UMALUSI

COUNCIL FOR QUALITY ASSURANCE IN GENERAL AND FURTHER EDUCATION AND TRAINING

Exemplar Book on Effective Questioning

Life Sciences

Compiled by the Statistical Information and Research (SIR) Unit

March 2018

PREFACE

The National Senior Certificate (NSC) examinations are set and moderated in part using tools which specify the types of cognitive demand and the content deemed appropriate for Life Sciences at Grade 12 level. Until recently, the level of cognitive demand made by a question was considered to be the main determinant of the overall level of cognitive challenge of an examination question.

However, during various examination evaluation projects conducted by Umalusi from 2008-2012, evaluators found the need to develop more complex tools to distinguish between questions which were categorised at the same cognitive demand level, but which were not of comparable degrees of difficulty. For many subjects, for each type of cognitive demand a three-level degree of difficulty designation, *easy, moderate and difficult* was developed. Evaluators first decided on the type of cognitive process required to answer a particular examination question, and then decided on the degree of difficulty, *as an attribute of the type of cognitive demand*, of that examination question.

Whilst this practice offered wider options in terms of *easy, moderate* and *difficult* levels of difficulty for each type of cognitive demand overcame some limitations of a one-dimensional cognitive demand taxonomy, other constraints emerged. Bloom's Taxonomy of Educational Objectives (BTEO) (Bloom, Engelhart, Furst, Hill, & Krathwohl, 1956) and the Revised Bloom's Taxonomy are based on the assumption that a cumulative hierarchy exists between the different categories of cognitive demand (Bloom et al., 1956; Bloom, Hastings & Madaus, 1971). The practice of 'levels of difficulty' did not necessarily correspond to a hierarchical model of increasing complexity of cognitive demand. A key problem with using the level of difficulty as an attribute of the type of cognitive demand of examination questions is that, questions recognised at a higher level of cognitive demand are not necessarily categorised as more difficult than other questions categorised at lower levels of cognitive demand. For example, during analyses a basic recognition or recall question could be considered more difficult than an easy evaluation question.

Research further revealed that evaluators often struggled to agree on the classification of questions at so many different levels. The finer categorization for each level of cognitive demand and the process of trying to match questions to pre-set definitions of levels of difficulty made the process of making judgments about cognitive challenge overly procedural. The complex two-dimensional multi-level model also made findings about the cognitive challenge of an examination very difficult for Umalusi Assessment Standards Committee (ASC) to interpret.

In an Umalusi Report, *Developing a Framework for Assessing and Comparing the Cognitive Challenge of Home Language Examinations* (Umalusi, 2012), it was recommended that the type and level of cognitive demand of a question and the level of a question's difficulty should be analysed separately. Further, it was argued that the ability to assess cognitive challenge lay in experts' abilities to recognise subtle interactions and make complicated connections that involved the use of multiple criteria simultaneously. However, the tacit nature of such judgments can make it difficult to generate a common understanding of what constitutes criteria for evaluating the cognitive challenge of examination questions, despite descriptions given in the policy documents of each subject.

The report also suggested that the Umalusi external moderators and evaluators be provided with a framework for thinking about question difficulty which would help them identify where the main sources of difficulty or ease in questions might reside. Such a framework should provide a common language for evaluators and moderators to discuss and justify decisions about question difficulty. It should also be used for building the capacity of novice or less experienced moderators and evaluators to exercise the necessary expert judgments by making them more aware of key aspects to consider in making such judgments.

The revised Umalusi examination moderation and evaluation instruments for each subject draw on research and literature reviews, together with the knowledge gained through the subject workshops. At these workshops, the proposed revisions were discussed with different subject specialists to attain a common understanding of the concepts, tools and framework used; and to test whether the framework developed for thinking about question difficulty 'works' for different content subjects. Using the same framework to think about question difficulty across subjects will allow for greater comparability of standards across subjects and projects.

An important change that has been made to the revised examination evaluation instrument is that the analysis of *the type of cognitive demand* of a question and analysis of *the level of difficulty* of each question are now treated as two separate judgments involving two different processes. Accordingly, the revised examination evaluation instrument now includes assessment of difficulty *as well as* cognitive demand.

LIST OF ABBREVIATIONS

Abbreviation	Full name
ASC	Assessment Standards Committee
BTEO	Bloom’s Taxonomy of Educational Objectives
CAPS	Curriculum Assessment Policy Statement
DBE	Department of Basic Education
FET	Further Education and Training
IEB	Independent Examinations Board
NSC	National Senior Certificate
NQF	National Qualifications Framework
QAA	Quality Assurance of Assessment
QCC	Qualifications, Curriculum and Certification
SIR	Statistical Information and Research

LIST OF TABLES

ACKNOWLEDGEMENTS

This Life Sciences exemplar book is informed by Umalusi Research Reports of previous years, especially the report by Reeves (Umalusi, 2012) titled '*Developing a framework for assessing and comparing the cognitive challenge of Home Language examinations*'.

In addition, Life Sciences subject experts and practitioners are acknowledged for their contribution to the content of this exemplar book. Included in this group are: Umalusi External Moderators and Maintaining Standards Subject Teams and Team Leaders; together with the South African Comprehensive Assessment Institute and the Independent Examinations Board (IEB) Examiners and Internal Moderators.

We also acknowledge the contributions of the members of the Umalusi Quality Assurance of Assessment (QAA); Qualifications, Curriculum and Certification (QCC) and Statistical Information and Research (SIR) Units. We specifically acknowledge the contribution made by the individuals listed below:

- Ms Agnes Mohale, who was responsible for the management and coordination of the Exemplar Books Project.
- Dr Cheryl Reeves, who was responsible for developing the framework that underpinned the design of the exemplar books.
- Mr Thapelo Rangongo, Ms Sisanda Loni and Ms Shannon Doolings for their assistance and support in the administration of the project.
- The review team included the following members: Mr Xolile Taba, Prof. Shaheed Hartley, Prof. Rajendran Govender, Mr Doctor Phokwani, Mr Roy Mathe and Dr Stephan Mchunu.

This exemplar book was prepared by Dr Anna Crowe.

TABLE OF CONTENTS

1. INTRODUCTION

The rules of assessment are essentially the same for all types of learning because, to learn is to acquire knowledge or skills, while to assess is to identify the level of knowledge or skill that has been acquired (Fiddler, Marienau & Whitaker, 2006). Nevertheless, the field of assessment in South Africa and elsewhere in the world is fraught with contestation. A review of the research literature on assessment indicates difficulties, misunderstanding and confusion in how terms describing educational measurement concepts, and the relationships between them, are used (Frisbie, 2005).

Umalusi believes that if all role players involved in examination processes can achieve a common understanding of key terms, concepts and processes involved in setting, moderating and evaluating examination papers, much unhappiness can be avoided. This exemplar book presents a particular set of guidelines for both novice and experienced Life Sciences national examiners, internal and external moderators, and evaluators to use in the setting, moderation and evaluation of examinations at the National Senior Certificate (NSC) level.

The remainder of the exemplar book is organised as follows: First, the context in which the exemplar book was developed is described (Part 2), followed by a statement of its purpose (Part 3). Brief summaries of the roles of moderation and evaluation (Part 4) and cognitive demand (Part 5) an assessment. Examination questions selected from the NSC Life Sciences examinations of assessment bodies, the Department of Basic Education (DBE), and/or the Independent Examinations Board (IEB) are used to illustrate how to identify different levels of cognitive demand as required by the Curriculum and

Assessment Policy Statement (CAPS) Life Sciences document (Part 6). Part 7 explains the protocols for identifying different levels of difficulty within a question paper. Application of the Umalusi framework for determining difficulty described in Part 7 is illustrated, with reasons, by another set of questions from a range of Life Sciences examinations (Part 8). Concluding remarks complete the exemplar book (Part 9).

2. CONTEXT

Umalusi has the responsibility to quality assure qualifications, curricula and assessments of National Qualification Framework (NQF) Levels 1 - 5. This is a legal mandate assigned by the *General and Further Education and Training Act (Act 58 of 2001)* and the *National Qualification Framework Act (Act 67 of 2008)*. To operationalize its mandate, Umalusi, amongst other things, conducts research and uses the findings of this research to enhance the quality and standards of curricula and assessments.

Since 2003, Umalusi has conducted several research studies that have investigated examination standards. For example, Umalusi conducted research on the NSC examinations, commonly known as 'Matriculation' or Grade 12, in order to gain an understanding of the standards of the new examinations (first introduced in 2008) relative to those of the previous NATED 550 Senior Certificate examinations (Umalusi, 2009a, 2009b). Research undertaken by Umalusi has assisted the organisation to arrive at a more informed understanding of what is meant by assessing the cognitive challenge of the examinations and of the processes necessary for determining whether the degree of cognitive challenge of examinations is comparable within a subject, across subjects and between years.

Research undertaken by Umalusi has revealed that different groups of examiners, moderators and evaluators do not always interpret cognitive demand in the same way, posing difficulties when comparisons of cognitive challenge were required. The research across all subjects also showed that using the type and level of cognitive demand of a question *only* as measure for judging the cognitive challenge of a question is problematic because cognitive demand levels on their own do not necessarily distinguish between degrees of difficulty of questions.

The new Umalusi framework for thinking about question difficulty described in this exemplar book is intended to support all key role players in making complex decisions about what makes a particular question challenging for Grade 12 examination candidates.

3. THE PURPOSE OF THE EXEMPLAR BOOK

The overall goal of this exemplar book is to ensure the consistency of standards of examinations across the years in the Further Education and Training (FET) sub-sector and Grade 12, in particular. The specific purpose is to build a shared understanding among teachers, examiners, moderators, evaluators, and other stakeholders, of methods used for determining the type and level of cognitive demand as well as the level of difficulty of examination questions.

Ultimately, the common understanding that this exemplar book seeks to foster is based on the premise that the process of determining the type and level of cognitive demand of questions and that of determining the level of difficulty of examination questions are two separate judgements involving two different processes, both necessary for evaluating the cognitive challenge of examinations. This distinction between cognitive demand and difficulty posed by questions needs to be made in the setting, moderation, evaluation and comparison of Life Sciences examination papers.

The exemplar book includes an explanation of the new Umalusi framework which is intended to provide all role-players in the setting of Life Sciences examinations with a common language for thinking and talking about question difficulty. The reader of the exemplar book is taken through the process of evaluating examination questions; first in relation to determining the type and level of cognitive demand made by a question, and then in terms of assessing the level of difficulty of a question. This is done by providing examples of a range of questions which make different types of cognitive demands on candidates, and examples of questions at different levels of difficulty.

Each question is accompanied by an explanation of the reasoning behind why it was judged as being of a particular level of cognitive demand or difficulty, and the reasoning behind the judgements made is explained. The examples of examination questions provided were sourced by Life Sciences evaluators from previous DBE and the IEB Life Sciences question papers, pre- and post- the implementation of CAPS during various Umalusi workshops.

This exemplar book is an official document. The process of revising the Umalusi examination evaluation instrument and of developing a framework for thinking about question difficulty for both moderation and evaluation purposes has been a consultative one, with the DBE and the IEB assessment bodies. The new framework for thinking about question difficulty is to be used by Umalusi in the moderation and evaluation of Grade 12 Life Sciences examinations, and by all the assessment bodies in the setting of the question papers, in conjunction with the CAPS documents.

4. MODERATION AND EVALUATION OF ASSESSMENT

A fundamental requirement, ethically and legally, is that assessments are fair, reliable and valid (American Educational Research Association [AERA], American Psychological Association [APA] and National Council on Measurement in Education [NCME], 1999). Moderation is one of several quality assurance assessment processes aimed at ensuring that an assessment is fair, reliable and valid (Downing & Haladyna, 2006). Ideally, moderation should be done at all levels of an education system, including the school, district, provincial and national level in all subjects.

The task of Umalusi examination **moderators** is to ensure that the quality and standards of a particular examination are maintained each year. Part of this task is for moderators to alert examiners to details of questions, material and/or any technical aspects in examination question papers that are deemed to be inadequate or problematic and that therefore, challenge the validity of that examination. In order to do this, moderators need to pay attention to a number of issues as they moderate a question paper – these are briefly described below.

Moderation of the technical aspects of examination papers includes checking correct question and/or section numbering, and ensuring that visual texts and/or resource material included in the papers are clear and legible. The clarity of instructions given to candidates, the wording of questions, the appropriateness of the level of language used, and the correct use of terminology need to be interrogated. Moderators are expected to detect question predictability, for example, when the same questions regularly appear in different examinations, and bias in examination papers. The adequacy and accuracy of the marking memorandum (marking guidelines) need to be checked to ensure that they reflect and correspond with the requirements of each question asked in the examination paper being moderated.

In addition, the task of moderators is to check that papers adhere to the overall examination requirements as set out by the relevant assessment body with regard to the format and structure (including the length, type of texts or reading selections prescribed) of the examination. This includes assessing compliance with assessment requirements with regard to ensuring that the content is examined at an appropriate level and in the relative proportions (weightings) of content and/or skills areas required by the assessment body.

The role of Umalusi examination **evaluators** is to perform analysis of examination papers after they have been set and moderated and approved by the Umalusi moderators. This type of analysis entails applying additional expert judgments to evaluate the quality and standard of finalised examination papers before they are written by candidates in a specific year. However, the overall aim of this evaluation is to judge the comparability of an examination against the previous years' examination papers to ensure that consistent standards are being maintained over the years.

The results of the evaluators' analyses, and moderators' experiences provide the Umalusi Assessment Standards Committee (ASC) with valuable information which is used in the process of statistical moderation of each year's examination results. Therefore, this information forms an important component of essential qualitative data informing the ASC's final decisions in the standardisation of the examinations.

In order for the standardisation process to work effectively, efficiently and fairly, it is important that examiners, moderators and evaluators have a shared understanding of how the standard of an examination paper is assessed, and of the frameworks and main instruments that are used in this process.

5. COGNITIVE DEMANDS IN ASSESSMENT

The *Standards for educational and psychological testing* (AERA, APA, & NCME, 1999) require evidence to support interpretations of test scores with respect to cognitive processes. Therefore, valid, fair and reliable examinations require that the levels of cognitive demand required by examination questions are appropriate and varied (Downing & Haladyna, 2006). Examination papers should not be dominated by questions that require reproduction of basic information, or replication of basic procedures, and under-represent questions invoking higher level cognitive demands.

Accordingly, the Grade 12 CAPS NSC subject examination specifications state that examination papers should be set in such a way that they reflect proportions of marks for questions at various level of cognitive demand. NSC examination papers are expected to comply with the specified cognitive demand levels and weightings. NSC examiners have to set and NSC internal moderators have to moderate examination papers as reflecting the proportions of marks for questions at different levels of cognitive demand as specified in the documents. Umalusi's external moderators and evaluators are similarly tasked with confirming compliance of the examinations with the CAPS cognitive demand levels and weightings, and Umalusi's revised examination evaluation instruments continue to reflect this requirement.

Despite that, subject experts, examiners, moderators and evaluators are familiar with the levels and explanations of the types of cognitive demand shown in the CAPS documents, Umalusi researchers have noted that individuals do not always interpret and classify the categories of cognitive demand provided in the CAPS the same way. In order to facilitate a common interpretation and classification of the cognitive demands made by questions, the next section of this exemplar book provides a clarification of each cognitive demand level for Life Sciences followed by illustrative examples of examination questions that have been classified at that level of cognitive demand.

6. EXPLANATIONS AND EXAMPLES OF QUESTIONS ASSESSED AT THE DIFFERENT COGNITIVE DEMAND LEVELS IN THE LIFE SCIENCES TAXONOMY ACCORDING TO CAPS

The taxonomies of cognitive demand for each school subject in the CAPS documents are mostly based on the Revised Bloom's Taxonomy (Anderson and Krathwohl, 2001) but resemble the original Bloom's taxonomy in that categories of cognitive demand are arranged along a single continuum.

Bloom's Taxonomy of Educational Objectives (BTEO) (Bloom, Engelhart, Furst, Hill, & Krathwohl, 1956) and the Revised Bloom's Taxonomy imply that each more advanced or successive category of cognitive demand subsumes all categories below it. The CAPS Taxonomies of Cognitive Demand make a similar assumption (Crowe, 2012).

Note:
In classifying the type and level of cognitive demand, each question is classified at the highest level of cognitive process involved. Thus, although a particular question involves recall of knowledge, as well as comprehension and application, the question is classified as an 'analysis' question if that is the highest level of cognitive process involved. If evaluating' is the highest level of cognitive process involved, the question as a whole should be classified as an 'evaluation' question. On the other hand, if one of more sub-sections of the question and the marks allocated for each sub-section can stand independently, then the level of cognitive demand for each sub-section of the question should be analysed separately.

The CAPS documents for many subjects also give examples of descriptive verbs that can be associated with each of the levels of cognitive demand. However, it is important to note that such 'action verbs' can be associated with more than one cognitive level depending on the context of a question.

The Life Sciences CAPS document states that Grade 12 NSC Life Sciences examination papers should examine three levels of cognitive demand (Table 1).

TABLE 1: THE LIFE SCIENCES TAXONOMY OF COGNITIVE DEMAND LEVELS FOR THE LIFE SCIENCES NSC EXAMINATIONS

Level of cognitive demand	**Knowing** science	**Understanding** science	**Applying** scientific knowledge	**Evaluating, Analysing** and **Synthesising** scientific knowledge

Source: CAPS (DBE, 2011a, p.67)

It is important to note that the four categories of cognitive demand required by this Life Sciences Taxonomy of Cognitive Demand do not explicitly allocate these levels to either lower-, middle- or higher order cognitive levels as is done in most other NSC subjects. Also, while this taxonomy uses categories of cognitive demand that are recognisable in the BTEO (Bloom et al.,1956) the categories are not used consistently with respect the descriptions in the BTEO, nor are they used consistently within the CAPS document (DBE, 2011a).

To facilitate reading of this section, each of the cognitive demand levels in the Life Sciences Taxonomy shown in Table 1 above is explained using BTEO (Bloom et al.,1956). Each explanation is followed by **three** examples of questions from previous Life Sciences NSC examinations classified at each of the levels of cognitive demand. These examples were selected to represent the **best and clearest** examples of each level of cognitive demand that the Life Sciences experts could find. The discussion below each example question explains the reasoning processes behind the classification of the question at that particular type of cognitive demand (Table 2 to Table 5).

Note:
Be mindful that analyses of *the level of cognitive process* of a question and *the level of difficulty* of each question are to be treated as two separate judgments involving two different processes. Therefore, whether the question is easy or difficult should not influence the categorisation of the question in terms of the type and level of cognitive demand. Questions should NOT be categorised as higher order evaluation/synthesis questions because they are difficult questions. Some questions involving the cognitive process of recall or recognition may be more difficult than other recall or recognition questions. Not all comprehension questions are easier than questions involving analysis or synthesis. Some comprehension questions may be very difficult, for example explanation of complex scientific processes. For these reasons you need to categorise the level of difficulty of questions separately from identifying the type of cognitive process involved.

TABLE 2: EXPLANATION AND EXAMPLES THE *KNOWING SCIENCE* LEVEL OF COGNITIVE DEMAND QUESTIONS USING THE LIFE SCIENCES TAXONOMY

Knowing science

Knowing science refers to the ability to **recognise** or **recall explicit information** such as details, facts, formulae, terms, definitions, concepts, procedures, or representations **from memory** or **from material provided** in the question. No evidence of understanding is required.

Here**, explicit information** refers to the acquired Life Sciences knowledge as stated in CAPS (DBE, 2011a), or given in the question.

Example 1:

Question 1.1.3, 2014 DBE Paper 1 (DBE, 2014a)

1.1 Various options are given as possible answers to the following questions. Choose the answer …

1.1.3 The number of chromosomes found in a human sperm cells is …

A. 23
B. 22
C. 46

D. 47

(2)

Discussion:

This is a multiple-choice question that requires candidates to recognise a biological fact, i.e., the correct number of chromosomes in a human sperm cell from the four options given. This question is categorised as 'Knowing science'.

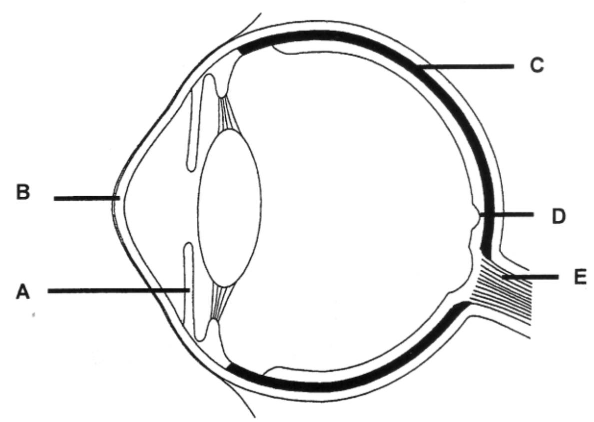
Memorandum/Marking guidelines

1.1.3 A √√ (2)

Example 2:

Question 1.4, 2014 DBE Paper 1 (DBE, 2014a)

1.4 The diagram below represents the structure of the human eye.



Structure of the human eye

Give the LETTER and the NAME of the part which:

1.4.1 Regulates the amount of light entering the eye.

1.4.2 Supplies food and oxygen to the eye.

1.4.3 Transmits impulses to the brain.

1.4.4 Contains cones and is the area of clearest vision.

1.4.5 Assists in the refraction of light rays. (5 x 2)

Discussion:

Candidates are required to recognise the structures and recall the name of the structures and matching functions given with the appropriate labels. This is therefore categorised as 'Knowing science'.

Memorandum/Marking guidelines

1.4.1 A √ iris √
1.4.2 C √ choroid √
1.4.3E √ optic nerve √
1.4.4 D √ fovea √ / macula lutea
1.4.5 B √ cornea √ (5 x 2)

Example 3:

Question 1.2, 2014 DBE Paper 2 (DBE, 2014b)

1.2 Give the correct biological term for each of the following descriptions.

1.2.1 An allele that does not influence the phenotype when found in the heterozygous condition. (1)

1.2.2 A section of a DNA molecule that codes for a specific characteristic. (1)

Discussion:

Candidates are given familiar definitions or descriptions of biological terms. They are expected to recognise the description and recall the associated term from memory. This question is therefore categorised as 'Knowing science'.

Memorandum/Marking guidelines

1.2.1 recessive √ (1)

1.2.2 gene √ (1)

TABLE 3: EXPLANATION AND EXAMPLES OF THE *UNDERSTANDING SCIENCE* LEVEL OF COGNITIVE DEMAND QUESTIONS USING THE LIFE SCIENCES TAXONOMY

UNDERSTANDING SCIENCE Understanding science refers to the ability to show evidence of an **understanding** of recognised or recalled explicit information learned or given in the question (Knowing science, above). **Understanding** refers to the ability to **identify** and **express** or **explain relationships between ideas,** and the way in which concepts are organised or structured, in ways that make sense biologically. Students can demonstrate basic understanding by reorganising or stating **what has been learned** or **given in a question in a different way**, or **in a different form** (e.g., diagram to words, words to diagrams, tables to graphs etc.), or by **explaining, categorising, exemplifying, paraphrasing**, **comparing**, or **consolidating** some ideas, concepts or procedures which have been learned or given in the question.
Example 1: <u>Question 3.3.3, 2014 DBE Paper 2</u> (DBE, 2014b) Explain TWO advantages of bipedalism. (4) **<u>Discussion</u>:** Students are required to state the anatomical characteristics associated with bipedalism (knowledge what has been learned) and to explain the relationship between each characteristic and how it confers advantage on a bipedal organism (not required to be learned). This answer assesses if students are able to demonstrate an understanding of the crucial biological relationship between structure and their function. This question is therefore classified as 'Understanding science'. Note – this question is not classified as 'Applying scientific knowledge' because the context is neither new, nor unfamiliar.
<u>Memorandum/Marking guidelines</u> 3.3.3 • Frees the arms√ so that they can carry offspring√ /tools/food/manipulate things. • Allows ability to see further √ to spot danger √/food.

- Reduces the surface area exposed to the sun √ so less heat absorbed√ /less heat lost / thermoregulation.
- Expose the genitals √ to attract the opposite sex √
- Efficient locomotion √ allows to travel longer distances √

(Mark first TWO only) Any 2 x 2 (4)

Example 2:

Question 2.2.2, 2012 Version 1 Paper 2 (DBE, 2012a)

2.2 In an investigation a learner was asked to put a cotton thread through the eye of a needle 10 times with both eyes open and then with only the right eye open. This was done under the same light intensity and at a distance of 50 cm from the eyes. The results of the time taken to thread the needle are shown in the table below.

	TIME TAKEN TO THREAD THE NEEDLE IN SECONDS (s)	
Attempts	**Two eyes open**	**Only right eye open**
1	12	38
2	12	35
3	10	37
4	11	36
5	9	34
6	9	33
7	10	30
8	8	31
9	7	29
10	7	28

2.2.2 State a general conclusion that can be drawn from the results above. (2)

Discussion:

The candidates are required to draw an obvious logical conclusion from presented information. They do not need to analyse the data in any detail, but they do more than just describe the data. The data for two eyes versus one eye opened needs to be compared before the data is consolidated into a conclusion. This question requires candidates to read and understand the information given – they are not required to explain their general conclusion with respect to how the eye functions. Therefore, this question is classified as the cognitive level 'Understanding science'.

Memorandum/Marking guidelines

2.2.2 It takes more time √ to thread a needle with one eye open compared to having both eyes open. √

OR

It takes less time √ to thread a needle with both eyes open compared to having one eye open. √

OR

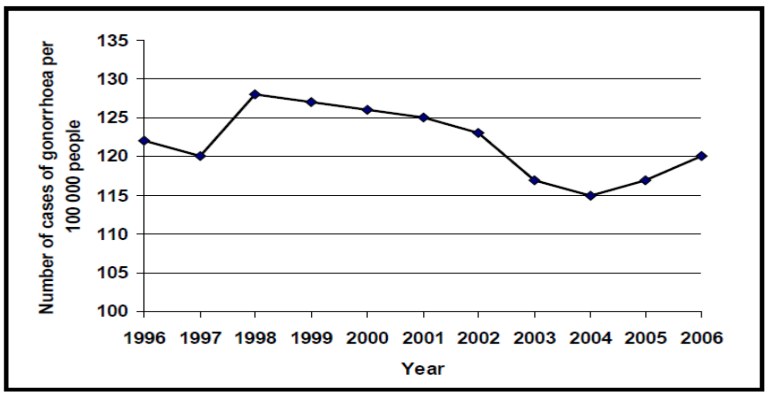
The more attempts undertaken √ to thread the needle the less time it takes. √

Any (1 x 2)(2)

Example 3:

Question 3.3.2, 2012 Version 2 Paper 1 (DBE, 2012b)

The graph below shows the number of cases of gonorrhoea in a country per 100 000 people between 1996 and 2006. Describe the pattern of cases of gonorrhoea between 1997 and 2002.



(2)

Discussion:

This question requires candidates to change information from one form of representation (a graph) to another (words). To do this, candidates need to interpret the graph, one of the ways in which students demonstrate understanding. Therefore, this question is classified as 'Understanding science'.

Memorandum/Marking guidelines

3.3.2 After a sharp increase √ in 1997 [to 1998] the numbers gradually decreased √ until 2002. (2)

TABLE 4: EXPLANATION AND EXAMPLES OF THE *APPLYING SCIENCE* LEVEL OF COGNITIVE DEMAND QUESTIONS USING THE LIFE SCIENCES TAXONOMY

APPLYING SCIENTIFIC KNOWLEDGE

Applying scientific knowledge refers to the ability to **use knowledge** (Knowing science, and Understanding science, above) **in a new way** (not explicated in CAPS) or to **transfer knowledge from a familiar context** (explicated in CAPS) **to an unfamiliar context**.

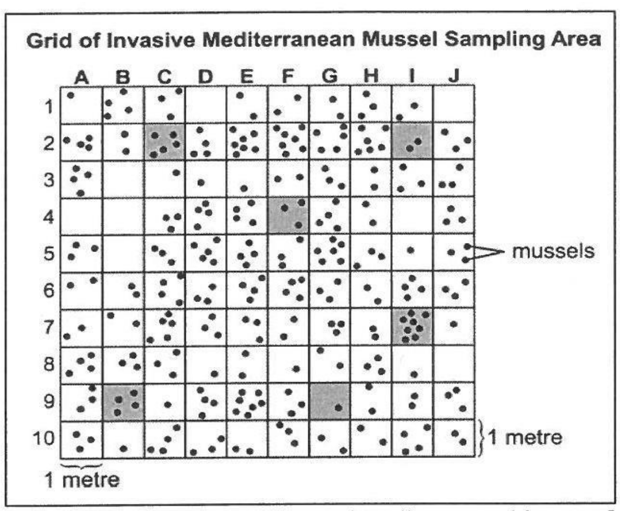
Students can demonstrate applying scientific knowledge when they can **perform** or follow a **standard procedure** (rule, method or operation) in a new context.

***Note**:* Applying scientific knowledge here excludes the evaluation, analysis and synthesis of scientific knowledge - the forth group described in Table 5.

Example 1:

Question Q2.4, 2017 IEB Supplementary Paper 2

2.4 Ecologists wanted to gain an idea of how many invasive alien mussels were growing on the rocks at a breeding site. They marked out an area as seen below and counted the mussels in a number of random samples.



[Source: <http://www.oocities.org>]

2.4.2 Areas where quadrats were placed are shaded on the diagram above. Use these quadrats to estimate the number of mussels in the marked-out area. Show all working. (4)

Discussion:

Candidates are required to perform a basic procedure, i.e., carry out a series of calculations. In order to perform these calculations, they need to extract the relevant details to perform the calculations from the diagram given, which is an unfamiliar context. These cognitive processes are included in the cognitive level 'Applying scientific knowledge'.

Memorandum/Marking guidelines

2.4.2 Total 6 quadrats/m² = 5 + 2 + 3 + 4 + 8 + 1 = 23 √

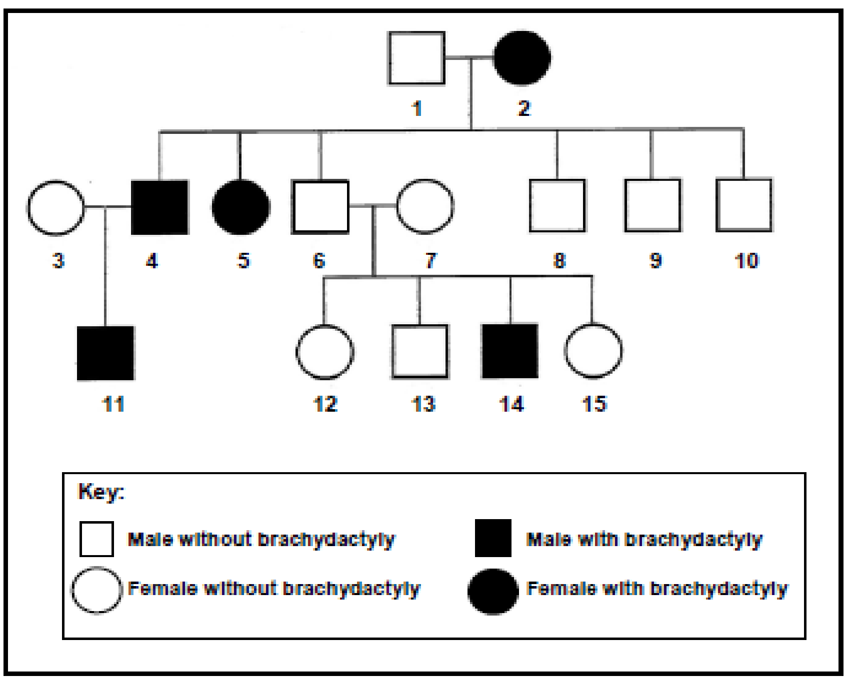
Average in 1 m² = 23 ÷ 6 = 3,8 √

Total area = 10 m² x 10 m² = 100 m² √

Total n. Of mussels = 100 x 3,8 = 380 √ (4)

Example 2:

Question 1.5.1, 2012 DBE Version 2 Paper 1 (DBE, 2012b)



1.5.1 Study the pedigree diagram above and state whether brachydactyly is caused by a dominant or recessive allele. (2)

Discussion:

Candidates are required to apply their understanding of pedigree diagrams (a known context) to an unfamiliar context (inheritance of brachydactyly). This question requires candidates to follow a standard procedure (substituting genotypes and phenotypes) and apply their understanding of the concepts and procedure to interpretation of the given pedigree. Therefore, this question is categorised as 'Applying scientific knowledge'.

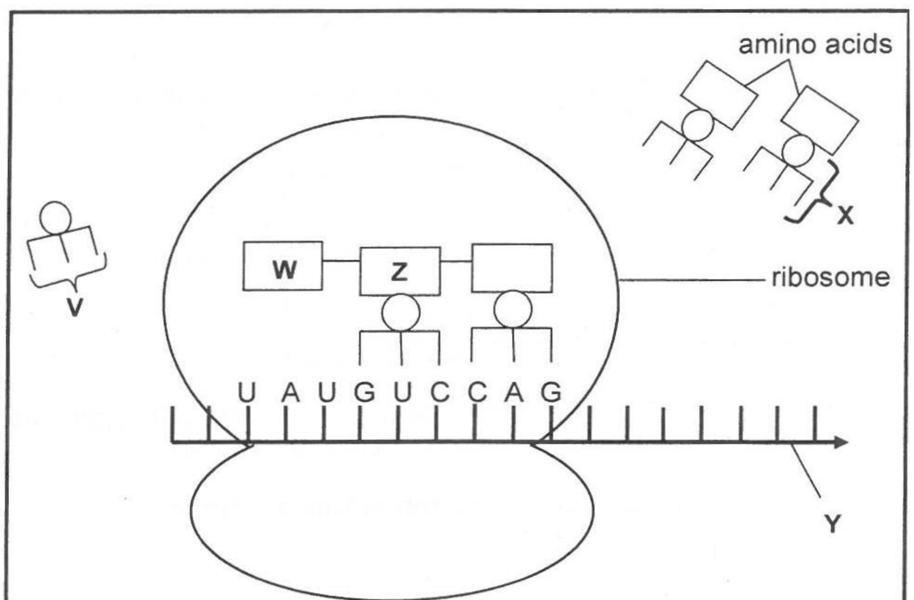
Memorandum/Marking guidelines

1.5.1 Recessive √√ (2)

Example 3:

Question 2.1.5 ,2014 DBE Paper 2 (DBE, 2014b)

2.1 Study the diagram below which shows a part of the process of protein synthesis.



Protein synthesis at the ribosome

2.1.5 Use the table below to identify amino acid **W**.

tRNA	Amino acid
GUC	glutamine

UAA	isoleucine
AUA	tyrosine
CCC	glycine
GGG	proline
CAG	valine

(2)

Discussion:

Candidates are required to demonstrate that they understand the learned process by which specific amino acids are combined to form a polypeptide which is given in the diagram – the polypeptide in an unseen structure and is therefore unfamiliar. These cognitive processes are categorised as 'Applying scientific knowledge'.

Memorandum/Marking guidelines

2.1.5 Tyrosine √√ (2)

TABLE 5: EXPLANATION AND EXAMPLES OF THE *EVALUATING, ANALYSING AND SYNTHESISING SCIENCE* LEVEL OF COGNITIVE DEMAND QUESTIONS USING THE LIFE SCIENCES TAXONOMY

ANALYSING, EVALUATING AND **SYNTHESISING** SCIENTIFIC KNOWLEDGE

Analysing, evaluating and synthesising refer to the ability to **engage in more abstract interpretations** or **reasoning,** using knowledge (Knowing science, Understanding science, and Applying scientific knowledge, above) in **new** or **unfamiliar** **contexts**. For example: generating hypotheses; interpreting relationships, patterns, trends and results; predicting consequences; deducing reasons; inferring causes; suggesting plausible explanation and drawing conclusions.

Note: Creating solutions to open-ended or novel problems, engaging in original thought or generating and supporting one's own ideas or arguments are not explicitly catered for in the CAPS Life Sciences Taxonomy of Cognitive Demand. Questions requiring theses cognitive skills should be included in this category.

Example 1:

Question 2.4, 2016 DBE Supplementary, Paper 2 (DBE, 2016a)

2.4 Read the passage below and answer the questions that follow.

USE OF STEM CELLS

Dr Orly Lachan-Kaplan of Monash Immunology and Stem Cell Laboratories has used stem cell manipulation to create an ovary-like structure containing ova. Although it is not yet clear if the cells of this ovary-like structure are functional. She hopes that this method can be used to develop functional human ova.

[Source: http//monash.edu/news/releases/308]

2.4.3Explain ONE possible advantage of creating an ovary-like structure. (3)

Discussion:

To answer this question candidates have to read the text and link this to what they have learned about stem cells to understand what an ovary-like structure is. They then have to infer from this understanding what the advantage of this technology is. This question is therefore categorised as 'Evaluating, analyzing and synthesising scientific knowledge'.

Memorandum/Marking guidelines

i.

to produce ova √ which could be used in cases where females do not have functional ovaries √ and are therefore infertile √ and thereby allowing them to have children√ Any 3 (3)

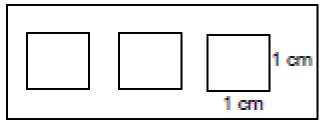
Example 2:

Question 2.2.6, 2010 DBE Paper 2 (DBE, 2010)

2.2 A group of learners performed the following investigation to measure the presence of particles which cause pollution in the air in their town.

The following procedure was followed:

1. Three squares of 1 cm^2 were drawn on each of three glass microscope slides, using a permanent marking pen, as shown in the diagram below.



Microscope slide

2. The other side of each microscope slide was covered with a thin layer of petroleum jelly/Vaseline.
3. The microscope slides were placed, with the side covered with petroleum jelly/Vaseline facing up, in three different outdoor locations A, B and C, and left for one week.

   Location A: Central area in the town
   Location B: 20 km away from the centre of town
   Location C: A fishing spot 40 km away from the centre of town

4. After a week the slides were collected and examined using a hand lens (magnification x20).
5. The visible particles in each square were counted and recorded in a table.
6. The average number of visible particles per square (1 cm^2) for each of the three locations were then calculated.

2.2.6 Name THREE ways in which the validity of the investigation [above] can be improved. (3)

Discussion:

Candidates are required to make critical descriptions of an investigation that is described in the question. They have to apply their knowledge of scientific method in order to make suggestions as to what improvements need to be made. To judge and critique covers the cognitive process evaluate. This question is therefore categorised as 'Evaluating, analysing and synthesising scientific knowledge'.

Memorandum/Marking guidelines

- All other environmental conditions in A, B and C must be similar √ /the slides must be similarly exposed.
- Increase the number of glass slides √ placed at each location.
- Use more locations between the centre of town and the fishing spot √ /repeat the investigation.
- Decrease the time √ of exposure of the slides from a week to daily.
- Increase the period √ of time /different seasons.
- Use one counter √ /trained counters.
- Ensure that non-particle pollutants and material sticking on the slide do not interfere √ with the particle collection.
- Use more advanced equipment √ to collect and measure data.

(Mark first THREE only) (3)

Example 3: <u>Question 1.7.4, 2017 IEB Supplementary Paper 2</u> (IEB 2017) 1.7 In 2007 lodge owners attempted to help the wildlife survive a crippling drought by digging waterholes. 1.7.4 Do you think that the establishment of artificial waterholes is interfering with the carrying capacity of the environment? Explain your answer. (2) **<u>Discussion</u>**: Candidates can take one of two positions about a context not explicitly stated in CAPS, and generate their own ideas to argue for the position that they have chosen This question is therefore categorised as 'Evaluating, analysing and synthesizing scientific knowledge'.
<u>Memorandum/Marking guidelines</u> 1.7.4 YES, more zebra surviving will place extra pressure on grass √; this can degrade the grass so that it cannot re-grow √; whole ecosystem will be under pressure and carrying capacity will go down √; this will have long-term implications for herbivores/other grazers could die out too. √ NO, it will help zebra survive only a short while as they look for other grazing √; a lack of grass, which will be in short supply √ will ultimately be the carrying capacity factor/environmental pressure/causing their death √; a few extra waterholes in such a large area √ will have no impact on long-term survival √ ; lack of grass will cause them to die. (Two good facts for the same position or one well explained) (2)

To accomplish the goal of discriminating between high achievers, those performing very poorly, and all candidates in between, examiners need to vary the challenge of examination questions. Until recently, the assumption has been that 'alignment' with the allocated percentage of marks for questions at the required cognitive demand levels meant that sufficient examination questions were relatively easy; moderately challenging; and difficult for candidates to answer.

However, research and candidate performance both indicate that a range of factors other than type of cognitive demand contribute to the cognitive challenge of a question. Such factors include the level of content knowledge required, the language used in the question, and the complexity or number of concepts tested. In other words, cognitive demand levels on their own do not necessarily distinguish between degrees of difficulty of questions.

This research helps, to some extent, explain why, despite that some NSC examination papers have complied with the specified cognitive demand weightings stipulated in the policy, they have not adequately distinguished between candidates with a range of academic abilities in particular between higher ability candidates. As a result, examiners, moderators and evaluators are now required to assess the difficulty of level of each examination question in addition to judging its cognitive demand.

Section 7 explains the new protocol introduced by Umalusi for analysing examination question difficulty.

7. ANALYSING THE LEVEL OF DIFFICULTY OF EXAMINATION QUESTIONS

When analysing the level of difficulty of each examination question, there are six important protocols to note. These are:

1. Question difficulty is **assessed independently** of the type and level **of cognitive demand**.
2. Question difficulty is assessed against **four levels of difficulty**.
3. Question difficulty is determined against the assumed capabilities of the **ideal** '**envisaged'** Grade 12 Life Sciences NSC examination **candidate**.
4. Question difficulty is determined using **a common framework** for thinking about question difficulty.

5. Question difficulty entails **distinguishing unintended sources of difficulty** or ease **from intended sources of difficulty** or ease.
6. Question difficulty entails identifying **differences** in levels of difficulty **within a single question**.

Each of the above protocols is individually explained and discussed below.

7.1 Question difficulty is assessed independently of the type and level of cognitive demand

As emphasised earlier in this exemplar book, the revised Umalusi NSC examination evaluation instruments separate the analysis of the type of cognitive demand of a question from the analysis of the level of difficulty of each examination question. Cognitive demand describes the *type of cognitive process* that is required to answer a question, and this does not necessarily equate or align with the *level of difficulty* of other aspects of a question, such as the difficulty of the content knowledge that is being assessed. For example, a recall question can ask a candidate to recall very complex and abstract scientific content. The question would be categorised as Level 1 in terms of the cognitive demand taxonomy but may be rated as 'difficult' (Level 3 Table 7 below).

> ***Note:***
> Cognitive demand is just one of the features of a question that can influence your comparative judgments of question difficulty. The type and level of cognitive process involved in answering a question does not necessarily determine how difficult the question would be for candidates. Not all evaluation/synthesis/analysis questions are more difficult than questions involving lower-order processes such as comprehension or application.

7.2 Question difficulty is assessed at four levels of difficulty

The revised Umalusi NSC examination evaluation instruments require evaluators to exercise expert judgments about whether each examination question is 'Easy', 'Moderately challenging', 'Difficult' or 'Very difficult' for the envisaged Grade 12 learner to answer. Descriptions of these categories of difficulty are shown in Table 6.

TABLE 6: LEVELS OF DIFFICULTY OF EXAMINATION QUESTIONS

1	2	3	4
Easy for the envisaged Grade 12 student to answer.	**Moderately challenging** for the envisaged Grade 12 student to answer.	**Difficult** for the envisaged Grade 12 student to answer.	**Very difficult** for the envisaged Grade 12 student to answer. The skills and knowledge required to answer the question allow for the top students (*extremely* high-achieving/ability students) to be discriminated from other high achieving/ability students).

Note:
The forth level, 'very difficult' has been included in the levels of difficulty of examination questions to ensure that there are sufficient questions that discriminate well amongst higher ability candidates.

7.3 Question difficulty is determined against the assumed capabilities of the ideal 'envisaged' Grade 12 Life Sciences NSC examination candidate

The revised Umalusi NSC examination evaluation instruments require evaluators to exercise expert judgments about whether each examination question is 'Easy', 'Moderately challenging', 'Difficult' or 'Very difficult' (Table 6) for the '**envisaged**' Grade 12 learner to answer. In other words, assessment of question difficulty is linked to a particular target student within the population of NSC candidates, that is, the Grade 12 candidate of average intelligence or ability.

The Grade 12 learners that you may have taught over the course of your career cannot be used as a benchmark of the 'envisaged' candidate as we cannot know whether their abilities fall too high, or too low on the entire spectrum of all Grade 12 Life Sciences candidates in South Africa. The revised Umalusi NSC examination evaluation instruments thus emphasise that, when rating the level of difficulty of a particular question, your conception of the 'envisaged' candidate needs to be representative of the entire population of candidates for all schools in the country, in other words, of the overall Grade 12 population.

Most importantly, the conception of this 'envisaged' candidate is a learner who has been taught the whole curriculum adequately by a teacher who is qualified to teach the subject, in a functioning school. There are many disparities in the South African education system that can lead to very large differences in the implementation of the curriculum. Thus this 'envisaged' learner is not a typical South African Grade 12 learner – it is an intellectual construct (an imagined person) whom you need to imagine when judging the level of difficulty of a question. This ideal 'envisaged' Grade 12 learner is an aspirational ideal of where we would like all Life Sciences learners in South Africa to be.

Note:
The concept of the ***ideal envisaged Grade 12 candidate is*** that of an imaginary learner who has the following features:

a. Is of average intelligence or ability.
b. Has been taught by a competent teacher.
c. Has been exposed to the entire examinable curriculum.

This envisaged learner represents an imaginary person who occupies the middle ground of ability and approaches questions *having had all the necessary schooling.*

7.4 Question difficulty is determined using a common framework for thinking about question difficulty

Examiners, moderators and evaluators **in all subjects** are now provided with a common framework for thinking about question difficulty to use when identifying sources of difficulty or ease in each question, and to provide their reasons for the level of difficulty they select for each examination question.

The framework described in detail below provides the main sources of difficulty or 'ease' inherent in questions. The four sources of difficulty which must be considered when thinking about the level of difficulty of examination questions in this framework are as follows.

1. **'Content difficulty'** refers to the difficulty inherent in the subject matter and/or concept/s assessed.
2. **'Stimulus difficulty'** refers to the difficulty that candidates confront when they attempt to read and understand the question and its source material. The demands of the reading required to answer a question thus form an important element of 'stimulus difficulty'.
3. **'Task difficulty'** refers to the difficulty that candidates confront when they try to formulate or produce an answer. The level of cognitive demand of a question forms an element of 'Task difficulty', as does the demand of the written text or representations that learners are required to produce for their response.

4. **'Expected response difficulty'** refers to difficulty imposed by examiners in a marking guideline, scoring rubric or memorandum. For example, mark allocations affect the amount and level of answers students are expected to write.

This framework derived from Leong (2006) was chosen because it allows the person making judgments about question difficulty to grapple with nuances and with making connections. The underlying assumption is that judgment of question difficulty is influenced by the interaction and overlap of different aspects of the four main sources of difficulty. Whilst one of the above four sources of difficulty may be more pronounced in a specific question, the other three sources may also be evident. Furthermore, not all four sources of difficulty need to be present for a question to be rated as difficult.

The four-category conceptual framework is part of the required Umalusi examination evaluation instruments. Each category or source of difficulty in this framework is described and explained in detail below (Table 7). Please read the entire table very carefully.

TABLE 7: FRAMEWORK FOR THINKING ABOUT QUESTION DIFFICULTY

CONTENT/CONCEPT DIFFICULTY
Content/concept difficulty indexes the difficulty in the **subject matter, topic or conceptual knowledge** assessed or required. In this judgment of the item/question, difficulty exists in the **academic and conceptual demands** that questions make and/or the **grade level** boundaries of the various 'elements' of domain/subject knowledge (topics, facts, concepts, principles and procedures associated with the subject).
For example:
Questions that assess **'advanced content'**, that is, subject knowledge that is considered to be in advance of the grade level curriculum, are *likely* to be difficult or very difficult for most candidates. Questions that assess subject knowledge which forms part of the core curriculum for the grade are *likely* to be moderately difficult for most candidates. Questions that assess **'basic**

content' or subject knowledge candidates would have learnt at lower grade levels, and which would be familiar to them are *unlikely* to pose too much of a challenge to most candidates. Questions that require general everyday knowledge or knowledge of 'real life' experiences are *often* easier than those that test more **specialized** school **knowledge**. Questions involving only concrete objects, phenomena, or processes are *usually* easier than those that involve more **abstract constructs, ideas, processes or modes.** Questions which test learners' understanding of theoretical or **de-contextualised issues or topics**, rather than their knowledge of specific examples or contextualised topics or issues *tend* to be more difficult. Questions involving familiar, contemporary/current contexts or events are *usually* easier than those that are more **abstract or** involve **'imagined' events** (e.g. past/future events) or **contexts** that are **distant from learners' experiences.** Content difficulty may also be varied by changing **the number of knowledge elements or operations assessed**. *Generally,* the difficulty of a question increases with the number of knowledge elements or operations assessed. Questions that assess learners on two or more knowledge elements or operations are *usually* (but not always) more difficult than those that assess a single knowledge element or operation. Assessing learners on **a combination of knowledge elements or operations that are seldom combined** *usually* increases the level of difficulty.
EXAMPLES OF INVALID OR UNINTENDED SOURCE OF CONTENT DIFFICULTY
• Testing obscure or unimportant concepts or facts that are not mentioned in the curriculum, or which are unimportant to the curriculum learning objectives. • Testing very advanced concepts or operations that candidates are extremely unlikely to have had opportunities to learn.

STIMULUS DIFFICULTY
Stimulus difficulty refers to the difficulty of the linguistic **features of the question** (**linguistic** complexity) and the challenge that candidates face when they attempt to read, interpret and understand the words and phrases in the question AND when they attempt to read and understand the **information or 'text' or source material (diagrams, tables and graphs, pictures, cartoons, passages, etc.) that accompanies the question**.
For example:
Questions that contain words and phrases that require only simple and straightforward comprehension are *usually* easier than those that require the

candidate to understand **subject specific phraseology and terminology** (e.g. idiomatic or grammatical language not usually encountered in everyday language), or that require more technical comprehension and specialised command of words and language (e.g. everyday words involving different meanings within the context of the subject).

Questions that contain information that is 'tailored' to an expected response, that is, questions that contain no irrelevant or distracting information, are *generally* easier than those that require candidates to select relevant and appropriate information or **unpack a large amount of information** for their response. A question **set in a very rich context** *can* increase question difficulty. For example, learners *may* find it difficult to select the correct operation when, for example, a mathematics or accountancy question is set in a context-rich context.

Although the level of difficulty in examinations is *usually* revealed most clearly through the questions, text complexity or the degree of **challenge or complexity in written or graphic texts** (such as a graph, table, picture, cartoon, etc.) that learners are required to read and interpret in order to respond *can* increase the level of difficulty. Questions that depend on reading and selecting content from a text *can* be more challenging than questions that do not **depend on actually reading the accompanying text** because they test reading comprehension skills as well as subject knowledge. Questions that require candidates to **read a lot** *can* be more challenging than those that require limited reading. Questions that tell learners where in the text to look for relevant information are *usually* easier than those where **learners are not told where to look**.

The level of difficulty *may* increase if texts set, and reading passages or other **source material** used are challenging for the grade level, and make **high reading demands** on learners at the grade level. Predictors of textual difficulty include:

- **semantic content** – for example, if vocabulary and words used are typically outside the reading vocabulary of Grade 12 learners, 'texts' (passage, cartoon, diagram, table, etc.) are *usually* more difficult. 'Texts' are *generally* easier if words or images are made accessible by using semantic/context, syntactic/structural or graphophonic/visual cues.
- **syntactic or organisational structure** – for example, sentence structure and length. For example, if learners are likely to be *familiar with the structure* of the 'text' or resource, for example, from reading newspapers or magazines, etc. 'texts' are *usually* easier than when the structure is unfamiliar.
- **literary techniques** – for example, abstractness of ideas and imagery – and **background knowledge required**, for example, to make sense of allusions.
- if the **context** is **unfamiliar** or remote, or if candidates do not have or are

not provided with access to the context which informs a text (source material, passage, diagram, table, etc.) they are expected to read, and which informs the question they are supposed to answer and the answer they are expected to write, then constructing a response is *likely* to be more difficult than when the context is provided or familiar.

Questions which require learners to **cross-reference different sources** are *usually* more difficult than those which deal with one source at a time.

Another factor in stimulus difficulty is presentation and visual appearance. For example, type face and size, use of headings, and other types of textual organisers etc. can aid **'readability'** and make it easier for learners to interpret the meaning of a question.

EXAMPLES OF INVALID OR UNINTENDED SOURCES OF STIMULUS DIFFICULTY

- Meaning of words unclear or unknown.
- Difficult or impossible to work out what the question is asking.
- Questions which are ambiguous.
- Grammatical errors in the question that could cause misunderstanding.
- Inaccuracy or inconsistency of information or data given.
- Insufficient information provided.
- Unclear resource (badly drawn or printed diagram, inappropriate graph, unconventional table).
- Dense presentation (too many important points packed in a certain part of the stimulus).

TASK DIFFICULTY

Task difficulty refers to the **difficulty that candidates confront when they try to formulate or produce an answer**.

For example:

In most questions, to generate a response, candidates have to work through the steps of a solution. *Generally*, questions that **require more steps in a solution** are more difficult than those that require fewer steps. Questions involving only one or two steps in the solution are *generally* easier than those where several operations required for a solution**.**

Task difficulty may also be mediated by the **amount of guidance present in the question**. Although question format is not necessarily a factor and difficult questions can have a short or simple format, questions that provide guided steps or cues (e.g. a clear and detailed framework for answering) are *generally* easier than those that are more open ended and require candidates to form or tailor their **own response strategy** or argument, work out the steps **and maintain the strategy for answering** the question by themselves. A high degree of prompting (a high degree of prompted recall, for example) *tends* to reduce difficulty level.

Questions that test specific knowledge are *usually* less difficult that **multi-step,**

multiple-concept or operation questions.

A question that requires the candidate to **use** a **high level of** appropriate **subject specific, scientific or specialised terminology in** their **response** *tends* to be more difficult than one which does not.

A question requiring candidates to **create a complex abstract (symbolic or graphic) representation** is *usually* more challenging than a question requiring candidates to create a concrete representation.

A question requiring writing a one-word answer, a phrase, or a simple sentence is *often* easier to write than **responses that require more complex sentences, a paragraph or a full essay or composition**.

Narrative or descriptive writing, for example where the focus is on recounting or ordering a sequence of events chronologically, is *usually* easier than **writing discursively** (**argumentatively or analytically**) where ideas need to be developed and ordered logically. Some questions reflect task difficulty simply by **'creating the space' for A-grade candidates** to demonstrate genuine insight, original thought or good argumentation, and to write succinctly and coherently about their knowledge.

Another element is the **complexity in structure of the required response**. When simple connections between ideas or operations are expected in a response, the question is *generally* easier to answer than a question in which the significance of the relations between the parts and the whole is expected to be discussed in a response. In other words, a question in which an unstructured response is expected is *generally* easier than a question in which **a relational response** is required. A response which involves **combining or linking a number of complex ideas or operations** is *usually* more difficult than a response where there is no need to combine or link ideas or operations.

On the other hand, questions which require continuous prose or extended writing *may* also be easier to answer correctly or to get marks for than questions that require no writing at all or single letter answer (such as multiple choice), or a brief response of one or two words or short phrase/s because they **test very specific knowledge**.

The **cognitive demand** or **thinking processes** required form an aspect of task difficulty. Some questions test thinking ability, and learners' capacity to deal with ideas, etc. Questions that assess inferential comprehension or application of knowledge, or that require learners to take ideas from one context and use it in another, for example, *tend* to be more difficult than questions that assess recognition or retrieval of basic information. On the other hand, questions requiring recall of knowledge are *usually* more difficult than questions that require simple recognition processes.

When the **resources for answering** the question are included in the examination paper, then the task is *usually* easier than when candidates have to **use and**

select their own internal resources (for example, their own knowledge of the subject) or transform information to answer the question.

Questions that require learners to take or **transfer** ideas, **skills or knowledge from one context/subject area and use them in another** *tend* to be more difficult.

EXAMPLES OF INVALID OR UNINTENDED SOURCES OF TASK DIFFICULTY

- Level of detail required in an answer is unclear.
- Context is unrelated to or uncharacteristic of the task than candidates have to do.
- Details of a context distract candidates from recalling or using the right bits of their knowledge.
- Question is unanswerable.
- Illogical order or sequence of parts of the questions.
- Interference from a previous question.
- Insufficient space (or time) allocated for responding.
- Question predictability or task familiarity. If the same question regularly appears in examination papers or has been provided to schools as exemplars, learners are likely to have had prior exposure, and practised and rehearsed answers in class (for example, when the same language set works are prescribed each year).
- Questions which involve potential follow-on errors from answers to previous questions.

EXPECTED RESPONSE DIFFICULTY

Expected response difficulty refers to difficulty imposed by examiners in a **mark scheme and memorandum**. This location of difficulty is more applicable to 'constructed' response questions, as opposed to 'selected' response questions (such as multiple choice, matching/true-false).

For example:

When examiners expect few or no details in a response, the question is *generally* easier than one where the mark scheme implies that **a lot of details are expected**.

A further aspect of expected response difficulty is the clarity of the **allocation of marks**. Questions are *generally* easier when the allocation of marks is explicit, straight-forward or logical (i.e. 3 marks for listing 3 points) than when the **mark allocation is indeterminate or implicit** (e.g. when candidates need all 3 points for one full mark or 20 marks for a discussion of a concept, without any indication of how much and what to write in a response). This aspect affects difficulty because candidates who are unclear about the mark expectations in a response may not produce sufficient amount of answers in their response that will earn the marks that befit their ability.

Some questions are more difficult/easy to mark accurately than others. Questions that are **harder to mark and score objectively** are *generally* more difficult for candidates than questions that require simple marking or scoring strategies on the part of markers. For example, recognition and recall questions are *usually* easier to test and mark objectively because they usually require the use of matching and/or simple scanning strategies on the part of markers. More complex questions requiring analysis (breaking down a passage or material into its component parts), evaluation (making judgments, for example, about the worth of material or text, or about solutions to a problem), synthesis (bringing together parts or elements to form a whole), and creativity (presenting own ideas or original thoughts) are *generally* harder to mark/score objectively. The best way to test for analysis, evaluation, synthesis and creativity is usually through extended writing. Such extended writing *generally* requires the use of more cognitively demanding *marking* strategies such as interpreting and evaluating the logic of what the candidate has written.

Questions where **a wide range of alternative answers or response/s** is possible or where the correct answer may be arrived at through different strategies *tend* to be more difficult. On the other hand, questions may be so open-ended that learners will get marks even if they engage with the task very superficially.

EXAMPLES OF INVALID OR UNINTENDED SOURCES OF EXPECTED RESPONSE DIFFICULTY

- Mark allocation is unclear or illogical. The weighting of marks is important in questions that comprise more than one component when components vary in levels of difficulty. Learners may be able to get the same marks for answering easy component/s of the item as other learners are awarded for answering the more difficult components.
- Mark scheme and questions are incongruent. For example, there is no clear correlation between the mark indicated on the question paper and the mark allocation of the memorandum.
- Question asked is not the one that examiners want candidates to answer. Memorandum spells out expectation to a slightly different question, not the actual question.
- Impossible for candidate to work out from the question what the answer to the question is (answer is indeterminable).
- Wrong answer provided in memorandum.
- Alternative correct answers from those provided or spelt out in the memorandum are also plausible.
- The question is ‘open’ but the memo has a closed response. Memo allows no leeway for markers to interpret answers and give credit where due.

The framework described above does not provide you with explicit links between the different sources of difficulty, or show relationships and overlaps between the different categories and concepts in the framework. This is because it is impossible to set prescribed rules or pre-determined combinations of categories and concepts used for making judgments about the source of difficulty in a particular examination question.

The intention behind the framework is to allow you to exercise your sense of judgment as an expert. The complexity of your judgment lies in your ability as an expert to recognise subtle interactions and identify links between different categories of a question's difficulty or ease. For example, a question that tests specific knowledge of your subject can actually be more difficult than a multi-step question because it requires candidates to explain a highly abstract concept, or very complex content. In other words, although questions that test specific knowledge are *usually* less difficult than multiple-concept or operation questions, the level of difficulty of the content knowledge required to answer a question can make the question more difficult than a multi-step or multi-operation question.

Not all one-word response questions can automatically be assumed to be easy. For example, multiple-choice questions are not automatically easy because a choice of responses is provided – some can be difficult. As an expert in your subject, you need to make these types of judgments about each question.

Note:
It is very important that you become extremely familiar with the framework explained in Table 7, and with each category or source of difficulty provided (i.e. content difficulty, task difficulty, stimulus difficulty, and expected response difficulty). You need to understand the examples of questions which illustrate each of the four levels (Table 8 to Table 11). This framework is intended to assist you in discussing and justifying your decisions regarding the difficulty level ratings of questions. You are expected to **refer to all four categories or sources of difficulty** in justifying your decisions.

When considering question difficulty ask:

- How difficult is the **knowledge** (content, concepts or procedures) that is being assessed for the envisaged Grade 12 candidate? (*Content difficulty*)
- How difficult is it for the envisaged Grade 12 candidate to formulate the answer to the question? In considering this source of difficulty, you should **take into account the type of cognitive demand** made by the task. (*Task difficulty*)
- How difficult is it for the envisaged Grade 12 candidate to **understand the question and the source material** that need to be read to answer the particular question? (*Stimulus difficulty*)
- What does the **marking memorandum and mark scheme** show about the difficulty of the question? (*Expected response difficulty*)

7.5 Question difficulty entails distinguishing unintended sources of difficulty or ease from intended sources of difficulty or ease

Close inspection of the framework for thinking about question difficulty (Section 7.4, Table 7) above, shows that, for each general category or source of difficulty, the framework makes a distinction between 'valid' or intended, and 'invalid' or unintended sources of question difficulty or ease. Therefore, defining question difficulty entails identifying whether sources of difficulty or ease in a question were intended or unintended by examiners. Included in Table 7 are examples of unintended sources of difficulty or ease for each of the four categories.

Valid difficulty or 'easiness' in a question has its source in the requirements of the question, and is ***intended*** by the examiner (Ahmed and Pollit, 1999). Invalid sources of difficulty or 'easiness' refer to those features of question difficulty or 'easiness' that were ***not intended*** by the examiner. Such unintended 'mistakes' or omissions in questions can prevent the question from assessing what the examiner intended, and are likely to prevent candidates from demonstrating their true ability or competence, and can result in a question being easier or more difficult than the examiner intended.

For example, grammatical errors in a question that could cause misunderstanding for candidates are unintended sources of question difficulty because the difficulty in answering the question could lie in the faulty formulation of the question, rather than in the intrinsic difficulty of the question itself (for example, because of stimulus difficulty). Candidates "may misunderstand the question and therefore not be able to demonstrate what they know" (Ahmed and Pollit, 1999, p.2). Another example is question predictability (when the same questions regularly appear in examination papers or textbooks) because familiarity can make a question which was intended to be difficult, less challenging for examination candidates.

Detecting unintended sources of difficulty or ease in examinations is largely the task of moderators. Nevertheless, evaluators also need to be vigilant about detecting sources which could influence or alter the intended level of question difficulty that moderators may have overlooked.

Note:
When judging question difficulty, you should distinguish **unintended sources of question difficulty or ease from** those sources that are **intended**, thus ensuring that examinations have a range of levels of difficulty. The framework for thinking about question difficulty allows you to systematically identify technical and other problems in each question. Examples of problems might be: unclear instructions, poor phrasing of questions, the provision of inaccurate and insufficient information, unclear or confusing visual sources or illustrations, incorrect use of terminology, inaccurate or inadequate answers in the marking memorandum, and question predictability. You should **not** rate a question as difficult/easy if the source of difficulty/ease lies in the 'faultiness' of the question or memorandum. Instead, as moderators and evaluators, you need to alert examiners to unintended sources of difficulty/ease so that they can improve questions and remedy errors or sources of confusion before candidates write the examination.

7.6 Question difficulty entails identifying differences in levels of difficulty within a single question

An examination question can incorporate more than one level of difficulty if it has subsections. It is important that the components of such questions are 'broken down' into to their individual levels of difficulty.

Note:
Each subsection of a question should be analysed separately so that the percentage of marks allocated at each level of difficulty and the weighting for each level of difficulty can be ascertained as accurately as possible for that question.

8. EXAMPLES OF QUESTIONS AT DIFFERENT LEVELS OF DIFFICULTY

This section provides **three** examples of questions from previous Life Sciences NSC examinations (Table 8 to Table 11) categorised at each of the four levels of difficulty described in Section 7 (Table 7) above. These examples were selected to represent the **best and clearest** examples of each level of difficulty that the Life Sciences experts could find. The discussion below each example question tries to explain the reasoning behind the judgments made about the categorisation of the question at that particular level of difficulty.

TABLE 8: EXAMPLES OF QUESTIONS AT DIFFICULTY LEVEL 1 – EASY

Example 1: <u>Question 3.2.1, 2014 DBE Paper 1</u> (DBE, 2014a) What is meant by the following terms: (a) Carbon footprint (2) (b) Food security (2) **<u>Discussion</u>**: The question is short, clear and concisely written **(*stimulus*)**. In each subsection of the question one biological term is used (carbon footprint or food security), and each of these terms should be familiar to Grade 12 students **(*content*)**. Students just need to recognise each term and recall the definition of the term **(*task*).** The answer can be written in a single sentence; one mark is awarded for each of two factual points in (a), and one mark is awarded for any two of three factual points (b) **(*expected response*)**. This question is thus easy in regard to all four sources of difficulty in the framework.
<u>Memorandum/Marking guidelines</u> 3.2.1 (a) - Carbon footprint is a measure of the total amount of greenhouse gas

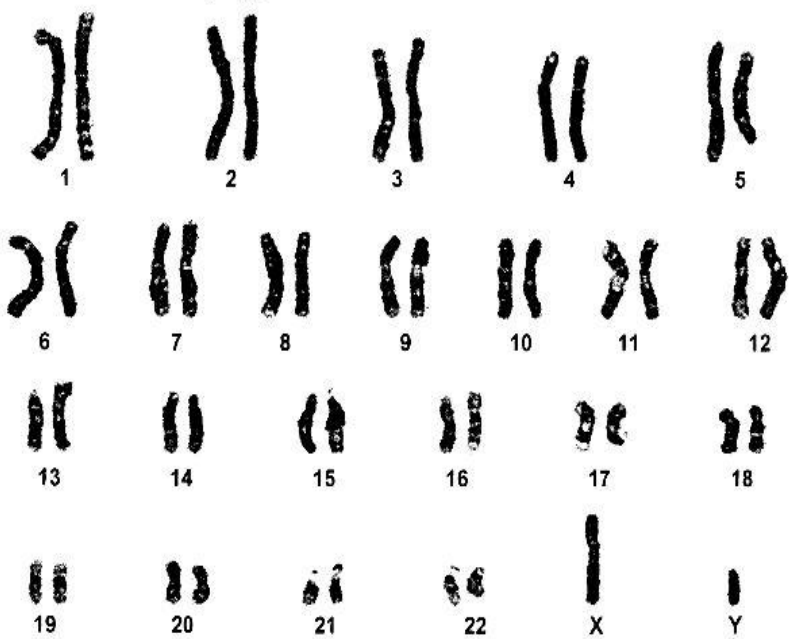
emissions √/ (example of greenhouse gas).
- of an individual √ / defined population/ company per year (2)

(b) - Food security refers to the availability and access. √
- to adequate, safe and nutritious food√ to people at all times. √
Any (2)

Example 2

Question 2.1.1, 2014 IEB Paper 1 (IEB, 2014)

Karyotype of an individual with Cri-du-chat



[Adapted: <http://teacherweb.com>]

2.1.1 Is this karyotype from a male or a female individual? Give a reason for your answer. (2)

Discussion:

The question with two parts (identifying the sex, and reason for the identification) is short and easy to understand. It is written clearly, the gonosomes are labelled, and the diagram is clear **(*stimulus*)** and contains one technical term {karyotype) which should be familiar to Grade 12 students **(*content*).** Students just need to recognise that the two gonosomes in the diagram are from a male, and recall why they are recognised as such **(*task*)**. The first part of the answers required one word for one mark (candidates have a 50% chance of guessing correctly); the second part requires a short statement, also for one mark **(*expected response*)**.

This question is thus easy in regard to all four sources of difficulty in the framework.

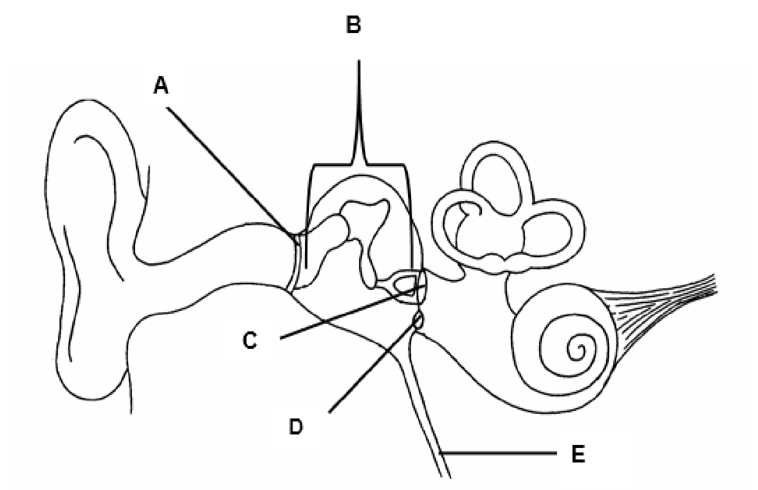
Memorandum/Marking guidelines

2.1.1 male ✓ presence of Y chromosome ✓ (2)

Example 3:

Question 2.1.1, 2011, DBE Version 1 Paper 2 (DBE, 2011b)

2.1 Study the diagram below showing a portion of the human ear and answer the questions that follow.



Portion of the ear

2.1.1 Provides labels for parts A, C and D respectively. (3)

Discussion*:*

The diagram provided in this question is a standard diagram that recurs in examination papers, and should therefore be familiar to Grade 12 candidates **(*content*)**. The question is written in simple language, with no scientific terms; the diagram is clear, and label lines indicate clearly which structures are required **(*stimulus*)**. Candidates are required to recognise each of the structures and recall the name of each **(*task*)**. Single words are required to answer the question for one mark each **(*expected response*)**. This question is thus easy in regard to all four sources of difficulty in the framework.

<u>Memorandum/Marking guidelines</u>

2.1.1

A = tympanic membrane √ /eardrum
C = oval window √
D = round window √ (3)

TABLE 9: EXAMPLES OF QUESTIONS AT DIFFICULTY LEVEL 2 – MODERATELY DIFFICULT

Example 1:

<u>Question 2.3.1, 2012 DBE Version 1 Paper 1</u> (DBE, 2012c)

2.3 A geneticist wanted to find out which corn colour is dominant in a species of maize. The species has two phenotypes for colour, yellow and white. She performed four genetic crosses and recorded the colour of the offspring in the table below.

Genetic crosses	Parent phenotypes	Offspring phenotypes
1	Yellow x yellow	All yellow
2	White x white	51 white and 17 yellow
3	White x yellow	32 white and 34 yellow
4	White x white	All white

2.3.1 According to the results, which colour is dominant? (1)

<u>Discussion</u>:

The information provided in the source material is fairly complex, with two 'white x white' crosses and includes a significant number of biological terms, which make the source material moderately difficult **(*stimulus*)**. However, the question itself is phrased simply, and contains terms that should be familiar to Grade 12 students **(*content*)**. Candidates need to apply their scientific knowledge to a new context, corn colour **(*task*)**. The answer is either white or yellow, so the level of difficulty of the ***expected response*** is easy.

This question is thus considered to be moderately challenging because of the levels of the stimulus difficulty and the task difficulty.

Memorandum/Marking guidelines

2.3.1 White √ (1)

Example 2:

Question 1.3, 2011 DBE Version 1Paper 2 (DBE, 2011b)

1.3 Indicate whether each of the statements in COLUMN I applies to **A ONLY, B ONLY, BOTH A AND B**, or **NONE** of the items in COLUMN II. Write **A only, B only, both A and B**, or **none** next to each question number (1.3.1 – 1.3.6) in the ANSWER BOOK.

COLUMN I	COLUMN II
1.3.1 Development in which a hatchling is incapable of moving around on its own and needs to be fed by its parents.	A: Altricial B: Precocial
1.3.2 Method of reproduction in which the foetus is retained in the mother's uterus and is nourished through an umbilical cord.	A: Ovipary B: Ovovivipary
1.3.3 A dominant and haploid sporophyte generation alternates with a dependant gametophyte generation	A: Moss B: Angiosperm

(3 x 2)

Discussion:

The instructions for this question appear to be complex and difficult to understand. However, this type of question has appeared in almost every NSC Life Sciences examination, and exemplar examination (and DBE and provincial Grades 10 and 11 exemplar examinations) published since 2008. Therefore, the instructions should be familiar and not as difficult as might first appear.

The biological terms and statements in the question contain knowledge that should be familiar to Grade 12 **(*content*)**. The statements in Column I of the table are terminologically dense, and are therefore not very easy to read **(*stimulus*)**. Candidates are required to recognise, recall and match the terms and statements where appropriate. However, this is not a straight forward matching exercise because the answer needs to be given in combinations of A and B **(*task*)**. To answer, a candidate has to only choose from the four possible answers: A only, B only, both A and B, or none **(*expected response*)**.

Despite the easy content and short answers this question is thus considered to be moderately challenging because of the levels of the stimulus and task difficulty.

Memorandum/Marking guidelines

1.3.1 A only √√
1.3.2 None √√

1.3.3 None √√	(3 x2)

Example 3:

Question 4.1.2, 2012 DBE Version 1 Paper 1 (DBE, 2012c)

4.1 Study the information about the discovery of a new species of Australopithecine in South Africa.

DISCOVERY OF A NEW SPECIES IN SOUTH AFRICA

Professor Lee Berger a paleoanthropologist at the University of the Witwatersrand, with the help of his 9-year-old son, found two fossils in South Africa on 15 August 2008.

Berger and about 60 of his colleagues from all over the world, studied the fossilized bone fragments before they announced their findings to the public on 8 April 2010. Their findings were published in a scientific journal.

They presented it as a new species called *Australopithecus sediba*, dated to approximately 1,78 to 1,95 million years ago. It consists of many primitive features characteristic of other australopithecine species and more advanced features typical of later *Homo* species.

The almost 2-million-year-old partial skeletons are thought to possibly be that of the transitional species between *Australopithecus africanus* (such as the famous Mrs Ples) and either *Homo habilis* or *Homo erectus*, the early ancestors of humans.

Berger said that the brain, hand and foot have characteristics of both modern and early pre-human forms. It represents a model that could have led to the human genus Homo.

It was noted that the brain of *A. sediba* is small, like that of a chimpanzee, but with a re-organisation more human-like, particularly with an expansion behind and above the eyes.

[Adapted from http://www.sciencemag.org]

4.1.1 Explain why scientists took a long time to present their findings to the public.

(2)

4.1.2 Explain why it was important to publish their findings in a scientific journal.

(2)

Discussion:

The following discussion refers only to Question 4.1.2 Here, the source material is relatively complex with many scientific terms. The genre is more scientific than Grade 12 candidates may be accustomed to ***(stimulus)***, Although the answer could draw from learned knowledge, candidates have to read the text because the question refers to "their findings" in the source material – therefore candidates are directed to read the text. However, the question itself is phrased in simple language and can be answered based on an understanding of science, particularly of the scientific method rather than the actual text **(*content, task*)**. The response required is a short answer, requiring writing only one or two sentences **(*expected response*)**.

This question is therefore moderately challenging due to stimulus difficulty.

Memorandum/Marking guidelines

4.1.2 To inform people of their findings√ so that they can critique √ /verify their findings /use it for future research/acknowledge ownership of the findings (2)

TABLE 10: EXAMPLES OF QUESTIONS AT DIFFICULTY LEVEL 3 – DIFFICULT

Example 1:

Question 2.1.2, 2016 DBE Paper 2 (DBE 2016b)

2.1 The leaf colour in a plant is controlled by two alleles, green (**G**) and yellow (**g**). Thorns on plant stems are controlled by two alleles, presence of thorns (**T**) and no thorns (**t**).

Two plants with the genotypes **GGTT** and **ggtt** were crossed. Their offspring were then left to pollinate each other.

The table below shows the possible genotypes of the offspring of the second generation. Genotypes (**i**) and (**ii**) have been left out.

Gametes	**GT**	**Gt**	**gT**	**gt**
GT	GGTT	GGTt	GgTT	GgTt
Gt	GGTt	GGtt	**(i)**	Ggtt
gT	GgTT	GgTt	ggTT	ggTt
gt	GgTt	Ggtt	ggTt	**(ii)**

2.1.2 List the FOUR genotypes of the offspring of the second generation that would be phenotypically different from the original pair of parents. (4)

Discussion:

The ***stimulus*** material is relatively technical, and the question while clearly stated, is complex. The abstract nature of genetics makes it difficult for students to understand, especially dihybrid crosses **(*content*)**. Distinguishing between phenotypes and genotypes for two characteristics simultaneously requires higher order thinking and is not easy ***(task)***. In addition, the use of specific scientific conventions and the combination of genotypes required by the answer is difficult **(*expected response*)**. This question is considered difficult based on characteristics of the content, task and expected response.

This question is not considered very difficult because using a Punnett square diagram to predict an outcome of a particular cross or breeding experiment is relatively routine once the procedure has been practised and mastered.

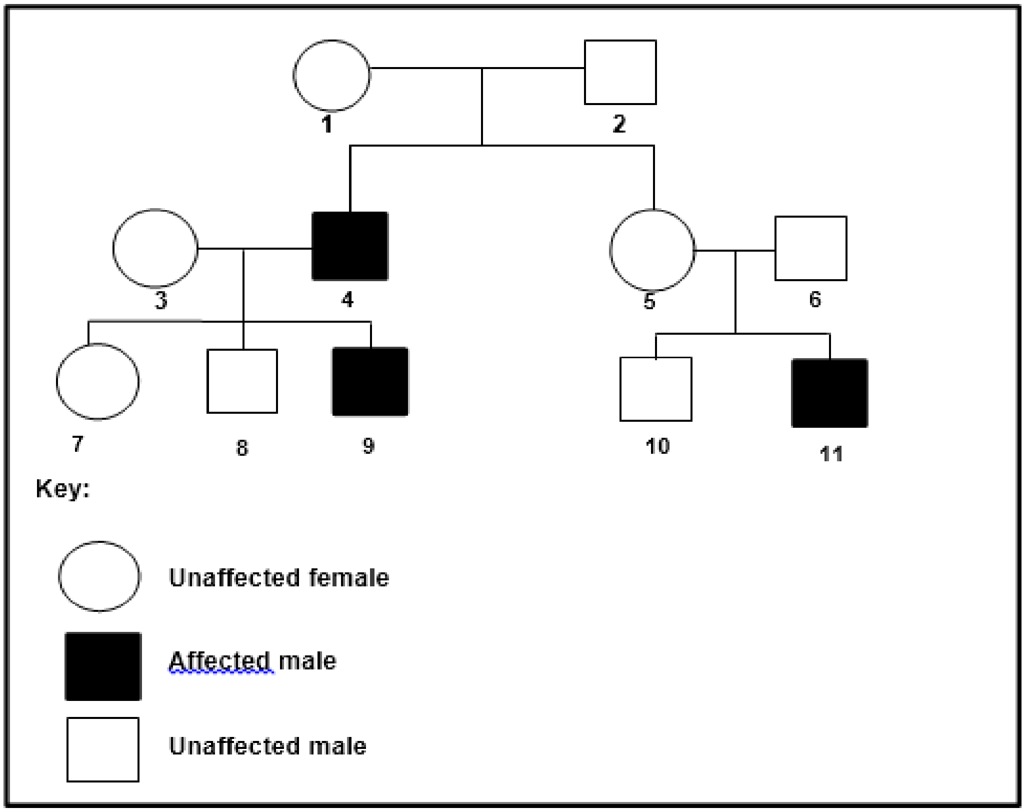
Memorandum/Marking guidelines

2.1.2 GGtt √
Ggtt √
ggTT √
ggTt √ (4)

Example 2:

Question 3.1.4, 2012 DBE Version 1 Paper 1 (DBE, 2012c)

3.1 The pedigree diagram below traces the inheritance of haemophilia in a family.



3.1.4 What are the chances of individual 10 and his wife, who is a carrier (not shown in the pedigree), having a son who is affected? (2)

Discussion:

The question uses only one scientific term ("carrier"), which should be familiar to students **(*content*)**. The wording of the question is quite complex, but the diagram is clear **(stimulus)**. To arrive at the one number (a percentage or a fraction) that the ***expected response*** requires, candidates to have an understanding of an abstract concept, that is, chance or probability – this concept is not easy for many students (most of whom study Mathematics Literacy, not Mathematics) (*task content*). The main difficulty of this item lies in the task and content. The fact that one individual (the wife of individual 10) is intentionally not shown on the pedigree diagram adds to the difficulty **(*stimulus*)**.

Memorandum/Marking guidelines

3.1.4 25% √√ / ¼ OR 50% √√/ ½ (2)

Example 3:

Question 3.2.5 2012 IEB Paper 1 (IEB, 2012)

3.2 DNA fingerprinting may be used in forensic science to solve crimes. Read the information below and answer the following questions.

> The first step to making a genetic fingerprint requires getting a sample of DNA. This sample can come from blood, semen, hair or saliva, and may be an extremely small sample. A blood sample contains white blood cells which are broken open using detergent, and the DNA is extracted.
>
> Next, restriction enzymes are used to cut the non-coding DNA into many pieces of different lengths. This non-coding DNA consists of short sequences of bases that are repeated. The number of times the non-coding segments are repeated varies from individual to individual.

[Adapted from: Taylor, Green, Stout: 1984]

3.2.5 Explain how a PCR (Polymerase Chain Reaction) would produce a larger amount of DNA if a forensic sample was too small for DNA analysis. (3)

Discussion:

The source text is fairly long and includes a number of scientific terms, 'PCR' is referred to in the question but not in the text above the question. While it might be possible to answer the question without reading the text candidates do not know this **(*stimulus*)**. The ***content*** on which this question is based is abstract and relatively technical for Grade 12 students which makes it more difficult for students to connect what they know about DNA analyses, what is in the text and PCR **(*task*)**. A short, concise and succinct answer is not easy, especially for Grade 12 students **(*expected response*)**. This question is judged to be difficult on the basis of the task and expected response sources of difficulty.

Memorandum/Marking guidelines

3.2.5 Small amount of DNA segments reproduced/increased (amplified) by adding sections of nucleotides to each DNA strand thereby making more of the specific DNA selected/required/needed for forensic analyses; catalysed by enzyme/DNA polymerase. 3 facts (3)

TABLE 11: EXAMPLES OF QUESTIONS AT DIFFICULTY LEVEL 4 – VERY DIFFICULT

Note:
During the development of the exemplar book some subject specialists argued that there is a fine line between a difficult and a very difficult question. It was also evident that in some subjects question papers did not have questions that could be categorised as very difficult. In order to cater for this category, subject specialists were requested to adapt existing questions and make them very difficult or create their own examples of very difficult question. However, it was noted that in some instances attempts to create very difficult questions introduced invalid sources of difficulty, which in turn rendered the questions invalid. Hence Umalusi acknowledges that the very difficult category may be problematic and therefore requires especially careful scrutiny.

Example 1:

Question 19, 2016 IEB Paper 3 (IEB, 2016)

19. **EXPERIMENTAL DESIGN**

RENNIN IS AN ENZYME FOUND IN MAMMALS

When babies are born and their diet consists of milk alone, the body produces rennin from the stomach wall to cause the milk to be changed from its liquid form into a semi-solid. This allows the milk to stay in the stomach for longer and slows down the digestion process. By doing this the baby can be satisfied longer on its very limited diet of milk in the early months of life.

When milk turns sour in the container if it is left out of the refrigerator for too long, it goes lumpy and we see this as we pour it into a glass. The lumps that are observed in souring milk are very similar to the effect we see when adding rennin to milk.

Rennin, in the form of 'Rennet', is used in the cheese-making process to coagulate (solidify) milk and make cheese. Commercial rennin or 'Rennet' can be bought in liquid form from cheese making suppliers and used in the laboratory to show its effect on coagulating milk. Rennet works best if it is at a temperature in the 60 °C to 70 °C range and at a slightly acidic pH. Adding a drop of lemon juice to a solution can lower pH.

Using equipment that you would normally find in a school laboratory, Rennet

obtained from chemical suppliers and any other materials you may need, design a simple experiment to show that when added to milk, Rennet will cause the milk to coagulate and form lumps. *In your investigation you will need to have both an experiment and a control.*

(Do not actually perform your experiment)

19.1 Formulate a hypothesis for this experiment that you are designing. (3)

19.2 State the aim of the experiment. (2)

19.3 Outline your own method using numbered points. (8)

Discussion*:*

Questions 19.1 and 19.2 provide the scaffolding to answer Question 19.3, Question 19.3 is the focus of the argument which follows.

The text given in the question is relatively complex and unfamiliar to Grade 12 candidates **(*stimulus*)**. Therefore, the student has no specific biological knowledge to recall and is forced to engage with the question by drawing on his/her knowledge of experimental design and procedures in the context of the scientific data given in question ***(content)***. Grade 12 students do not find it easy to design their own investigations, especially with respect to distinguishing between what constitutes an experiment and a control, and the importance of both **(*task*)**. The ***expected response*** for Question 19.3 requires students to specifically link their experimental design to their specific answers to Questions 19.1 and 19.2, and to describe an appropriate methodology without being able to test it practically – not easy for Grade 12 students. This question is judged to be at very difficult, because the stimulus, task, content and expected response are very difficult.

Memorandum/Marking guidelines

19.1 Rennet added to milk will cause the milk to coagulate. Statement. (3)

THE DESIGN MUST SPEAK TO THE HYPOTHESIS

19.2 To investigate whether milk in the presence of Rennet will coagulate and form lumps.

OR

To determine if adding Rennet will cause the milk to coagulate and form lumps. (2)

19.3 (1) Take two identical test tubes.
(2) Mark one test tube A and the other B.
(3) Using a 10ml syringe add 30 ml of milk to both A and B.
(4) Add 4 drops of lemon juice to both test tubes A and B.
(5) Into A add 4 drops of Rennet.
(One test tube has Rennet and the other test tube no Rennet.) THIS

IS A CRUCIAL POINT THAT NEEDS TO BE IN THE DESIGN.

(6) Take a 250ml beaker and to it add 200 ml water that is at a temperature of 70 C (water bath).

(7) Place both test tubes in the water bath for 10 minutes/set time.

(8) Remove the test tubes from the water bath and tilt to the side to observe for small lumps/coagulation.

(9) Record your observations in a table.

(10) Repeat the investigation to verify your findings.

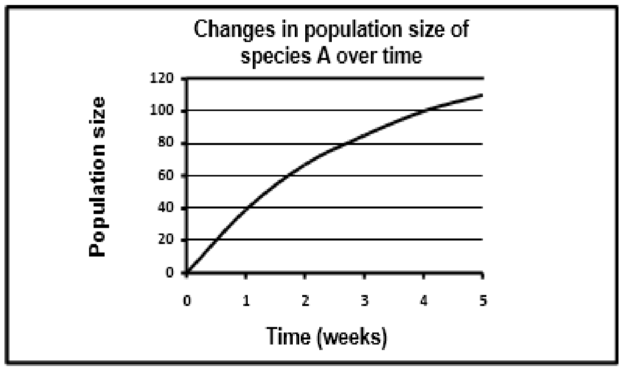
NB: The Rubric for this question is attached as Annexure A

Example 2:

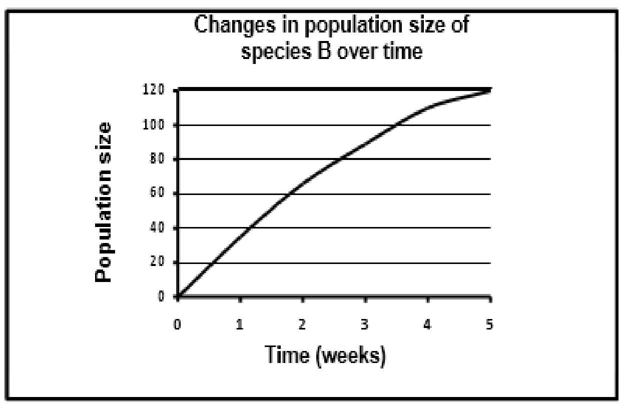
<u>Question 4.2.2, 2012 DBE Version 1 Paper 2</u> (DBE, 2012a)

4.2 The growth patterns of two closely related species (**A** and **B**) that rely on the same food source were investigated. At first the two species were separated and then the two species were kept in the same habitat for the same period of time. In all cases the organisms were provided with a limited food supply. The results are shown in the three graphs (**A**, **B** and **C**).

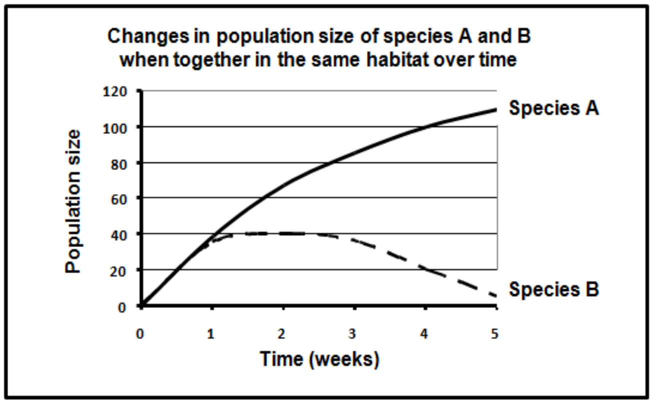
GRAPH A



GRAPH B



GRAPH C



4.2.2 Use GRAPHS A, B and C to explain the growth patterns of species A and species B when separated, compared to the growth patterns of species A and B together in the same habitat. (6)

Discussion:

The description of the investigation in the source material for this question is dense, complex and unfamiliar (the kinds of organisms are not identified) to Grade 12 candidates **(*stimulus*)**. The question itself is complex and it is difficult for students to work out what they have to do to answer the question **(*stimulus*)**. Insufficient information is deliberately not provided, e.g., the kinds of organisms are not identified. This means that the student has no specific biological knowledge to recall and is forced to confine their answer to only what is given in the question **(*content*)**. The *task* involves a comparison among three graphs which is also challenging for candidates. The ***expected response*** requires three short, concise and succinct sentences (one per comparison) – not easy for Grade 12 students.

This question is judged to be at very difficult, because the stimulus, task and expected response are difficult.

Memorandum/Marking guidelines

4.2.2 When grown alone population size of both species A and B increased greatly √*/ to over 100 because there is no competition √*
When the two species are grown together in the same habitat Population size of species A and B increases √ in the first week √/to about 40 since there was sufficient food √ for the low population size of both species. √
As food supply decreased, competition √ increased.
Thereafter, species A outcompeted species BΠ Resulting in species A increasingΠ/up to 110 while species B stabilized √/remained at 40 and then declined √

*2 compulsory marks + any 4 others (6)

Example 3:

<u>Question 5, 2014 IEB Paper 1 (IEB, 2014)</u>

Collaboration (working together) amongst scientists has been a feature of many important scientific findings.

Is there any benefit in scientists working together to find out more about the evolution of modern humans *(Homo sapiens)***?**

Read the source material carefully and present a debated argument to illustrate your point of view.

To answer this question, you are expected to:

- Select relevant information from Sources A to H below. Do not attempt to use all the detail provided.
- Integrate your own relevant biological knowledge. However, do not write an essay based solely on your own knowledge.
- Take a definite stand on the question and arrange the information to best develop your argument.
- Write in a way that is scientifically appropriate and communicates your point of view clearly.

Write an essay of not more than 1 ½ to 2 pages to answer the question. (20)

SOURCE A **Is *Australopithecus sediba* the Most Important Human Ancestor Discovery Ever?**

By Kate Wong April 24, 2013

RECONSTRUCTION OF *AUSTRALOPITHECUS SEDIBA*, BASED ON REMAINS FROM THREE INDIVIDUALS FOUND AT THE SITE OF MALAPA IN SOUTH AFRICA

[Image: Courtesy of Lee R. Berger and the University of the Witwatersrand]

Three years ago researchers added a new branch to the human family tree: *Australopithecus sediba*, a nearly two-million-year-old relative from South Africa. By all accounts it was a dazzling find – two partial skeletons, an adult female and young male. Any time human fossils, especially skeletons, are found it's a big deal, because such remains are so incredibly rare. *A. sediba* may just be the most important hominin (modern humans and their extinct relatives) discovered yet.

The way Berger and his collaborators are studying the finds and sharing what they learn represents a real difference to the cloak-and-dagger (secretive) manner in which palaeoanthropological investigations often work. Berger has made the project open access, with a policy of granting permission to any palaeoanthropologist who asks to see the original fossils. This can only improve the quality of the science that comes out of the project and may well inspire other teams to be more forthcoming with their own data.

At the end of the day we need loads of fossils, artifacts and DNA from different times and places to piece together the full story of our origins.

[Adapted: **@Scientific American,** a blog where several senior editors and managers will provide you with up-to-date updates on everything that is new at *Scientific American*]

SOURCE B **COLLABORATION IN SCIENCE**

PRESS RELEASE

2012-10-08

THE NOBEL ASSEMBLY AT KAROLINSKA INSTITUTE

has today decided to award

THE NOBEL PRIZE IN PHYSIOLOGY OR MEDICINE 2012

jointly to

John B. Gurdon and Shinya Yamanaka

for the discovery that mature stem cells can be reprogrammed to become pluripotent

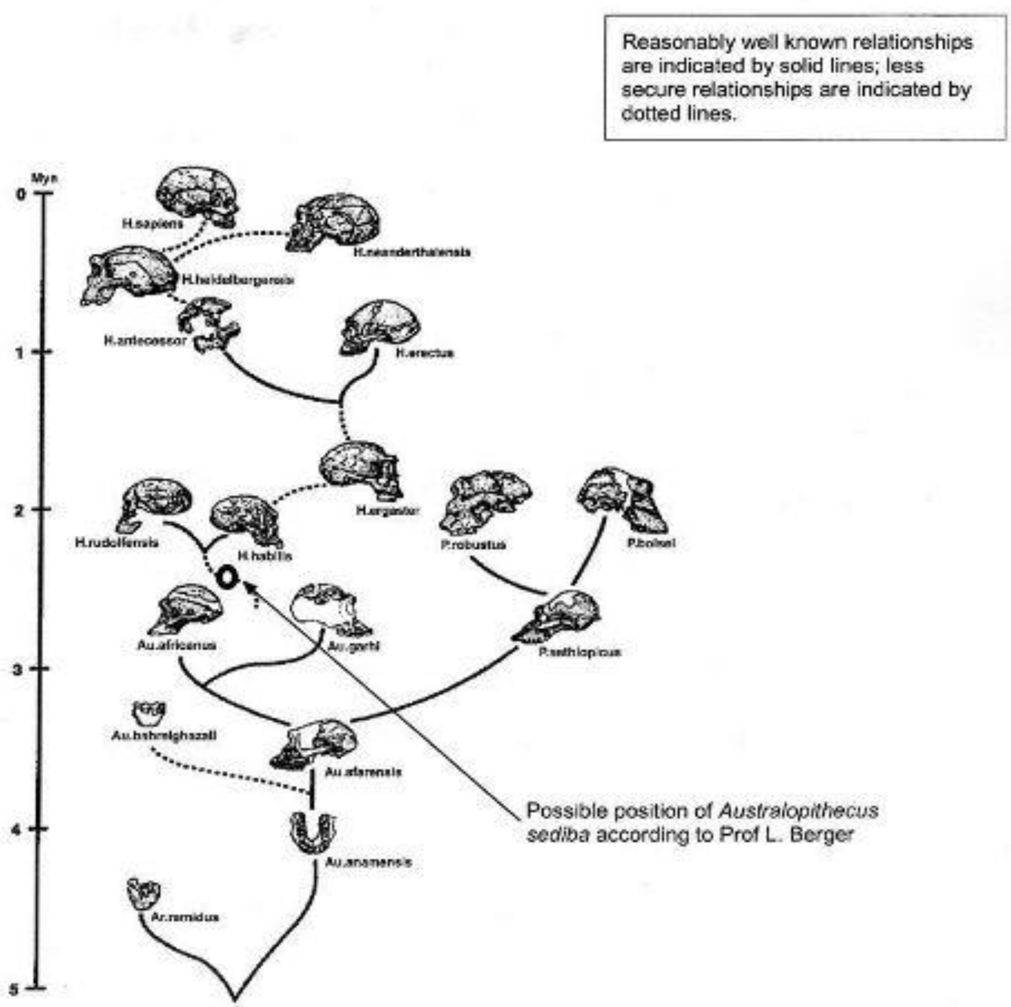
* **Sir John B. Gurdon** – Gurdon is currently at the Gurdon Institute at Cambridge University
* **Shinya Yamanaka** – Yamanaka is currently Professor at Kyoto University, where he directs the Centre for Stem Cell Research and Application

SOURCE C RELIABILITY OF EARLY FOSSIL FINDS

How reliable are palaeoanthropological finds and their interpretations? There are limitations – discoveries are fairly rare and have often been made under questionable circumstances, especially in the early days. As soon as something is dug up and taken elsewhere, essential elements are destroyed and one is dependent on the testimony of the discoverers. Sometimes early field research methods were extremely unscientific, but the resulting finds were taken seriously. When a certain discovery fits the currently ruling theory it will be accepted without much scrutiny. If something does not fit the pattern, it is either ignored or attacked and rejected, but not always on valid grounds.

[Adapted from: <www.theosophy-nw.org>]

SOURCE D HOMINID FAMILY TREE SHOWING A POSSIBLE LINE OF DESCENT



[<http://www.theosophy-nw.org 2013>

SOURCE E THE NATURE OF SCIENCE

SCIENCE IN THE REAL WORLD

Science is logical and objective – mostly. But let us get real ...

Hypothesis-test-revise is an idealised view of science. Like learning to drive, we all know what we should do, but after our test do we all drive perfectly?

Scientists are human too. The majority are honest, conscientious, hard-working and fair. Some, a tiny minority, cheat: they make things up, or fudge figures.

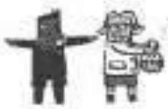
But scientists also tend to be attached to their own theories.

Science has many safety checks:

Peer review

Scrutiny of research by other experts before findings are published.

Whistle blowing



If researchers suspect foul play, they can raise their concerns with authorities.

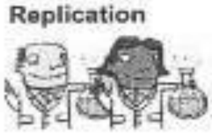
Regulation

There are many forms of regulation that scientists must adhere to – from their employers (e.g. universities), ethical committees and national laws.

Objectivity

Experiments should be objective – another scientist should be able to carry out the same research in the same way and get the same findings.

Replication



It is usually not enough for one group of scientists to report findings. These are often treated with caution until a second group has found the same thing.

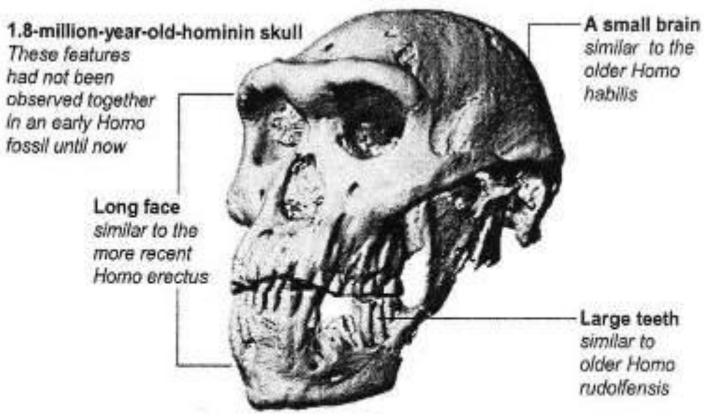
Testability

A good scientific theory is testable. If it cannot be tested, it is of little use in science.

[Adapted: <www.wellcomescience.ac.uk>]

SOURCE F **NEWLY FOUND SKULL CAUSES DEBATE ON HUMAN HISTORY**

A 1.8-million-year-old skull blends features of a number of early human species

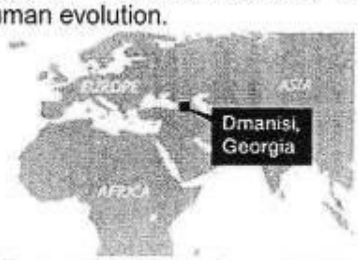


The blend of features suggests early humans were one species that had diverse facial and cranial characteristics.

This skull's features resemble those of both earlier and later humans. Photograph courtesy Georgian National Museum

Hominid bones located in Georgia could rule out *Australopithecus sediba* as a human ancestor. The five hominid fossils could remove a number of African hominid species from the tree of human evolution and rule out *Australopithecus sediba* as a possible human ancestor.

The Dmanisi skulls – hailed by international and local palaeo-scientists as 'spectacular' – cast doubt on the prevailing story of human evolution.



[Source: <http://news.nationalgeographic.com>2013]

SOURCE G **OPINIONS OF SCIENTISTS**

'Hominid bones located in Georgia could rule out *Australopithecus sediba* as a human ancestor.'

Professor Lee Berger said that it is a pity that comparisons of *Australopithecus sediba* were not included in the study of the new hominid fossils before the announcement was made.

Caley Orr (Midwestern University) and Dean Falk (palaeoanthropologist); both think that the new theory erasing the other *Homo* species is intriguing, but believe that more specimens and additional research are needed to fully validate it.

It does not matter who finds hominid fossils or proposes new theories on hominid evolution. There is much work to be done – the important goal is to find as many hominid fossils in the shortest possible time. Only then can we work on completing the hominid family tree.

(ADJ Taylor – palaeoanthropologist; NSW University)

[Adapted: <http://evolutionnews.org>]

SOURCE H THE REWARDS IN SCIENCE

Curiosity may drive people into science, but then reality quickly sinks in. To be successful, today's scientists must often be self-promoting entrepreneurs whose work is driven not only by curiosity but by personal ambition, political concerns, and quests for funding.

[Adapted: <http://www.nytimes.com>]

Perhaps the most difficult barriers to effective collaboration (working together) in science are concerns about authorship of results and ownership of ideas or data. Some scientists avoid collaboration altogether because of a fear that their contributions will go unrecognised.

[Adapted: <http://www.huffingtonpost.com/science>]

Billionaires with big ideas are privatising American Science

President Obama believes in using science 'to grow the American economy' and presented 'the next great American project': a $100 million initiative to probe the mysteries of the human brain. "We can't afford to miss scientific opportunities while the rest of the world races ahead," President Obama said. "I don't want the next job-creating discoveries to happen in China or India or Germany. I want them to happen right here." "For better or worse," said a political analyst at the American Association for the Advancement of Science, "the practice of science in the 21st century is becoming shaped more by the particular interests of individuals with huge amounts of money." For example, many are financing hunts of dinosaur bones.

[Adapted: <http://nytimes.com>]

Discussion:

This open-ended question is complex and long with a number of detailed, but concise instructions. The student needs to read and process eight sources (spread over four and a half pages) of information about human evolution and the nature of science **(*content*)** to answer this question. The sources are from different genres, with varying amounts of relevant information which add to the difficulty associated with reading and processing the sources **(*stimulus*)**. The student needs to select relevant information from the sources and integrate this information with their own relevant biological knowledge. In addition, the student needs to take a definite stand on the question to develop an empirical, and debated argument – a formidable ***task*** for many Grade 12 students. The marks for the ***expected response*** are spread as follows: content thoroughness (4), content relevance (4), supporting argument (4), fairness of argument (3), position (1) and style and tone of writing (4). A short, concise and balanced answer of this kind is very difficult for most Grade 12 students.

This question is judged to be very difficult in relation to all four sources of difficulty.

Memorandum/Marking guidelines

COLLABORATION IS IMPORTANT	COLLABORATION NOT REQUIRED
• Much unknown about human family tree. • Should pool information for greater understanding. • Science has important rules/code of conduct that all must adhere to. • Will regulate findings: 1. Peer review 2. Regulation by authorities 3. Objectivity 4. Replication 5. Testability • This will make 'family tree' more certain as all data is scrutinised by all scientists working on hominids. • Prof. Berger sharing *A. sediba* findings; good example to scientific world • Yamanaka and Gurdon won Nobel prize for jointly discovering mature stem cells can be programmed – led to important research for therapeutic uses in humans – only owing to collaboration. • Fossils so rare, important to share any findings. • Without collaboration hominid find are contradictory.	• Findings in Europe – scientists not obliged to inform each other • Scientists might not get recognition. • Much work still to be done, plenty of opportunities for all. • Scientists divided on new findings – could hamper future progress on Hominids if they do not pursue own research. • What is intellectual property – great findings belong to scientists who discover them. • Allows prestige, money for country where scientist residing – could increase tourism – revenue. • Fossils are from the past – future research should focus on innovative scientific research, e.g. stem cells. • Human origins are not as important as, e.g. cure for cancer • Political intervention would prevent collaboration – must pursue scientific opportunities for country. • Creates jobs for economy.

- many extra facts to use from own knowledge re: SAGS
  e.g. Franklin; Watson and Crick – structure of DNA.

NB: The rubric is attached as Annexure B.

(20)

9. CONCLUDING REMARKS

This exemplar book is intended to be used as a training tool and as a source of reference to ensure that all role players are using a common set of principles, concepts, tools and framework when assessing cognitive challenge when the NSC examinations are set, moderated and evaluated. We hope that the explanations given and the examples of examination questions selected to demonstrate different types and levels of cognitive demand and different levels of difficulty will help all to achieve this goal.

REFERENCES

Ahmed, A., and Pollitt, A. (1999). Curriculum demands and question difficulty. Paper presented at IAEA Conference, Slovenia.

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1999). *Standards for educational and psychological testing.* Washington, DC: Author.

Bloom. B. S., Engelhart, M. D., Furst, R. J., Hill, W. H., & Krathwohl, D. R. (1956). *Taxonomy of educational objectives: The classification of educational goals. Handbook 1: Cognitive domain.* New York: David McKay.

Bloom, B. S., Hastings, J. T., & Madaus, G. F. (1971). Handbook on formative and summative evaluation of student learning. New York: McGraw-Hill Book Company.

Crowe, A.A. (2012). *Standards of South African Senior Certificate Biology Examinations: 1994 to 2007.* Doctoral dissertation, University of Cape Town, Cape Town, South Africa.

Department of Basic Education. (2010). *National Senior Certificate Examination November 2010 Life Sciences Paper 2.* Pretoria, South Africa: Author.

Department of Basic Education. (2011a*). National Curriculum Statement (NCS)[:] Curriculum and Assessment Policy Statement [CAPS] [-] Life Sciences Further Education and Training Phase Grades 10-12*. Pretoria, South Africa: Author.

Department of Basic Education. (2011b). *National Senior Certificate Examination November 2011 Life Sciences Version 1 Paper 2*. Pretoria, South Africa: Author.

Department of Basic Education. (2012a). *National Senior Certificate Examination November 2012 Life Sciences Version 1 Paper 2*. Pretoria, South Africa: Author.

Department of Basic Education. (2012b). *National Senior Certificate Examination November 2012 Life Sciences Version 2 Paper 1*. Pretoria, South Africa: Author.

Department of Basic Education. (2012c). *National Senior Certificate Examination November 2012 Life Sciences Version 1 Paper 1*. Pretoria, South Africa: Author.

Department of Basic Education. (2014a). *National Senior Certificate Examination November 2014 Life Sciences Paper 1*. Pretoria, South Africa: Author.

Department of Basic Education. (2014b). *National Senior Certificate Examination November 2014 Life Sciences Paper 2*. Pretoria, South Africa: Author.

Department of Basic Education. (2016a). *National Senior Certificate Examination November 2016 Supplementary Life Sciences Paper 2*. Pretoria, South Africa: Author.

Department of Basic Education. (2016b). *National Senior Certificate Examination November 2016 Life Sciences Paper 2*. Pretoria, South Africa: Author.

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

Fiddler, M., Marienau, C. and Whitaker, U. (2006). *Assessing Learning: Standards, Principles & Procedures.* (2nd Edition). Chicago: Kendal/ Hunt Publishing Company.

Frisbie, D. A. (2005). Measurement 101: some fundamentals revisited. *Educational Measurement, Issues and Practice, 24*(3), 21-28.

Independent Examinations Board (2012) *IEB National Senior Certificate Examination November 2012 Life Sciences Paper 1*. Johannesburg, South Africa: Author.

Independent Examinations Board. (2014). *IEB National Senior Certificate Examination November 2014 Life Sciences Paper 1*. Johannesburg, South Africa: Author.

Independent Examinations Board. (2016). *IEB National Senior Certificate Examination November 2016 Life Sciences Paper 3*. Johannesburg, South Africa: Author.

Independent Examinations Board. (2017). *IEB National Senior Certificate Examination Supplementary 2017 Life Sciences Paper 2*. Johannesburg, South Africa: Author.

Leong, S. C. (2006). *On varying the difficulty of test items.* Paper presented at the 32nd Annual Conference of the International Association of Educational Assessment, Singapore. Retrieved August 2011, from http://www.iaea.info/documents/paper_1162a1d9f3.pdf

Umalusi. (2009a). *From NATED 550 to the new National Curriculum: maintaining standards in 2008* [–] *2008 Maintaining Standards Report – Part 1: Overview*. Pretoria, South Africa: Author.

Umalusi. (2009b). *From NATED 550 to the new National Curriculum: maintaining standards in 2008* [–] *2008 Maintaining Standards Report – Part 2: Curriculum evaluation*. Pretoria, South Africa: Author.

Umalusi (2012). *Developing a framework for assessing and comparing the cognitive challenge of Home Language examinations*. Pretoria, South Africa: Author.

Annexure A: Rubric for Example 1 of questions at Difficulty Level 4 – VERY DIFFICULT

Method Rubric Criteria	5	4	3	2	1	0
L Layout – appearance of method					Layout meets criteria below: neat and tidy and bulleted/ numbered.	Layout is untidy and hard to read. **OR** Method is not formatted correctly with bullet points or numbers.
A Aim – Method relates to prescribed experiment.				Method clearly tests an aim that relates to the prescribed experiment and achieves the required result.	Method relates to the prescribed aim given, but is a little confusing and does not achieve the required result.	Method does not relate to the prescribed aim or achieve the desired result. Method given is the same as the given experiment.
M Method – This needs to be appropriate and relevant to the aim, clearly logical and sequential. If apparatus is given in the examination paper, the method should resemble the one given in the marking guidelines.	All 5 criteria given below are met: 1. An original experiment provided. 2. Equipment is appropriate and used correctly. 3. Measuring of solutions, reagents and marking of equipment are explained and this assists in the control of variables. 4. Instructions are scientifically valid and ordered. 5. Instructions are complete to produce measurable results that are recorded.	An original experiment provided. Plus 3 of 5 criteria are met.	An original experiment provided. Plus 2 of 5 criteria are met.	An original experiment provided. Plus 1 of 5 criteria is met.	An original experiment provided.	None of the 5 criteria are met. **OR** Method a copy of the original, given experiment.

Annexure B: Rubric for Example 3 of questions at Difficulty Level 4 – VERY DIFFICULT

	1 mark	2 marks	3 marks	4 marks
Content: Thorough-ness	• Up to 1/3 of potential detail in sources cited (**e.g**. 1 to 4 facts)	• About half of potential detail in sources cited (**e.g**. 4 to 8 facts from sources)	• All main topics in sources covered • About ¾ of potential detail in sources cited (**e.g**. 9 to 12 facts = 11 + 1 original fact*) • One instance of significant information beyond the sources	• All main topics covered • Source detail very close to full potential • At least (x) significant instances of information beyond the sources (**e.g**. 13 – 16 facts; 2 must be original & beyond the sources) = 11/14 + 2
Content: Relevance	• Mostly digression and/or repetition	• Around half is digression and/or repetition	• Repetition mostly avoided • Some minor digression • Argument relevant	• Isolated incidences of minor repetition • No digression • Argument relevant
Supporting Argument, i.e. <u>for</u>	• Writing consists of facts with little linkage or reasoning • Reasoning incorrect	• **<u>Maximum if no clear decision to support</u>** • Reasoning correct, but hard to follow • Ordinary; some linkage is evident	• Supports the position • Reasoning is clear • Minor errors in flow • Solid but not compelling; linkage sometimes missed	• Strongly supports a clear position • Reasoning is very clear and succinct • Flow is logical, showing evidence of clear planning • Compelling with regular use of linking language
Fairness i.e. Argument <u>against</u>	• One counter opinion given	• Two counter opinions given	• Three or more counter opinions given	
Position	• **<u>Clear decision made</u>**			

Presentation	• Writing is almost unintelligible • Tone, language and terminology unscientific and exceptionally weak • Introduction **and/or** conclusion **not** present	• Tone, language and terminology is weak • Attempts at correct paragraphing • **Introduction** and **conclusion** present, no matter how weak	• Tone is consistent and suited to scientific argument • Good and appropriate language and terminology • Mostly appropriate paragraphing • **Introduction** and **conclusion** have merit.	• Tone mature and suited to scientific argument • Excellent and appropriate use of language and terminology • Correct paragraphing with good transitions • Interesting **introduction**, satisfying **conclusion**

The design grid above is the draft version. The final marking guidelines were expanded in consultation with the marking panel at the standardisation meeting.