoranda and examination scripts?

b) What distinguishes high performance from low/failure performance? What are the error patterns/knowledge gaps displayed in examination papers? How do learners respond to cognitive demands of examination papers?

c) What are the various ways in which language co-determine examination performance? How is context and local experiences reflected in examination papers and answers?

5. Research process

The setting up phase involved a thorough analysis of international literature on script analysis and school exit examination standards. Relevant Department of Education (DoE) policy documents and research reports by the DoE and Umalusi have been analysed. Based on the document analysis frameworks for the script analyses were developed per subject exam paper.

The actual analyses were mostly qualitative in nature, depending on the analysis frameworks, to include foci on knowledge/content, interpretation and language use, error patterns, textual presentation, and the “cognitive operations” displayed in examination texts. The analyses also included a detailed study of examination papers, memoranda, and examiner reports.

Apart from the analysis of scripts, some interviews with teachers and visits to a small selection of schools are planned, to develop some understanding of examination preparation practices, and to involve teachers in analyses and the interpretation of findings.

6. Outcomes of the research

The SAP produced six individual subject reports. Each of these includes subject specific script analysis frameworks, relevant to the SA NSC examinations. The individual reports also include findings in terms of trends and patterns identified per paper and question type. Patterns of examination performance in the different subjects, typical errors by learners, gaps in knowledge, language problems, and reasons for examination failures are also described. Implications for the improvement of learner performance are made.

7. Sub-reports:

A. Computer Applications Technology – Mr. Jerry Maseko

B. English first additional language – Prof. Gert van der Westhuizen

C. History – Ms. Kristen van Lelyveld

D. Life Science – Dr. Josef de Beer

E. Mathematics and Maths Literacy – Dr. Kakoma Luneta

F. Physical Science – Dr. Umesh Ramnarain