DIFFICULTIES IN THE COMPREHENSION AND INTERPRETATION OF A SELECTION OF GRAPH TYPES AND SUBJECT-SPECIFIC GRAPHS DISPLAYED BY SENIOR UNDERGRADUATE BIOCHEMISTRY STUDENTS IN A SOUTH AFRICAN UNIVERSITY

by

ANDRÉ VAN TONDER

submitted in accordance with the requirements for
the degree of

MASTER OF SCIENCE IN MATHEMATICS, SCIENCE AND TECHNOLOGY EDUCATION

in the subject

SCIENCE EDUCATION

at the

UNIVERSITY OF SOUTH AFRICA

SUPERVISOR: PROF. H.I. ATAGANA

NOVEMBER 2010

CONTENTS

CHAPTER 3.   METHODOLOGY

CHAPTER 4.   DATA ANALYSIS, RESULTS AND DISCUSSION

CHAPTER 5.   CONCLUSIONS AND RECOMMENDATIONS

(i)   List of tables

(ii)   List of figures

(iii) Abstract/Summary

A carefully constructed set of 16 graphical tasks related to key biochemistry concepts was designed and administered to a group of 82 students in their final year of B.Sc. study.

The test mean score of 48,3% ($\pm$ 12,1) was low and characterised by gender and ethnic differences. There was a moderate linear relationship between biochemistry grades obtained by the students over two years of study and their graphical literacy ($r = 0,433$). The majority of the students exhibited slope/height confusion and only seven students (8,5%) were able to answer the two items corresponding to Kimura's Level F, the most complex and difficult level of graphical literacy.

Eye tracking data gave valuable insights into different strategies used by students while interpreting graphs and is a valuable tool for assessing graphical literacy.

These findings confirmed other studies where researchers have found a widespread lack of graph comprehension among biological science students.

(iv) Acknowledgements

I would like to acknowledge the help and goodwill of the following people in this endeavour:

CHAPTER 1. INTRODUCTION

1.1 Background of the study

Presenting scientific data with graphs, diagrams and tables is a central practice in the natural and social sciences and is a basic skill of the scientist. Graphs in particular have the potential to represent complex scientific data more successfully than a combination of text and tables and students need to be graph-literate in order to utilise scientific data presented in graphs. Graphs play an important role in the modelling and understanding of complex biochemical systems and are ubiquitous in the introductory biochemistry curriculum. Graphs work well because humans are good at seeing things - our visual perception is particularly well-developed.

Biochemists use a variety of simple graphs that, in theory, nearly every student can comprehend. Graphs can be a visual way of illustrating enzyme catalysis, protein denaturation, buffer action, gene homology or metabolic energetics, for example. They are also used to visualise large sets of enzyme kinetic data and to help Biochemists understand important systems that might not be easy to understand just by perusing tables of data. In addition, secondary plots, where data obtained from a series of one type of graph is used to plot a different type of graph, are used extensively in enzymology. Graphical literacy is becoming more important for biochemistry researchers, post-graduate and under-graduate students as the advent of new technology provides access to more and different types of data. This abundance of biochemical information in various forms requires an increased capacity to evaluate critically what is being presented.

It has been suspected for a long time that students have difficulty in interpreting scientific biochemistry information as it is displayed visually in a graph format. Staff lecturing in biochemistry at the University of the Free State (UFS) assumed that abstract concepts and abstract elements used in graphs or depicted in scientific graphs are automatically understood by senior undergraduate students. However, the students' answers to test and examination questions, where there were graphs involved, pointed to a potential deficiency in the students' grasp of the basics of graphing or plotting as well as interpretation with, on average, about 25% of the students (predominantly females) scoring zero or less than full marks for the task. Deficiencies that seemed most prevalent were:

    (i)   Choosing the appropriate axis for a variable;

    (ii)   Placing ordered pairs on the plot;

(iii)   Determining scale and labelling the axes with numbers (difficult for many students);

(iv)   Determining the slope and interpreting it;

(v)   Identifying ordered pairs;

(vi)   Converting reciprocal values;

(vii)   Identifying maxima and minima; and

(viii)   Identifying or predicting a trend.

These difficulties are surprising since these students should have gained extensive experience of plots and graphs throughout their school careers as well as first year B.Sc. mathematics, chemistry and physics courses.

It was also theorised by the academic staff that the students' average academic performance correlated with their ability to interpret and comprehend graphs and instructional diagrams. In a preliminary study conducted by Van Tonder and De Lange (2002) the graphical and visual literacy of biochemistry students with regards to their ability to interpret scientific biochemistry information in a graph format as well as instructional diagrams was tested using a short graphical literacy test. The data that was gathered indicated poor pictorial capabilities among some biochemistry students. The graphical literacy test scores of the students were compared with their biochemistry grades and a weak positive correlation was found.

1.2 Statement of the problem

This research project explores the extent of what appears to be a deficiency in graph decoding skills exhibited by senior undergraduate biochemistry students in the University of the Free State. In particular, the research focuses on identifying specific deficiencies as well as attempts to identify the cognitive basis of these deficiencies.

1.3 Purpose of the study

The purpose of this study was therefore five-fold:

1.   to assess, by developing and applying a reliable graphical literacy test, the extent of what appears to be a problem with graph decoding skills exhibited by biochemistry students at the UFS;

2.   to confirm whether there is any correlation between the graphical literacy test results and the students' academic performance in biochemistry;

3.   to identify the effect, if any, of gender and ethnicity on graphical literacy;

4.   to attempt, using eye tracking, to identify the possible sources of deficiencies in graph decoding skills exhibited by the students, and;

5.   to guide the design and implementation of interventions that can be utilised as solutions for correcting or remediating the identified difficulties.

1.4 Significance of the study

The results obtained from a *biochemistry* graphical literacy test, capable of identifying the achievement levels of biochemistry students, can be used to guide interventions by teachers of undergraduate biochemistry. These interventions can take the form of tutorials in as well as extra time being spent during lectures on enzyme kinetics, for example. The results of this study could also inform other introductory courses in the life sciences with respect to graphical literacy difficulties experienced by students in the life sciences. In addition, eye tracking data will enable the researcher to gain insight into the student's cognitive processes (e.g. where they are searching for information and whether they are experiencing difficulties) when viewing the given graphical representations.

1.5 Research questions

The following research questions will be investigated in this study:

Question 1:   What are the subject-specific graphical literacy levels of senior undergraduate biochemistry students in their third year of study at the University of the Free State?

Question 2:   Are the graphical literacy levels of undergraduate biochemistry students reflected in their academic performance?

Question 3:   Is gender or ethnicity related to graphical literacy levels of senior undergraduate biochemistry students?

Question 4:   Is it possible, using eye gaze data obtained using subject-specific graphs, to discern a difference in the way senior undergraduate biochemistry students with high graphical literacy and students with low graphical literacy interpret graphs, and thus identify possible sources of graph-decoding deficiencies?

1.6 Statement of hypotheses

Four hypotheses were formulated for this study:

Hypothesis 1: Senior undergraduate biochemistry students at the UFS experience difficulty with graph comprehension and display poor graphical literacy levels when assessed, specifically an arithmetic mean score of less than 50% on a subject-specific graphical literacy test incorporating examples of common graph decoding tasks as well a selection of biochemistry concepts in Cartesian graphs. In addition, students perform poorly, if at all, at Kimura's Level F (an abstract level of graph interpretation requiring the creation of dimensionally new information 'beyond the data'). The null and alternative hypotheses are stated as follows:

Null hypothesis: H_o = adequate graphical literacy levels, including Kimura's Level F, and an arithmetic mean score > 50%.

Alternative hypothesis: H_1 = inadequate graphical literacy levels, including poor performance, if any, at Kimura's Level F, and an arithmetic mean score < 50%.

Hypothesis 2: The graphical literacy levels of senior undergraduate biochemistry students at the UFS are reflected in their academic performance in biochemistry, specifically evidenced by a moderate to strong positive correlation between graphical literacy test mean scores and the arithmetic mean of the students' grades in biochemistry modules. The null and alternative hypotheses are stated as follows:

Null hypothesis: H_o = Pearson's $r < 0{,}3$ at a confidence level of 95% ($p < 0{,}05$).

Alternative hypothesis: H_1 = Pearson's $r > 0{,}3$ at a confidence level of 95% ($p < 0{,}05$).

Hypothesis 3: There are gender and ethnicity differences (in the favour of white males) in the graphical literacy levels of senior undergraduate biochemistry students at the UFS, specifically evidenced by a higher mean score on the graphical literacy test. The null and alternative hypotheses are stated as follows:

Null hypothesis: H_o = arithmetic mean scores on graphical literacy test are not statistically different at a confidence level of 95% ($p < 0,05$).

Alternative hypothesis: H_1 = arithmetic mean scores on graphical literacy test are statistically different at a confidence level of 95% ($p < 0,05$).

Hypothesis 4: Eye gaze data obtained using subject-specific graphs can be used to discern differences in the way senior undergraduate biochemistry students with high graphical literacy and students with low graphical literacy interpret graphs. The null and alternative hypotheses are stated as follows:

Null hypothesis: H_o = eye gaze data revealed no differences in task response times and mean gaze times in predefined areas of interest at a confidence level of 95% ($p < 0,05$) and no difference in the patterns of heatmaps, gaze plots and scan paths.

Alternative hypothesis: H_1 = eye gaze data revealed differences in task response times and mean gaze times in predefined areas of interest at a confidence level of 95% ($p < 0,05$) and differences in the patterns of heatmaps, gaze plots and scan paths.

1.7 Operational description of terms.

Graphical literacy: defined in the literature as "the ability to construct, produce, present, read and interpret charts, maps, graphs, and other visual presentations and graphical inscriptions" (Readence, Bean & Baldwin, 2004), in this study graphical literacy refers to the ability to interpret (decode), predict and infer from a selection of graphical representations of biochemical data.

Eye tracking: typically defined as "…using a device to record exactly where in a picture a person is looking at any given moment in order to study the way in which the gaze wanders over a picture, pauses and fixes on certain points" (Pettersson, 1989), in this study eye tracking refers to the use of an eye tracker to monitor and track the test participant's eye movement and fixations in real time while viewing a selection of graphical representations of data.

1.8 Organisation of the study

The organisation of this report from Chapter 2 to the end can be visualised as follows:

LITERATURE REVIEW	→	Conceptual framework What is a graph? Graphical literacy Graph comprehension Elements of graph design Assessing levels of graphical literacy Eye tracking
METHODOLOGY	→	Research design Population of the study Sampling Measurement instruments Data collection procedures Methods of data analysis
DATA ANALYSIS, RESULTS AND DISCUSSION	→	Research question 1 Research question 2 Research question 3 Research question 4 Test of the hypotheses
CONCLUSIONS AND RECOMMENDATIONS	→	Summary of the study Recommendations Contributions to knowledge Limitations of the study Suggestions for further studies
REFERENCES		
APPENDICES	→	A: Questionnaire #1 B: Graph- % correct responses per question C: Graph- individual student performance at Kimura's levels D: Graph- individual student performance at Curcio's levels

CHAPTER 2.   REVIEW OF RELATED LITERATURE

2.1 Conceptual Framework

When scientific conceptions are perceived as conceptually challenging, students may show reasoning difficulties or form alternative conceptions, which are not scientifically acceptable and can have the potential to prevent the learning and appreciation of science. In the discipline of biochemistry it is important to identify and classify student difficulties so that remediation through teaching interventions can be initiated and formulated.   However, teachers of biochemistry have access to relatively few publications in the area of learning and conceptual understanding of biochemistry, unlike in the disciplines of physics (Pfundt & Duit, 1998; Harrison, Grayson & Treagust, 1999), chemistry (Garnett, Garnett & Hackling, 1995) and biology (Lazarowitz & Penso, 1992). This is despite the fact that biochemistry is generally regarded as being very cognitively demanding, both in terms of conceptual complexity and the scientific reasoning required (Wood, 1990; Vella, 1990). The relatively few publications in biochemistry education research have been limited mainly to the field of metabolism.   Hanson (1989), for example, has discussed the role of adenosine triphosphate (ATP) in metabolism while more extensive results have been published by Anderson and Grayson (1994) who investigated biochemistry students' difficulties experienced with studying carbohydrate metabolism.   Student difficulties with the interpretation of textbook diagrams is also an emerging field in biochemistry education research (e.g.: Schönborn, Anderson & Grayson, 2002).

An efficient methodological framework for the identification and classification of student conceptual or reasoning difficulties in science has been developed by Grayson, Anderson and Crossley (2001). A more recent initiative by the International Union for Biochemistry and Molecular Biology is that of developing a 'concept inventory' (CI) with the aim of assessing students' understanding of concepts before or in response to instruction or both (Sears, 2008). Sears (2008) emphasises that CI questions that include visual representations are problematic because analysis of student responses are complicated by the inability to distinguish between the students' conceptual understanding and the students' ability to interpret and make inferences from the symbolism used in the accompanying reaction schemes or graphs. Interestingly, Sears and Thompson (2007) used a titration curve (as a semi-logarithmic data plot) in a pre-instructional assessment of basic chemical literacy and basic

molecular literacy and concluded that students were revealing misconceptions in interpreting titration data. It would have been more revealing to investigate the students' familiarity with and ability to interpret semi-logarithmic plots in general since this type of graph is a challenge for most people, scientists included!

Research into the impact of diagrams and illustrations on the learning of science is not a new field. Hill (1988) found that chemistry students were likely to hold conceptions that were consistent with common errors found in science textbook illustrations. Hill (1988) claims that students interpret diagrams of relationships in textbooks qualitatively and globally and are unable to distinguish between the diagram and the model that it represents. The students are therefore highly likely to form alternative conceptions. Brasell and Rowe (1993) also reported that physics students who are from a rural area have more difficulty comprehending graphs and also that they had errors associated with linking graphs and verbal descriptions. Making sense of graphs appears to be a complex process. A group of researchers, working in this field, identified three components (known collectively as *graph sense*) of graph comprehension, namely, understanding the convention of a graph design; the decoding process of a graph; and the process of generalising or relating the information from the graph to the context of the situation (Friel, Curcio & Bright, 2001). They stated that more research is needed to understand what makes these three components difficult for students.

From an article by Lake (1999) it appears as if secondary and tertiary students in general have difficulty reading and interpreting graphs, are not always able to discern important issues illustrated by graphs and tables and that little attention is given to develop this area of literacy. In similar studies (Knuth, 2000; Hale, 2000) where graphs were involved, authors report that many students who have completed secondary school do not understand the connection between equations and graphs; even if students understand certain mathematical concepts, that they have difficulty with graphs portraying these concepts; and that students can read data from a graph but still have difficulty understanding the meaning of the graph.

Most of the work reported on graph interpretation appears to be in the field of physics, perhaps because of their widespread use as a teaching tool. Since graphs are such an efficient way of illustrating packages of data, they are used almost as a language by physics teachers. Unfortunately, many studies have indicated that students do not share the vocabulary of their teachers. Research by McDermott, Rosenquist & Van

Zee (1987), Beichner (1994) and many others has uncovered a consistent set of student difficulties with graphs of position, velocity, and acceleration versus time - graphs otherwise known as kinematic graphs in physics. These difficulties include misinterpreting graphs as pictures, confusion between the slope and height of a graph line, problems finding the slopes of lines not passing through the origin, and the inability to interpret the meaning of the area under various graph curves.

What can be done to address the difficulties students have with the interpretation of graphs? The first step is for teachers to become aware of the problem and, hence, studies such as this one will provide valuable information for teachers of Biochemistry.

2.2 What is a graph?

Graphs are a format used to display scientific data, or to illustrate information that is difficult to describe with text, in textbooks and other popular print or electronic media (Shah, Mayer & Hegarty, 1999; Van Tonder & De Lange 2000; Shah & Hoeffner 2002; Renshaw, Finlay, Tyfa & Ward, 2004).

The purpose of graphs, as with most data summarisation techniques, is to quickly convey information (Fischer, 2000) enabling fast and accurate data extraction (Hink, Eustace & Wogalter, 1998). Essentially: "Line graph construction and interpretation are very important because they are an integral part of experimentation, the heart of science" (McKenzie & Padilla, 1986). The importance of graphs in the practice of science cannot be overestimated, in fact it has been argued that there is no other statistical tool as powerful for facilitating pattern recognition in complex data (Chambers, Cleveland, Kleiner & Tukey, 1983).

Interestingly, the term "graph" was first introduced by J.J. Sylvester in a paper published in 1878 in Nature. Sylvester (1878) draws an analogy between atoms (valences or number of bonds) and binary quantics (algebraic forms), more specifically "quantic invariants" and "co-variants" of algebra and molecular diagrams, as follows:

> "[...] Every invariant and co-variant thus becomes expressible by a *graph* precisely identical with a Kekuléan diagram or chemicograph. [...] I give a rule for the geometrical multiplication of graphs, *i.e.* for constructing a *graph* to the product of in- or co-variants whose separate graphs are given. [...]" (italics as in the original).

Although the chemico-algebraic theory of Sylvester was short-lived, it was a major stimulus for the study of *graphs* as objects of interest in themselves (Biggs, Lloyd & Wilson, 1976).

The use of graphs and plots are key in introductory courses in which quantitative skills are emphasized because they are the essence of giving students multiple representations of mathematical concepts which can be expressed numerically, visually, and symbolically (Wenner, 2009). In fact, it can be said that graphs bridge the gap between perceptual and conceptual relations and can, therefore, support reasoning with respect to abstract concepts such as 'slope' and 'function' that are difficult to grasp directly (Gattis & Holyoak, 1996).

To summarise: A graph is a visual representation of a relationship between two (and sometimes, three) variables x and y (and z). The reason that we plot data is so that we can more easily observe trends or behaviour of the data.

2.3 Graphical literacy

The relatively newly-coined term *graphical literacy* (shortened by some authors to *graphicacy*) refers to the ability to construct, produce, present, read and interpret charts, maps, graphs, and other visual presentations and graphical inscriptions (Readence, *et al.*, 2004). It is a visual language or skill for enhancing learning. According to dual-coding theory, information is easier to retain and retrieve when it is coded both verbally and visually (Paivio, 1991) and, thus, adding graphics to text can improve learning (Clark & Mayer, 2002). Visualization is also a powerful cognitive tool in scientific discovery and invention, and essential to problem solving in daily life as it provides concrete means to interpret abstract images (Rieber, 1995). Visualisation tools such as physical and molecular models, photographs, micrographs, pictures, diagrams, metabolic maps, graphs and animated visuals, collectively known as external representations (Lohse, Walker, Biolsi & Rueter, 1991), are particularly valuable in the teaching and learning of biochemistry. However, biochemists often use a range of symbolism to represent the same phenomenon, a convention that requires biochemistry students to learn a "visual language" that is far more complex and potentially confusing than that used in other disciplines (Schönborn & Anderson, 2006).

Based on a review of experimental and theoretical work in this area, Winn (1987) describes the cognitive value of visual representations and visually-oriented processing as follows:

> Visual representations use an entirely different type of logic based on the meaningful use of space and the juxtaposition of elements in a graphic. ... Graphic forms, conveying information by means of visual argument, can induce the use of cognitive processes that are themselves "visual" in some way. ... The advantage of using visual argument lies in the application, by students, of cognitive abilities that are particularly suited to what has to be learned. ... Graphic forms encourage students to create mental images that, in turn, make it easier for them to learn certain types of material. ... Presenting information graphically allows students to scan it rapidly and quickly to discover patterns of elements within the diagram that are meaningful and that lead to the completion of a variety of cognitive tasks. ... Charts, diagrams and graphs are effective in instruction because they allow students to use alternative systems of logic. ... Certain physiological strengths of learners, such as pattern recognition and the ability to recognize geometric shapes, as well as the advantages of "right-brain" processing, can be exploited. (pp. 156, 157, 158, 159, 160)

These visually-oriented cognitive processes include, but are not limited to, the construction of mental models.

Not unexpectedly, the interpretation of visual language is a skill that depends, to some extent, on experience in a particular domain of knowledge. This skill is known as "visual literacy", the earliest definition of which states that *"visual literacy is the ability, through knowledge of the basic visual elements, to understand the meaning and components of the image"* (Dondis, 1973).

Visual literacy is a concept that has its roots in philosophy, art, linguistics, imagery theory, perceptual psychology and communication research (Moore & Dwyer, 1994). Dondis (1973) noted that it includes comprehension and that it must give meaning to the subject. The term "visual literacy" is, according to Messaris (1994), *"... an appropriate label to describe a viewer's familiarity with specific images or sets of images that have played a role in his or her culture's visual heritage ...".* A definition, proposed by Braden and Hortin (1982) states: *"Visual literacy is the ability to understand and use images, including the ability to think, learn, and express oneself in terms of images."* It is the ability to understand (read) a visual and to interpret (give meaning to) the visual in its context (its educational text). A person's level of visual literacy is influenced by his or her prior experience and exposure to visual media, for example television, magazines, advertising, comics and other forms of visual literature – media that are usually culturally biased. A person's level of visual literacy can furthermore be affected by educational and socio-economic variables (Peeck, 1993).

It follows, therefore, that poor visual literacy, or poor pictorial literacy, combined with the screen effect of culture can contribute to poor learning performance when using picture-text learning material.

It is perhaps necessary at this point to address the often erroneous confusion between visual literacy and visual intelligence (exacerbated by writers of "self-help"/improvement books). Visual/spatial intelligence (one of the multiple intelligences identified by Howard Gardner) is the process of forming mental representations of <u>three-dimensional reality</u> as a basis for *understanding one's environment and interacting with it effectively* - a type of intelligence crucial for success in professions such as architecture or carpentry, but it is also a vital ingredient of any person's everyday physical activities (Gardner, 1999). Visual literacy, on the other hand, is a skill or ability that can *contribute* to visual/spatial intelligence. It can also be involved in other intelligences such as bodily/kinesthetic, musical, interpersonal, intrapersonal, linguistic, and logical-mathematical. In fact, graphical literacy more than likely involves logical-mathematical intelligence.

Thus, graphical literacy appears to be a component of visual literacy and this author finds it useful to summarise and conceptualise this distinction as follows:

> *Visual literacy*   =   *"communicating with pictures"*
>
> *Visual intelligence*   =   *"thinking in pictures"*
>
> *Graphical literacy*   =   *"communicating with graphs"*

It is generally accepted that, particularly in our modern information age, there is a growing need for graphical/visual literacy. However, there is less attention paid to it at the elementary level than there is to reading and writing. There is also evidence that higher order visual literacy skills do not develop unless they are identified and "taught" (Avgerinou & Ericson, 1997). Visual presentations of abstract concepts tend to be difficult for students yet they are ignored in basic and other school texts (Readence, *et al.*, 2004). Fortunately, this shortcoming has been recognized in the field of biochemistry and Schönborn and Anderson (2006) have called for visual literacy teaching to be included in the biochemistry undergraduate curriculum. This echoes a call by researchers in education for increased attention to be paid to graphical instruction to help students become literate in practices related to the production and interpretation of graphics (Roth, 2002).

Graphs are essential communicative tools used in the biological sciences (Tairab & Khalaf Al-Naqbi, 2004) and are a central practice in the natural sciences (Roth, 2002). Researchers are in agreement that it is vital for science students to be able to construct, interpret, predict and infer from graphical representations. Researchers have also found a widespread deficiency of these abilities among biological science students (Tairab & Khalaf Al-Naqbi, 2004).

Difficulties with graphical literacy span the age spectrum as school-aged children (Shah & Hoeffner, 2002; Hadjidemetriou & Williams, 2000), tertiary students (Van Tonder & De Lange, 2002; Tairab & Khalaf Al-Naqbi, 2004), adults (Guthrie, Weber & Kimmerly, 1993; Shah & Hoeffner, 2002) and even scientists (Roth, 2002) have exhibited shortcomings.

2.4 Graph comprehension

According to Shah and Hoeffner (2002), and, more recently, Shah, Freedman and Vekiri (2005), the difficulty in comprehending a graph is not just a function of the characteristics of the graph itself or the graphic design elements :- comprehension is also influenced by how those features interact with the prior knowledge and objectives of the viewer.

The three component processes of graph comprehension according to the dominant cognitive model (Shah *et al.*, 1999) are:
i)   **encoding** the visual array and identifying important visual features (e.g. curved line or straight line slanting downward);
ii)   **relating** those features with the quantitative facts or conceptual relations that the represent (e.g. a decreasing linear relationship between x and y), and;
iii) **associating** those relations being quantified to the graphic variables depicted (e.g. population versus year) (Shah *et al.* 1999; Shah & Hoeffner, 2002) (see also Friel *et al.*, 2001).

The cognitive features mentioned above are, however, heavily influenced by other factors. Shah and Hoeffner (2002), in an extensive review of empirical research on graph comprehension, attempted to elucidate how viewers interpret graphs and the factors that aided or hindered the comprehension of these graphs. They concluded that graph comprehension is affected by three factors:

Firstly; the actual visual characteristics of the graph, for example the perceptual features such as colour and the format of the display such as two- or three-dimensional graphs (i.e graphic design);

Secondly; the viewers' prior knowledge and expectations about the graph affected how they decoded the graph (i.e. graphical literacy), and;

Thirdly; the viewers' prior knowledge about the information depicted in the graph.

The third factor, as identified by Shah and Hoeffner (2002), was the subjects' prior knowledge of the information depicted in the graphs. With reference to picture-text studies the published research indicates that the facilitating effect of pictures becomes evident if subjects have minimal prior knowledge of the subject or of related subjects (Mayer & Gallini, 1990; Mayer, 1993). A high prior knowledge might induce a ceiling effect and the potential benefit of supporting pictures might not be evident. This low prior knowledge principle appears initially to be in conflict with the results of Shah and Hoeffner (2002) who indicated that prior knowledge of information can assist in graph comprehension. These contradictory results are not, in fact, in opposition to each other. The facilitation effect of pictures becomes evident if students have low prior knowledge of the learning material. Pictures therefore act as a second cueing device and are supposed to act as a secondary source of information retrieval. High prior knowledge appears to assist students in comprehending a graph when the graph reflects content of this knowledge. Hence it is then hypothesized that graphs will also not show learning facilitation if the student has high prior knowledge of information depicted in the graph but that it will assist in comprehending the graph.

Another important factor that has an impact on graph comprehension is that of familiarity (Kumar & Benbasat, 2004). During the activity of reading, for example, the reader attempts to establish what the text is saying about some aspect of the world. For this to occur, the reader must be familiar with the words, concepts and real-world objects referred to. The activity of reading then goes beyond the material basis of the text to that which the text is said to be about. When the reader is unfamiliar with the content a text refers to, he or she is likely to experience difficulties comprehending the text. The same concept of familiarity applies to graphs. The reading of graphs by people, such as scientists, who are thoroughly familiar with the type of graph and that which it refers to, becomes transparent in the same way text does. When the graphs and phenomena are unfamiliar, problems in reading become apparent (Roth, 2002).

Thus, prior knowledge and experience in graphs, graphing practices and mathematical content (such as number concepts) becomes an absolute necessity as it determines the level at which the material is analysed, thereby providing insight into the level of graph literacy of the viewer (Friel *et al.*, 2001; Van Tonder & De Lange, 2002; Postigo & Pozo, 2004).

With the aforementioned in mind the following deductions can be made with reference to the comprehension of graphs by students:

i) **A student must be graph literate**. Prior knowledge of a graph, or graph literacy, will determine if a student can comprehend graphs and therefore utilize graphs effectively.

ii) **A student must be subject-literate**. Prior knowledge of the information depicted in the graph will determine if a learner can comprehend the information in the graph.

iii) **A graph must be correct**. Graph- and subject-literate students will be able to effectively comprehend and utilize a graph as well as the information depicted in the graph if the graph is graphically and textually correct. This correctness of the graph is directly related to the graphic and textual conventions that the student is accustomed to and the graphic and textual (subject information) literacy of the student.

Broadly speaking, a person's level of graph-visual literacy is therefore influenced by his or her prior experience, use of and exposure to graphs and graphing practices within a scientific educational environment.

Another (non-politically correct) factor that might influence graph comprehension is that of gender (Linn & Petersen, 1985). Performance differences between males and females on non-verbal or spatial graphics tasks have been extensively reported in the literature, however, the extent of these differences, the age when these differences occur and the nature of the tasks have raised considerable debate. Reasons for apparent performance differences between males and females include the confidence levels of females (Forgasz, Leder & Kauer, 2001); their attitude towards mathematics (Forgasz, Leder & Kloosterman, 2004); everyday experiences (Tracey, 1990); as well as the manner in which tasks are presented, for example short-answer questions tend to advantage males (Lokan, Greenwood & Cresswell, 2001). The most recent study in this field reported that boys aged 9-12 years outperformed girls on graphical

languages (using the Graphical Languages in Mathematics instrument) that required the interpretation of information represented on an axis and graphical languages that required movement between two- and three-dimensional representations (Lowrie & Diezmann, 2011).

Irrespective of the aforementioned factors, graph-based reasoning can also be profoundly influenced by working memory limitations during the time course of problem solving. Analysis of eye-movement data by Peebles and Cheng (2002) revealed fine grained fixation patterns produced as a natural consequence of the decay in activation of perceptual chunks over time. Peebles and Cheng (2002) have constructed a cognitive model (ACT-R/PM) of graph-based reasoning that predicts that graph users initially encode all three elements of the question (*1. find the start location determined by the given variable; 2. find the given location representing the given value; 3. find the required location representing the required variable*) but are required to re-encode parts of it as the problem progresses, with the probability of re-encoding increasing over time.

Eye movement data is becoming a valuable tool in the study of graph-based reasoning and this author predicts that more and more researchers will challenge long-held assumptions regarding factors that impact graph comprehension.

2.5 Elements of graph design affecting comprehension

2.5.1 Challenges associated with data presentation

There are two fundamental challenges to be faced when presenting data (Few, 2004). Firstly, determining the most appropriate medium of display (table, graph or text) and secondly, designing the components of the chosen medium in order to display the data and message clearly and without distraction. According to Few (2004), effective data presentation is made possible if there is a constant awareness of the goal – to communicate information, not entertain.

Showing data inadequately (e.g. not showing enough data), showing data inaccurately (e.g. using a bar graph instead of a line graph or not starting the y-axis at zero) and obscuring data (e.g. leaving in the grid or using a double axis graph with different scales) are three ways of displaying data poorly (Bracey, 2004). An example of

showing data in a misleading manner is the use by the advertising industry, as well as financial consultants, of the so-called 'zoom' graph. A zoom graph is a graph that does not start at the origin (0:0) thus making small changes seem big (Cox, 2010). The general consensus is that the use of additional or non-informative features in a graphic display should be kept to a minimum, as these are often unhelpful and distracting (Shah & Hoeffner, 2002). The graphic theorist Edward Tufte introduced a concept which he called the "data-ink ratio" (Tufte, 1983) that illustrates how the number of elements in a graph can be reduced without losing any information (Fischer, 2000; Bracey, 2004; Few, 2004). The "data-ink ratio" can be described as the proportion of ink (or pixels on a screen) used to display actual data compared to the total amount of ink (or pixels) in the entire display. The goal of the graph designer, then, is to maximise the "data-ink ratio" without eliminating the necessary elements required for effective data communication (Few, 2004). Data ink can be described as the core of a graphic display – non-erasable and necessary for communication (Few, 2004). Non-data ink is an additional, usually irrelevant, feature of a graphic display used to make it more attention-grabbing for the viewer (e.g. hard-to-read elaborate fonts, colourful or moiré shading and a pseudo third dimension) (Bracey, 2004).

For effective communication of graphical information, priority must be given to the clarification of data by emphasizing the essential data (Renshaw, Finlay, Tyfa & Ward, 2003) and eliminating that which does not support the data and its message in an effective manner (Few, 2004).

2.5.2 Visual features

Studies on the comprehension of quantitative information in graphical displays with particular reference to the influence of the following visual features have been extensively reviewed by Shah *et al*. (2005). The following subsections are therefore of necessity rather brief with the aim of giving the reader a short overview of findings with respect to certain visual features of graphs.

2.5.2.1   Colour

A visual feature commonly used by graph designers is that of colour (Shah & Hoeffner, 2002). Colour can be used as a contrasting attribute in order to highlight certain data sets (brighter colours stand out more than less saturated colours) (Few 2004), to represent quantitative information (the deeper the

colour the higher the population), to mark variables (red bars for girls and blue bars for boys) and to group elements in a display (Shah & Hoeffner, 2002). A study of viewer's eye fixations, using eye tracking, during graph interpretation showed that viewers continuously re-examined the labels to refresh their memory proving that viewers had difficulties keeping track of these graphic referents. Meaningful colour choices should be used to represent these dimensions, which might help the viewers to keep track of the variables (Shah & Hoeffner, 2002).

2.5.2.2   Labels

Renshaw *et al*. (2004) claim that it is easier to comprehend a visual display when emphasis is placed on simple pattern identification processes that avoid complex cognitive processes. This is achieved by using simple designs, avoiding too many dimensions in one graph and the use of direct labelling of lines. The general consensus is that graph designers should directly label graph features (lines and bars) according to their referents (taking care not to "clutter" the display) and to avoid using legends (keys) as they pose special demands on a viewer's working memory (Shah & Hoeffner, 2002).

2.5.2.3   Size

Size does matter! A commonly used visual feature is size (Shah & Hoeffner, 2002). Bigger objects such as bigger fonts draw more attention because they stand out more. The title of a graphic display is important because it provides the viewer with the context in which the data is presented and can be visually highlighted by using a large font size which will visually communicate its importance (Renshaw *et al*., 2004).

2.5.2.4   Irrelevant depth cues (adding a "third" dimension)

Although the general convention is to keep additional, non-informative features to a minimum, the use of a third "spatial" dimension as a perceptual, but totally decorative, characteristic is still common (Shah & Hoeffner, 2002) This is possibly as a result of commercially available graphing software packages supporting a variety of options to display any set of data as well as design

features such as colours and forms which are often used by non-graphic specialists to communicate information in a visual way (Renshaw *et al.*, 2004).

Text book or instructional material authors in favour of using these additional depth cues argue that it makes the graph more attractive and "attention-grabbing" which, in turn, enhances the depicted information's memorableness. Those opposed to this practice argue that graphs with an added "third dimension" appear more cluttered which can slow or retard comprehension (Fischer, 2000). Tufte (1983) even refers to these irrelevant depth cues as an example of "chart junk" which should be avoided (Shah & Hoeffner, 2002).

Studies undertaken to investigate the value of an added third dimension include, amongst others:

i)   Spence (1990), who discovered that relations between three-dimensional features such as cylinders and boxes were generally estimated equally accurate and even somewhat quicker than their two-dimensional counterparts such as discs, lines and bars.

ii)   Carswell, Frankenberger and Bernhard (1991), who investigated the role of irrelevant depth information on the processing of bar graphs, line graphs and pie charts and found a negative effect only where line graphs were concerned.

iii) Zacks, Levy, Tversky and Schiano (1998), who added irrelevant depth cues to bar graphs and found lower accuracy associated with the additional "third-dimension".

iv)   Siegrist (1996), who found an overall better performance for two-dimensional charts.

v)   Fischer (2000), who conducted a study using depth cues on bar graphs and measured the decision times to determine whether this added feature had an impact on graph comprehension times. In his findings Fischer (2000) states that irrelevant depth cues negatively affected decision times and that a considerable time cost was associated with three-dimensional frames over two-dimensional frames. When bars were also rendered in depth processing costs increased even more.

Thus, no processing benefits were obtained from consistency of depth between graph elements. Therefore, complexity rather than consistency affected decision times. Based upon these results, Fischer (2000) recommends avoiding the use of irrelevant depth cues altogether.

2.6 Assessing levels of graphical literacy

It is generally accepted that the following basic skills are essential for drawing and interpreting graphs (see for example *http://cstl.syr.edu/fipse/GraphA/Unit2/Unit2.html*):

- *Define the terms constant and variable.*
- *Identify whether an item is a constant or a variable.*
- *Identify whether an item is a dependent or independent variable.*
- *Identify the x and y axes.*
- *Identify the origin on a graph.*
- *Identify x and y coordinates of a point.*
- *Plot points in a graph.*
- *Draw a graph from a given equation.*
- *Determine whether a given point lies on the graph of a given equation.*
- *Define slope.*
- *Calculate the slope of a straight line from its graph.*
- *Identify if a slope is positive, negative, zero, or infinite.*
- *Identify the slope and y-intercept from the equation of a line.*
- *Identify the y-intercept from the graph of a line.*
- *Match a graph with its equation.*

These skills are rather narrow in the sense that they are limited to processing and converting data into graphs and reading simple quantitative information from graphs, tasks that would fall under the definition of *graph sense* coined by Friel *et al.* (2001). These skills can also be relatively easily assessed. The ability to interpret, predict and infer from graphical representations implies higher order cognitive processing skills that, together with *graph sense*, reflect the development of *graphical literacy*.

Kimura (1999), a distinguished Japanese science educator, has suggested that a key component of statistical literacy is the ability to extract qualitative information from quantitative information, and/or to create new information from qualitative and quantitative information with an emphasis on graphical representations of statistical phenomena. He also proposed a hierarchy of statistical understanding in relation to graphical representations.

Kimura's (1999) proposed hierarchy reflected a growing trend among researchers in education or literacy to use frameworks denoting increasing sophistication in thinking and reasoning skills that have emerged from cognitive and developmental psychology (Shaughnessy, Garfield & Greer, 1996). A well-known example of such a framework is the SOLO (<u>S</u>tructure of <u>O</u>bserved <u>L</u>earning <u>O</u>utcome) Taxonomy developed by Biggs and Collis (1982, 1991), with its focus on the structure of observed learning outcomes. The SOLO Taxonomy comprises the following five criteria that are used to describe levels of increasing complexity in student's understanding of subjects or cognitive tasks:

1. *Prestructural:* The student engages in the task, but is distracted or mislead by an irrelevant aspect.

2. *Unistructural:* The student focuses on the relevant domain, but picks up only one aspect to work with.

3. *Multistructural:* The student picks up more and more relevant or correct features, but does not or is unable to integrate them.

4. *Relational:* The student now integrates several relevant or correct features with each other, so that the whole has a coherent structure and meaning.

5. *Extended* abstract: The student now generalises the structure to take in new and abstract features, representing a new and higher mode of operation. (Biggs & Collis, 1991, p. 65)

Another popular framework is that of Curcio (1981, 1987) who sought to identify the predictors of graph comprehension based on a schema theory model. She defined three levels of graph reading. These were:

i) **"Reading the data"**: extracting information explicitly stated in the data display, recognising graphical conventions, and making connections between context and data (lifting the data).

ii) **"Reading between the data"**: this involves using mathematical operations to combine or compare data (interpretation and integration of the data)

iii) **"Reading beyond the data":** this requires students to predict from the data by tapping their existing schema for information that is not explicitly stated in the data (extend, predict or infer).

Kimura (1999) argues that in order to make maximum use of the power of statistics, students have to learn how to derive new information from statistical information. Kimura (1999) proposed the following six levels of statistical ability, henceforth known as Kimura's hierarchy (shown in Table 1).

Table 1. Categories of Statistical Ability (Kimura, 1999)

Level	Features
Level A1	Basic reading of tables and graphs
Level A2	Reading key features from graphs
Level A3	Comparing information from two graphs
Level A4	Reading a simple trend in graphs
Level B	Knowing what constitutes an appropriate source of data for a given question
Level C	Statistical computation skills
Level D	Reading global trends in graphs
Level E	Extracting qualitative information from quantitative information
Level F	Creating new dimensional information

The order from A to F does not imply a strict sequence of development. Levels A, B, and C, for example, can be seen as different facets of the same ability, akin to the *graph sense* of Friel *et al.* (2001). Levels A, D, E, and F, however, do imply a developmental sequence according to Kimura, in relation to the interpretation of statistical information.

Aoyama and Stephens (2003) have interpreted Kimura's paper (originally published in Japanese) as follows:

> *Level A captures to a large extent the basic skills that students are taught in school mathematics in Japan and in other countries.*
>
> *Level A1 is about reading the title or theme of a graph, or naming its units, or locating particular values. Curcio's (1981, 1987) 'Read the Data' corresponds to Level A1.*
>
> *Level A2 is about reading maximum and minimum values in graphs, or differences between values, or finding ratios between several values.*
>
> *Level A3 refers to the ability to compare two features in a graph or graphs.*
>
> *Level A2 and Level A3 seems to correspond to what Curcio (1981, 1987) refers to as the ability to 'Read between the Data'.*

Level A4 involves reading a clear trend from data presented in a simple non-fluctuating graph.

[*Levels B and C* are not discussed and seem to be irrelevant with respect to (w.r.t.) graph comprehension]

Reading a trend from fluctuating phenomena, like stock prices, corresponds to *Level D*. Curcio's (1981, 1987) reference to 'Reading beyond the Data' appears to correspond most closely to this level. Level D is especially important in dealing with information presented as a time series. Basic questions need to be asked, such as, "Is there a trend?" and "What predictions can be made on the basis of the trend?"

Level E involves extracting qualitative information from quantitative information. Students are required to extract qualitative information, such as "the supply of tomatoes influences their price" or "the price of tomatoes becomes cheaper when they are in more plentiful supply," from the quantitative information. This information is not new, however, because it is directly implied by the information embedded in the graphical representation.

By contrast, Kimura's *Level F* involves creating dimensionally new information from, or imposing new information on, pre-existing qualitative and quantitative information. Two interpretations can be offered to elucidate Kimura's likely explanation of a trend shown in a graph. In this sense, it may refer to a capacity to provide an explanation that goes 'beyond the data'. In the other sense, that of imposing new information, it may involve asking questions about the quality or reliability of the data or the conclusions drawn from data, such as represented in opinion polls. Level F can be viewed as the signature feature of Kimura's levels. It is also the most difficult to interpret and analyse.

Level F, therefore, essentially involves the creation of a scenario to account for the data and is in contrast with Level E which simply involves an extraction of qualitative information from pre-existing quantitative data. The created scenario contains new information - Kimura's "new dimension" - in order to supply an explanatory context or set of conditions to the data. Similar to the ability to take a critical stance in evaluating statistical information, Kimura's Level F presupposes a capacity by the individual to step back from what is directly implied by the data, and to bring to that stance additional knowledge of other factors and of likely interpretative frameworks.

According to Aoyama and Stephens (2003), Kimura (1999) used a range of questions to carry out a pilot study among Japanese students, across Grades 4, 5, and 6, in order to show relationships among his categories. Tasks used in the pilot study were intended to test performance on Levels A, D, E, and F only (Levels B and C apparently being deemed irrelevant to *interpreting* graphs and essentially different facets of the same ability as Level A). Approximately 120 students were involved at each grade level. Kimura's findings indicated high levels of performance overall on tasks representing Level A. Performances diminished for Level D, with a range of 49%

considered to be operating at this Level in Grade 4 to 69% in Grade 6. For Level E, the range was 31% to 60%, and further diminished for Level F where only 35% in Grade 4 were considered at Level F and 41% in Grade 6.

Aoyama and Stephens (2003) identified shortcomings in Kimura's work and elected to build upon Kimura's pilot study, looking especially at Level F tasks, which are the most complex and difficult tasks for students. Kimura's questionnaire items did not require students to give reasons for their choices while the study carried out by Aoyama and Stephens (2003) required students to articulate and to justify their choices. A theoretical framework based on the levels of the SOLO Taxonomy (Biggs and Collis, 1982, 1991) was used to analyse students' performances more deeply and to verify achievement of Level F tasks. Aoyama and Stephens (2003) reported that the proportion of students giving Extended Abstract Level (or Level F) responses was considerably smaller than Kimura's data had suggested: if Level F responses are best characterised as Extended Abstract, using the SOLO Taxonomy, then it appears that few, if any, elementary grade students operate in this fashion.

Aoyama and Stephens (2003) conclude by advocating the administering of their questionnaire to older secondary students and to tertiary students with the goal of seeing how performances change with age and how "critical" statistical thinking develops.

2.7 Eye tracking – a powerful new tool in cognitive science

This, at-first-glance-irrelevant topic, has been included here in order to familiarise the reader with technology that will be used in phase 2 of this study.

Vision is an active, exploratory process that is an obvious requirement for both visual and graphical literacy. By using eye tracking devices that allows the researcher to record exactly where in a picture a person is looking at any given moment it is possible to study the way in which the gaze wanders over a picture, pauses and fixes on certain points (Pettersson, 1989). The location of each fixation in a picture filed influences how a picture is interpreted and later remembered. Eye movements can also be employed as an index of visual learning (Wolf, 1970) and can supply information on where, how long and how often subjects look at different parts of a picture.

Several basic types of eye movements have been defined in the literature. Eye movements that are applicable in eye gaze tracking have been summarized by Renshaw *et al.* (2003) as follows:

Fixation - A relatively stable position within some threshold of dispersion (typically 1 – 2 degrees) over some minimum duration (typically 100 – 200 ms), and with a velocity below some threshold (typically 15 – 100 degrees per second).

Saccade - Rapid ballistic movements of the eye from one point of interest to another, whose trajectory cannot be altered once begun. Saccades take between 30 and 120 ms and may cover between 1 – 40 degrees of visual angle.

Gaze - A series of consecutive fixations within an area of interest.

Scan path - A spatial arrangement of a sequence of fixations. The saccade-fixation-saccade sequence of eye movement defines a scan path during search and processing tasks.

Data related to the abovementioned eye movements that can be used to assess usability, according to Goldberg and Kortval (1998), are:

Number of fixations:- The number of fixations overall is thought to be negatively correlated with search efficiency.

Fixation duration:- Longer fixations are believed to be an indication of the difficulty a participant has in extracting information from a display.

Scan path:- This indicates areas of interest, cognitive load, and various search strategies. The optimal scan path for a task is a straight line to a desired target with a short duration fixation at the target.

Scan path length:- An unduly long scan path length could indicate non-meaningful representation or poor layout.

Visual information is typically extracted during a fixation when the eye stabilises the retina over an object of interest. No information is obtained during a saccade as there is insufficient time for visual feedback to guide the eye to its final position. (Tzanidou, 2003)

The idea of tracking human gaze and the subsequent study of eye movements is not a new domain in cognitive research. Eye tracking as a tool to gain insight into cognitive

processes by analysing eye movement patterns pre-dates the widespread use of computers by nearly a century (Cowen, 2001; Jacob & Karn, 2003; Milekic, 2003; Tobii User Manual, 2004; Ewing, 2005).

Eye tracking methods initially made use of invasive procedures requiring direct mechanical contact with the cornea (Jacob & Karn, 2003; Renshaw *et al.*, 2003) or the placement of a reflective white dot onto the eye (as well as tiny mirrors by way of a small suction cup) and electrodes positioned around the eye (also known as the electro-oculogram or EOG) (Milekic, 2003; Morimoto & Mimica, 2004). Later, less invasive but equally cumbersome techniques required the participant to wear head mounted equipment or glasses (Morimoto, Amir & Flickner, 2002). With the advent of the personal computer, high-speed data processing became possible, an innovation allowing the use of eye tracking data in real-time. An added complication was that the aforementioned methods precluded subjects who wore spectacles or contact lenses. The current underlying technology of all modern eye trackers is corneal reflection of infra-red light.

A breakthrough in eye tracking technology came with the development of the *Tobii 50 Series* of integrated eye trackers by *Tobii Technology*. With the *Tobii* eye trackers it is possible to achieve high quality tracking without interfering with the user environment as this technology is totally non-intrusive. The system utilises near-infra-red diodes to generate reflection patterns on the corneas of the eyes of the user. These reflection patterns, together with other visual information about the person, are collected by image sensors (an infra-red camera). Together with very sophisticated image analysis and mathematical algorithms a gaze point (where the user is looking) on the screen can be calculated. Unlike previous techniques, the *Tobii* eye trackers tolerate fairly large head movements, which provide an unrestrained user experience and can track the eye gaze of virtually anyone irrespective of visual acuity and corrective measures such as spectacles or contact lenses (Tobii User Manual, 2004).

Eye tracking has proved a valuable technological tool in our modern information-driven society and has found wide application. Disciplines where eye and gaze tracking technology are applied include psychophysics, psychology (notably psycholinguistics), ergonomics, market research, the medical field (neuroscience), web, television and advertising evaluation and human-computer interaction (HCI) (Hansen, Hansen, Nielsen, Johansen & Stegmann, 2002; Jacob & Karn, 2003; Schiessl, Duda, Thölke & Fischer, 2003; Tobii User Manual, 2004). Cognitive science is one area of research

where eye tracking has been extensively utilised, essentially based upon the assumption that eye movement patterns reflect the cognitive states of individuals (Cowen, 2001; Jacob & Karn, 2003; Qvarfordt & Zhai, 2005; Peebles & Chen, 2002).

The visual cortex is able to acquire vast amounts of visual information through the retinal signal transduction pathway housed in the eyes. A person's goals influence his/her eye gaze trajectory (Roy, Ghitza, Bartelma & Kehoe, 2004) and he/she will focus their eyes and also their attention (Holsonova, 2004) on areas that are most likely to be informative. Thus, the empirically validated assumption is made that data on visual behaviour can be used as a gauge of cognitive processes (Gulliver & Ghinea, 2004; Holsonova, 2004). An example of this approach was the study by Peebles & Chen (2002) of fine grained fixation patterns that reveal the effects of working memory limitations during the time course of problem solving.

Eye tracking technology (for detecting and collecting eye gaze data) was, until recently, unavailable to local researchers. During January 2006 the Department of Computer Science and Informatics, UFS, imported the latest in eye tracking technology from Sweden, the *Tobii 1750 Eye Tracker* developed by *Tobii Technology*.

The UFS-owned eye tracker was recently upgraded (June 2009) to a higher resolution model, the Tobii T120. The *Tobii Eye Tracker* utilises near-infra-red light-emitting diodes and infra red cameras integrated (at the bottom) into a 17" TFT (Thin Film Transistor) monitor to monitor and track the subject's eye movements in real time after calibration (Fig. 1).

Figure 1. The Tobii T120 Eye Tracker

Use of the *Tobii T120 Eye Tracker* together with the *Tobii Studio 1.58* eye tracking analysis and visualisation software allows integration of recorded eye gaze data with video from a video camera or web cam, microphone sounds, keystrokes, mouse clicks and other data streams in a single solution to give the researcher a holistic view of behavior. *Tobii Studio* provides tools for creating visualizations of eye gaze data, such as gaze plots, heatmaps and bee swarms.

The Area of Interest (AOI) tool makes it possible to aggregate and compare quantitative gaze data from large numbers of test participants. This allows a researcher to accurately calculate a number of fixation measures in AOI data from a given stimuli where those AOIs are defined (Tobii User Manual, 2006).

These measures, as they are mentioned in the Tobii user manual, are:

- fixation count: number of fixations in each of the AOIs;
- gaze time: total time of all fixations in each AOI;
- average fixation duration: average length of all fixations during all recordings in each AOI;
- time to first fixation: the time from the beginning of each recording until each AOI was first fixated upon, and;
- fixation order: the order in which each respective AOI was fixated upon.

Use of the above-mentioned data sources enables a researcher to gain insight into the test subject's cognitive processes (e.g. where they are searching for information and whether they are experiencing difficulties) when viewing the given graphical representations.

2.8 Summary

The foregoing literature review can be summarised as follows:

1. Presenting scientific data with graphs, diagrams and tables is a central practice in the natural and social sciences and is a basic skill of the scientist enabling fast and accurate data extraction.

2. Secondary and tertiary students in general have difficulty reading and interpreting graphs, are not always able to discern important issues illustrated by graphs and tables and little attention is given to develop this area of literacy.

3. Making sense of graphs appears to be a complex process consisting of three components (known collectively as graph sense): understanding the convention of a graph design; the decoding process of a graph; and the process of generalising or relating the information from the graph to the context of the situation.

4. Graphical literacy, a visual language or skill, refers to the ability to construct, produce, present, read and interpret charts, maps, graphs, and other visual presentations and graphical inscriptions.

5. Graph comprehension is affected by three factors: the visual characteristics of the graph, the viewers' prior knowledge and expectations about the graph, and; the viewers' prior knowledge about the information depicted in the graph.

6. Graph-based reasoning can also be influenced by working memory limitations and gender.

7. A number of hierarchies, or frameworks, identifying predictors of graph comprehension have been proposed. Two popular frameworks are those of Kimura and Curcio. Kimura's 'categories of statistical ability' consists of six levels of ability ranging from 'basic reading of graphs and tables' to 'creating dimensionally new information'. Curcio's framework defines three levels of graph reading: 'reading the data'; 'reading between the data'; and 'reading beyond the data'.

8. Eye tracking devices allow a researcher to record exactly where in a picture a person is looking at any given moment in order to study the way in which the gaze wanders over a picture, pauses and fixes on certain points. Eye tracking is a valuable tool that has been applied in a wide range of fields but has been extensively utilised in cognitive science essentially based upon the assumption that eye movement patterns reflect the cognitive states of individuals.

CHAPTER 3. METHODOLOGY

3.1 Research design and procedure

3.1.1 Research design

The research design that has guided the direction of the research effort in terms of formulation of research questions and hypotheses as well as the design of the research instrument is essentially *ex post facto*. This type of quasi-experimental design, where a particular characteristic of a certain group is investigated with the aim of identifying its antecedents "after the fact" is oriented to specific problems, supported by specific data, and given direction in by underlying hypotheses (Leedy, 1989, p228).

3.1.2 Procedure

It was decided to follow a mainly quantitative analytical descriptive (or survey) cross-sectional method (questionnaires and statistical analysis) mixed with some qualitative aspects (written responses) in the form of a graphical literacy test. Results obtained were analysed using Kimura's (1999) hierarchy of statistical understanding to verify achievement levels. An eye tracker was used to investigate selected student's gaze patterns with regards to scan paths and fixations during graph viewing in an attempt to identify possible detractors and sources of confusion as well as possible differences between students with low graphical literacy and students with high graphical literacy. It was also decided not to test graph encoding (or drawing) as this is a relatively low level task and more problems with graph interpretation seemed to be experienced by students.

With research questions 1 to 4 in mind it was decided that the project would be executed in two phases using two separate assessment instruments:

Phase 1: A graphical literacy test (#1) in the form of a questionnaire (on paper) with 16 tasks, each dealing with a different Cartesian $(x - y)$ graph, that was administered to the whole 3rd year biochemistry class of approximately 80 students in a class test situation. The results of this phase were used to answer research questions 1, 2 and 3 and to identify test subjects for phase 2.

Phase 2: A graphical literacy test (#2), in the form of a slide show, taken in conjunction with eye gaze data using students identified from the results of Phase 1 as scoring in the bottom 10% of the class and students identified as scoring in the top

10% of the class w.r.t. their overall graphical literacy in biochemistry. The results of the answers and eye gaze data was used to answer research question 4.

3.1.3 Time dimension

Questionnaire #1:   as a group, approximately 40 minutes, administered during a BOC334 (Protein module) practical session scheduled for the evening of Monday 21 May 2009 in the PC laboratory, UFS.

Questionnaire #2:   individual, approximately 15 minutes each, administered in the usability laboratory, Department of Computer Science and Informatics, UFS, during the week of 19-23 October 2009.

3.2 Population of the study

Because of their previous exposure to biochemistry topics and concepts where graphs are often used to illustrate certain important facets of, for example, protein denaturation, enzyme kinetics, energetics or protein purification, it was decided to use all of the biochemistry students in their third year of tertiary study (i.e. senior undergraduates) at the UFS as the population for this study. In addition, these students were the group of students identified as experiencing problems with graph encoding (drawing) and decoding (interpretation) during their second year of study (the previous year – 2008) and the first semester of their 3rd year (2009).

3.3 Sample and Sampling Technique

3.3.1 Design

Whole population: the entire 3rd year biochemistry class registered at the University of the FreeState. This can be regarded as a defined generalized group.

3.3.2 Units of analyses

The units of analyses differed for the two phases differed as follows:

Phase 1 (Quiz):   Intact, whole class of students (82 in number), exhibiting a range of abilities.

Phase 2 (Eye tracking):   Individuals (top 10 and bottom 10)

3.3.3 Demographics

The demographics of the test samples are summarised in Tables 2 and 3.

Table 2. Demographics of test sample for phase 1.

Ethnicity	Gender		Totals
	Male	Female	
White	8	9	17
Black	27	38*	65
Totals	35	47	82

*One 'coloured' female student and one asian student were absent on the day of the study.

Table 3. Demographics of test sample for phase 2.

Ethnicity	Gender per group				Totals
	Male		Female		
	Top 10 grp	Bottom 10 grp	Top 10 grp	Bottom 10 grp	
White	3	0	4	0	7
Black	3	1	0	9	13
Totals	6	1	4	9	20

3.4 Measurement instruments

3.4.1 Test development

Two graphical literacy tests to examine the students' level of graphical literacy according to the hierarchy of Kimura (1999) were developed. One test was in the form of a questionnaire administered to the whole class (Questionnaire #1) and one test was in the form of a slide show used in conjunction with the eye tracker (Questionnaire #2). Questionnaire #1 was administered to a group of four B.Sc. Honours students in a pilot study with the aim of identifying sources of confusion w.r.t. question formulation and graph design of the tasks. Results were not analysed statistically and the students provided verbal feedback at the end of the test. This feedback resulted in a few questions being reformulated and one graph being changed from a line graph to a

bar graph, thus ensuring clarity. This pilot study also provided an indication of the time required to complete the test: 60 minutes.

The type of questionnaire used (the final, corrected version) was a graphical literacy test incorporating examples of common graph decoding tasks (eliciting the value of one variable corresponding to a given value of another) as well as a selection of biochemistry concepts (dependence of enzyme activity on pH, temperature, substrate and enzyme concentration, enzyme allostery, product formation and substrate depletion of an enzyme-catalysed reaction, buffer action, chromatography and energetics) in Cartesian graphs and designed to assess the following levels of Kimura's hierarchy: Levels A1-A4 and D-F. Levels B and C and were not applicable as the students were not required to draw graphs (the aim was to assess the higher level cognitive skills of decoding graphs). In each case, the graphs chosen were designed so as to exclude any extraneous effects of graph design (see section 2.5) on graph comprehension.

Questions were formulated in simple English so that all three items of information (the given variable, the given value, and the required variable) were clear and unambiguous.

All graphs were drawn using Microsoft Excell and imported into the questionnaire set up in Microsoft Word (for Questionnaire #1) or the slide show set up in Tobii Studio (Questionnaire #2).

The process followed during the design and application of the graphical literacy test is summarised in Figure 2.
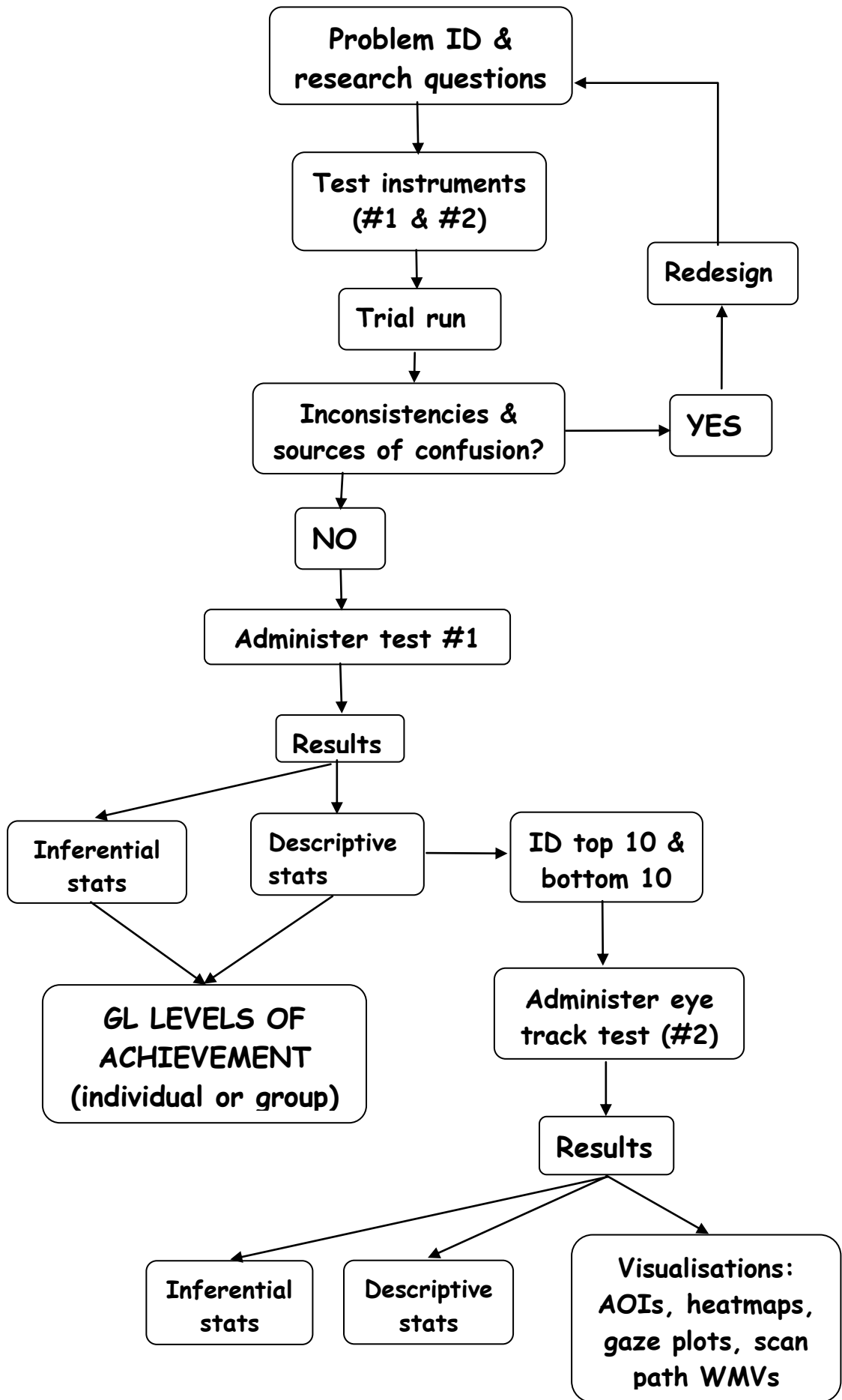
Figure 2. Flowchart of graphical literacy test development and application.

Reliability of the test

Internal consistency (reliability) of the instrument was established using Cronbach's alpha (Cronbach, 1951). This was calculated with the aid of a Microsoft Excel spreadsheet (obtained from Del Siegle: dsiegle@uconn.edu) using the results obtained with questionnaire #1 and resulted in a Cronbach's alpha value of 0,797. An internal consistency coefficient (alpha) of 0,7 or higher is generally accepted as an indicator of acceptable reliability (based upon the rules of thumb by George and Mallery, 2003, where "_ > 0.9 – Excellent, _ > 0.8 – Good, _ > 0.7 – Acceptable, _ > 0.6 – Questionable, _ > 0.5 – Poor, and _ < 0.5 – Unacceptable", p. 231) and an alpha of 0,8 is a reasonable goal (see also Simon, 2008, for an insightful discussion). The graphical literacy test (Questionnaire #1) developed by the researcher can, therefore, be considered to be reliable, i.e. it can be considered a fairly precise measurement of graphical literacy (scores on similar items are related or are internally consistent). However, the greater the number of test items, the closer alpha tends towards the ideal even when the set of items measures several unrelated latent variables (Cortina, 1993), a potential shortcoming. A more appropriate index of the extent to which all of the items in a test measure the same latent variable may be the hierarchical "Coefficient omega" (McDonald, 1999). Unfortunately, this researcher was unable to source an MS Excel worksheet to calculate this coefficient and elected to establish alpha using existing software.

3.4.2 Synopsis of test tasks

A synopsis of the individual test tasks comprising each questionnaire follows.

3.4.2.1   Questionnaire #1

Questionnaire #1 was composed of 16 graphs (16 tasks) with a total of 43 questions consisting of 16 multiple choice items, 24 single answer items, three descriptive items and eight items where an explanation for the given answer was required (students were required to articulate and justify their answers). For scoring purposes, single answer items scored one point each with an additional point being scored for the correct explanation. A 100% score would be obtained with 52 points. Some items requiring a description of more than one feature on the graph scored two points.

The questionnaire was administered to the whole class in a test situation. A detailed synopsis of each task and correct responses follows (**NOTE**: the complete test has been attached as **APPENDIX A**):

Task 1 is a warm-up type of task and is intended to examine students' graphical literacy at the lowest level (A1 - locating values; read the data). The correct responses are 1.1: Point A; 1.2: Point B; 1.3: Points A & C. This task was chosen as it is so widely performed and because the procedures involved are relatively simple: elicit the value of one variable which corresponds to a given value of another variable.

Task 2, identifying features on a pH optimum graph, examines students' responses at two levels (A2 and A3) and assesses whether they can read between the data. Correct responses and levels tested are as follows:

2.1	Answer:	1 mM/min	(A2 – locate maximum values)
2.2	Answer:	8,5 mM/min	(A2 – read between the data)
2.3	Answer:	7 & 10	(A2 – find ratios between values)
2.4	Answer:	Velocity	(A3 – compare two features on a graph – the axes)
2.5	Answer:	half	(A2 – find ratios)

The **third task**, identifying features on a substrate-dependence graph, examines students' responses at three levels (A1, A2 and A3) and analyses whether they can read the data or between the data. Correct responses and levels tested are as follows:

3.1	Answer:	c (mM)	(A1 – locate axis, name units)
3.2	Answer:	70 U/ml	(A2 – find ratio, read off y-axis)
3.3	Answer:	50 U/ml	(A3 – comparing features; students must correctly note that even if [S] changes, ½Vmax is a fixed feature of the graph and will not change)

Task 4, identifying features on an oxygen binding curve, examines students' responses at two levels (A3 and E) and analyses for the first time whether they can read beyond the data. Correct responses and levels tested are as follows:

4.1	Answer:	pO_2	(A3 – compare two features on a graph: the dependent and independent axes)
4.2	Answer:	Θ	(A3 – compare two features on a graph: the dependent and independent axes)
4.3	Answer:	c (decreases)	(E – extracting qualitative information from quantitative information and requires deeper thinking). It is clear that the as the graph shifts to the right, so the affinity decreases – less oxygen binds at the same partial pressure of oxygen).

Task 5, identifying features on a progress curve, examines students' responses at two levels (A3 and D) and also assesses the presence of the classic slope/height confusion exhibited by physics students interpreting kinematics graphs. To correctly answer this task students must realise that they must look at the slope (velocity in terms of ΔAbs/sec) and not at the height of the graphs. Correct responses and levels tested are as follows:

5.1   Answer: greater   (A3 – compare two features on a graph: in this case the slopes.) If students answered "velocity of A less than B" because "absorbance value of B at 20 sec is higher than that of A" then they looked at the height of the graph and not the slope.

5.2   Answer: No   (D – reading a trend: slopes are not the same.) If students answered "Yes, at 40 secs" because "the graphs intersect and both have the same absorbance value" then they are confusing height and slope again. This task can also be classified as "reading beyond the data".

The **sixth task**, identifying features on a concentration versus time graph, examines students' responses at three levels (A2 and D) and also requires that they note the slope of the graphs. Correct responses and levels tested are as follows:

6.1   Answer: a (0-40 secs)   (A2 – identify steepest slope)

6.2   Answer: c   (E – extracting qualitative information from quantitative information)

6.3   Answer: Yes   (D – prediction of trend.) Explanation: it is clear, based on the negative slope of the graph, that the substrate will eventually reach a concentration of zero.

6.4   Answer: Will stop increasing   (D – prediction of trend). Explanation: it is clear, based on the trend of the graph, that the product formation is dependent on the substrate consumption and when the substrate is depleted the product concentration will reach a maximum and stop increasing.

Task 7, requiring students to match the information in a narrative passage to a graphical representation, examines students' responses at Kimura's level A4 but also requires that they refer both to the three graphs and the problem statement in order to choose the correct graph. The correct response is 'b' (graph 2 - a clear trend) and is a Level A4 task –: identify a graph with a direct relationship between [E] and rate, i.e. a straight line.

Task 8, identifying features on a temperature-dependence graph, examines students' responses at three levels (A2, D and E) and analyses whether they can read the data, read between the data or read beyond the data. Correct responses and levels tested are as follows:

8.1   Answer: d (45°C)   (A2 - identify max y-value, read off x-axis)
8.2   Answer: c   (E – extracting qualitative information from
                                quantitative information)
8.3   Answer: b (close to 0) (D - prediction made on the basis of a trend:
                                eventually no activity will be detectable)

Task 9, identifying features on a titration graph, examines students' responses at two levels (A3 and D) and analyses whether they can compare slopes at different points on the fluctuating graph. This also analyses whether they can read between the data and beyond the data. Correct responses and levels tested are as follows:

9.1   Answer: c   (A3 - compare two features on a graph: the
                                dependent and independent axes)
9.2   Answer: D and F followed by B   (D – reading a trend. These points
                                have the lowest slope and correspond to a small
                                increase in pH with addition of OH⁻ If the students
                                answer 'A', the lowest point on the graph, then they
                                have confused the slope of the graph with the height
                                of the graph)
9.3   Answer: C and E   (D – reading a trend. These points have the
                                steepest slope and correspond to a large increase in
                                pH with addition of OH⁻. If the students answer 'G',
                                the highest point on the graph, then they have
                                confused the slope of the graph with the height of
                                the graph)

The **tenth task** requires students to match the quantitative information in a pH vs time graph with a semi-logarithmic x-axis scale to a different graphical representation, examines students' responses at Kimura's level A4 but also requires that they be able to recognise different graphing practices with regard to scale in order to choose the correct graph. The correct response is "a" (graph 1). This is a Level A4 task (identifying a clear trend). If the students choose c as their answer then they are confused with respect to scale.

Task 11, identifying features on a column graph showing reaction velocity as a function of substrate type, examines students' responses at Kimura's level A2 but also analyses for the first time the extent of 'zoom graph confusion'. Correct responses and level identifiers are as follows:

11.1   Answer: d (none)   (A2 – finding ratios between several values.) If the
                                students answer anything else then they have not
                                noted that the y-axis does not commence at zero – a
                                zoom effect – and are only comparing the portion of
                                the column visible in the graph displayed
11.2   Answer: c (10%)   (A2 – finding ratios between several values.) If the
                                students answer anything else then they have not
                                noted that the y-axis does not commence at zero – a
                                zoom effect – and are only comparing the portion of
                                the column visible in the graph displayed. Since the

velocity obtained with fructose is 11 U/ml and that obtained with glucose is 10 U/ml it follows that the difference (1 U/ml) is 1/10 or 10% and the velocity obtained with fructose is 10% higher than that obtained with glucose.

Task 12, identifying features on an elution profile, examines students' responses at four levels (A1, A2, A3 and E). A second y-axis (which does not commence at zero) is introduced for the first time and students must be able to discriminate between the two graphs, a potential source of confusion. Correct responses and level identifiers are as follows:

12.1 Answer: c (2.5-5 mM) (A3 - compare two features on a graph: the start and the end of a gradient)
12.2 Answer: tube 9 (A2 – locating a maximum value)
12.3 Answer: 3.5 mM (A1 – locating a particular value. Must be able to locate points on the x- and 2nd y-axis)
12.4 Answer: a (E – extracting qualitative information from quantitative information)

Task 13, identifying features on a double reciprocal plot with two graphs, examines students' responses at three levels (A2, A3 and A4). The task is complicated, however, by the reciprocal nature of the data. The final answer, therefore, requires a simple calculation (inversion and negative inversion) to answer the question. Correct responses and level identifiers are as follows:

13.1 Answer: Vmax = 1/0.1 = 10 (A2 – finding a ratio between values)
13.2 Answer: B (A3 – comparing information from two graphs). Explanation for answer: because of the reciprocal nature of the data, the lowest – not the highest - graph has a higher velocity at all substrate concentrations.
13.3 Answer: B (A3 – comparing information from two graphs). Explanation for answer: because of the reciprocal nature of the data, the highest value on the x-axis represents the lower Km value.
13.4 Answer: No (A4 – reading a simple trend). Explanation for answer: the graphs don't intersect in the positive quadrant – only in the negative quadrant – which means that enzyme B has higher velocities at all substrate concentrations.

The **fourteenth task**, identifying features on a free energy versus reaction progress graph, examines students' responses at one level (D) but is complicated by the fluctuating nature of the slope of the graph which can be represented in a different format. The task requires students to match the change in slope (ΔG) of each line segment (a-b, b-c, c-d, d-e) to a different graphical representation in order to choose the correct graph. The correct response is "Graph 3" (D – reading a trend from fluctuating phenomena). Explanation for answer: Graph 3 shows that the change in G

(ΔG) between a and b is zero, between b and c positive, between c and d negative and between d and e zero.

Task 15, identifying features on a pH optimum graph, examines students' responses at three levels (A3, E and F). This task also assesses for the first time whether the students can think critically and posit a likely explanation for the presence of two pH optimum peaks in the graph. Correct responses and level identifiers are as follows:

15.1   Answer: c (10 µM/min)   (A3 – compare two features on a graph). Any other answer indicates confusion with the y-axis scale.

15.2   Answer: The students should give a description along the lines of "An increase in velocity with an increase in pH but after decreasing around pH 8 another increase is observed resulting in two peaks". This is a Level E task (extracting qualitative information from quantitative information).

15.3   Answer: "Two peaks indicates the presence of two enzymes, each with their own pH optimum". To achieve this level F task (creating dimensionally new information from, or imposing new information on, pre-existing qualitative and quantitative information. May refer to a capacity to provide an explanation that goes 'beyond the data') students need to read and understand all the velocity and pH information presented graphically, and then draw on their own knowledge related to enzymology, and finally supply a suitable scenario to explain all information from the graph.

The **sixteenth task**, the most complex and difficult so far, examines students' responses at level F. The bar graph in this task shows an initial decrease in reaction velocity with assay number followed by an increase in reaction velocity. Although no specific data regarding temperature is supplied, the possible influence of temperature is hinted at in the description of the experiment ('carried out in a thermostatically controlled spectrophotometer'). Students are required to offer a likely explanation for the presence of the dip in velocity. There are two likely scenarios for the observed results. Because a thermostatically-controlled cuvette holder allows fairly rapid temperature adjustments this means that the researcher could have decreased the temperature incrementally to a minimum for assays 1 to 10, causing a decrease in enzyme activity, and then increased the temperature to the original temperature setting for assays 11-20. Alternatively, the researcher could have increased the temperature incrementally up to a maximum for assays 1-10, causing thermal denaturation and loss of activity, and then decreased it to the initial temperature setting. Other scenarios are possible, for example addition of inhibitors, activators, removal of essential metal ions etc, but because mention is made of temperature, I am specifically looking for an explanation involving the effect of temperature on enzyme activity and how a fluctuating temperature environment could be brought about. To

achieve this level F task, students need to read and understand all the velocity and time information presented graphically, take note of the thermostatically-controlled cuvette, and then draw on their own knowledge related to enzymology and the effect of temperature on enzyme activity, and finally supply a suitable scenario to explain how the data was obtained.

A summary of task level identifiers according to Kimura as well as the corresponding descriptors of Curcio are tabulated in Table 4.

Table 4. Summary of task levels in Questionnaire #1

TASK	QU	Read the data	Read between the data		Reading beyond the data			
		Kimura's levels						
		A1	A2	A3	A4	D	E	F
1	1.1	X						
	1.2	X						
	1.3	X						
2	2.1		X					
	2.2		X					
	2.3		X					
	2.4			X				
	2.5		X					
3	3.1	X						
	3.2		X					
	3.3			X				
4	4.1			X				
	4.2			X				
	4.3						X	
5	5.1			X				
	5.2					X		
6	6.1		X					
	6.2						X	
	6.3					X		
	6.4					X		
7					X			
8	8.1		X					
	8.2						X	
	8.3					X		
9	9.1			X				
	9.2					X		
	9.3					X		
10					X			
11	11.1		X					
	11.2		X					
12	12.1			X				
	12.2		X					
	12.3	X						
	12.4						X	
13	13.1		X					
	13.2			X				
	13.3			X				
	13.4				X			
14						X		
15	16.1			X				
	16.2						X	
	16.3							X
16								X
TOTALS: Kimura		5	11	10	3	7	5	2
TOTALS: Curcio		5	21		17			

This quiz was presented as a slide show (set up in *Tobii Studio 1.58* with graphs imported as JPEG files) with each task presented separately, i.e. the graph is repeated with each question.   The questions were answered verbally while the student was seated in front of the *Tobii T120 EyeTracker*.

Questionnaire #2 comprised three graphs (three tasks) with a total of 12 questions consisting of 8 single answer items (two of which required students to articulate and justify their answers) and 4 descriptive items.   For scoring purposes, single answer items score one point with an additional point being scored for the correct explanation. A 100% score would be obtained with 18 points.   Some items requiring a description of more than one feature on the graph scored two points.

A synopsis of each task and correct responses follows (score = ✔):

Task 1, an expanded version of Task 8 in Questionnaire #1 (identifying features on a temperature-dependence graph), examines students' responses at six levels (A1, A2, A3, D, E and F) and analyses whether they can read the data, read between the data and read beyond the data.   An example of one of the six slides for this task (question 1.1) is shown in Figure 3 with the six questions listed separately (to save space) in Figure 4.   Correct responses, score allocated per correct response and level identifiers are as follows:

1.1   Answer: 20 μM/min ✔ (A1 - locating values; read the data)
1.2   Answer: 90 μM/min ✔ (A2 – identify max value, read off y-axis)
1.3   Answer: Velocity axis ✔ (A3 – compare two features on a graph: the dependent and independent axes)
1.4   Answer: The following description is acceptable: "The velocity of an enzyme-catalysed reaction increases initially to a maximum and then decreases with an increase in temperature" ✔ OR "shows a peak with lower velocity on either side of the temperature optimum".   This is a Level E task (extracting qualitative data from quantitative data)
1.5   Answer: "Velocity will reach or be close to zero" OR "eventually no activity will be detectable" ✔ (D - prediction made on the basis of a trend)
1.6   Answer: Most likely explanation is that the reaction velocity increases with an increase in temperature ✔ (a feature of thermodynamics) until a point of maximum velocity after which any increase in temperature causes the enzyme protein to denature ✔ (thermal denaturation) and lose activity.   This can be seen as a tapering off in velocity above 45°C.   To achieve this level F task, students

need to read and understand all the velocity and temperature information presented graphically, and then draw on their own knowledge related to enzymology and the effect of temperature on enzyme activity, to finally supply a suitable explanation for the observed trend in the graph.
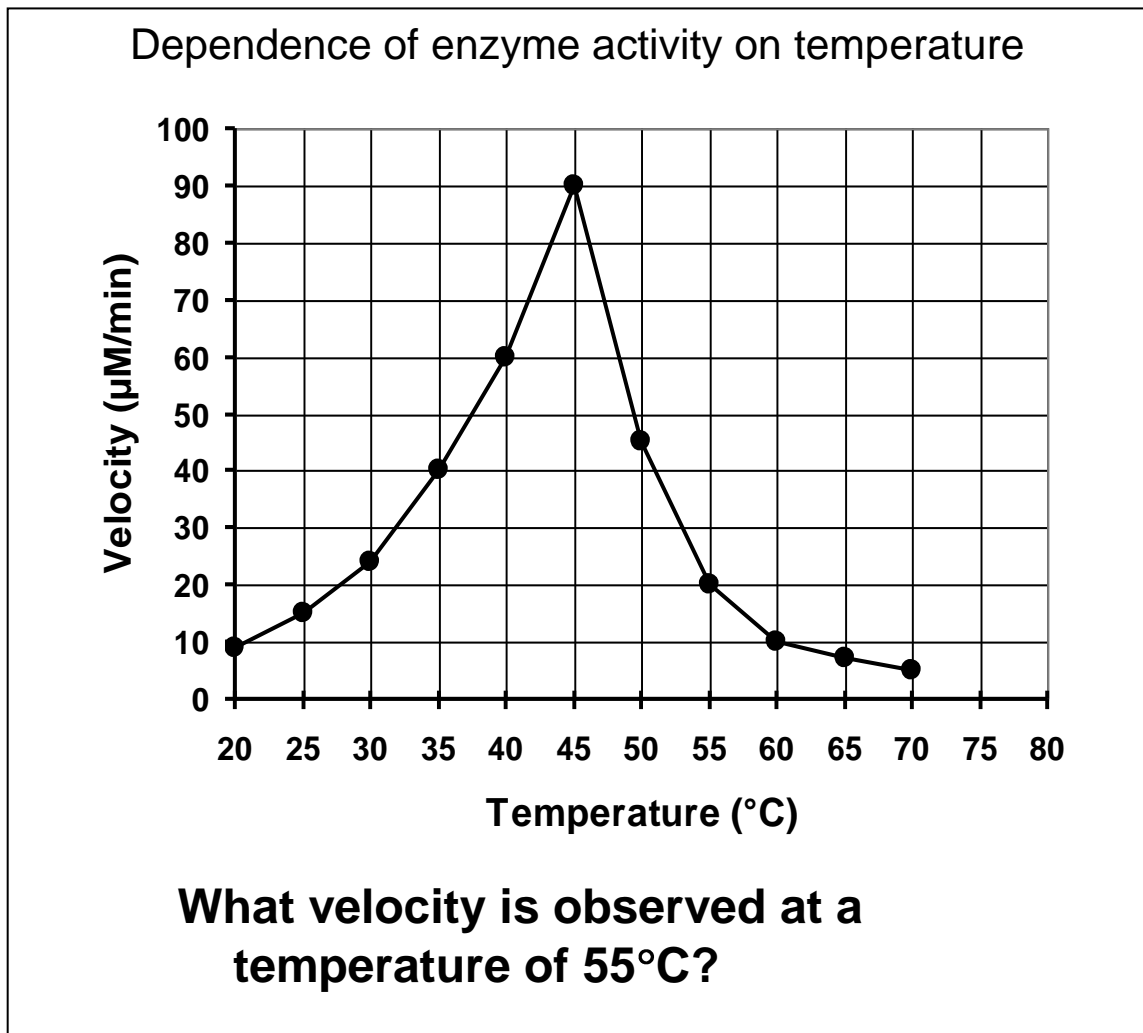

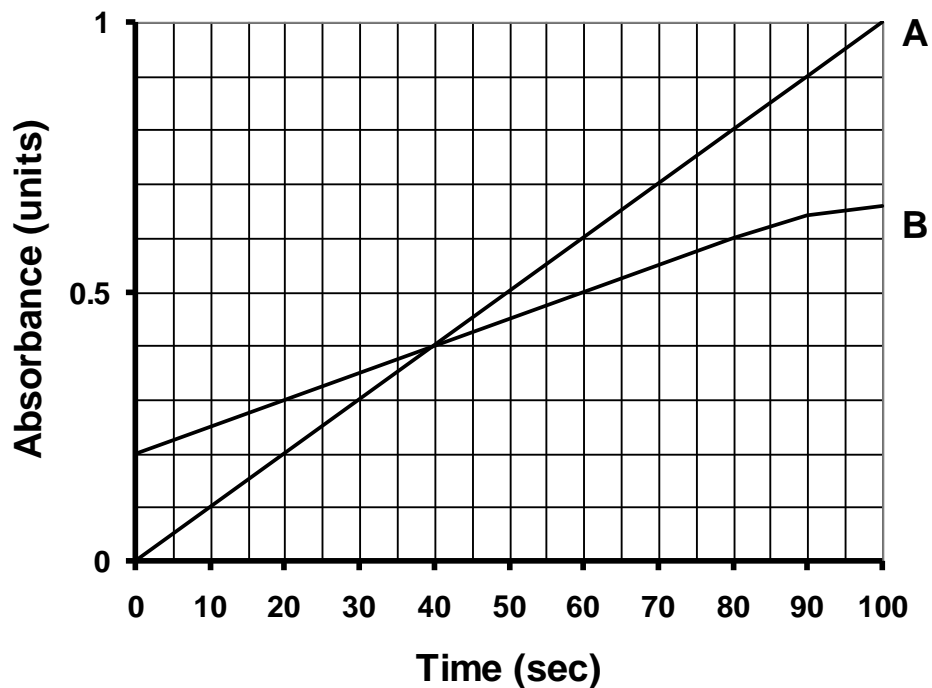*Figure 3.* Sample slide from eye track test Task 1 showing question 1.1.

1.1	What velocity is observed at a temperature of 55°C?
1.2	What is the maximum reaction velocity?
1.3	Which axis is the dependent axis?
1.4	How would you describe the trends in the graph?
1.5	What do you think will happen at a temperature of 90°C?
1.6	Why do you think the graph reaches a maximum and then decreases to a minimum?

Figure 4. Questions used on the six slides comprising eye track test **Task 1**.

Task 2 (which is Task 5 from Quiz #1), identifying features on a progress curve, examines students' responses at two levels (A3 and D) and also assesses the classic slope/height confusion exhibited by physics students interpreting kinematics graphs. To correctly answer this task students must realise that they must look at the slope (velocity in terms of ΔAbs/sec) and not at the height of the graphs. This can be confirmed using the eye tracking data. An example of one of the two slides for this task (question 2.1) is shown in Figure 5 with the two questions listed separately in Figure 6. Correct responses and level identifiers are as follows:

2.1	Answer: greater ✔ *Why: slope of A > B* ✔	(A3 – compare two features on a graph: in this case the slopes). If students answered "velocity of A less than B" because "absorbance value of B at 20 secs is higher than that of A" then they looked at the height of the graph and not the slope.
2.2	Answer: No ✔ *Why: slopes are* *not the same.* ✔	(D – reading a trend: slopes are not the same). If students answered "Yes, at 40 sec" because "the graphs intersect and both have the same absorbance value" then they are confusing height and slope again. This task can also be classified as 'reading beyond the data'.

Absorbance vs time for two different enzyme catalysed reactions

At instant t = 20 secs, is the velocity (measured as change in absorbance per second or ∆Abs/sec) of enzyme A greater than, less than, or equal to that of enzyme B?

Why?

Figure 5. Sample slide from eye track test Task 2 showing question 2.1.

2.1   At the instant t = 20 sec, is the velocity (measured as change in absorbance per second or ∆Abs/sec) of enzyme A greater than, less than, or equal to that of enzyme B?

Why?

2.2   In the time frame 0 to 100 seconds, do enzymes A and B ever have the same velocity, and if so, after what time interval?

Why?

Figure 6. Questions used on the two slides comprising eye track test **Task 2**.

Task 3, identifying features on a concentration versus time graph with three sets of data, examines students' responses at four levels (A3, A4, E and F) and analyses whether they can read between the data and read beyond the data. An example of one of the four slides for this task (question 3.1) is shown in Figure 7 with the four

questions listed separately in Figure 8. Correct responses and level identifiers are as follows:

3.1   Answer: sugar a ✓   (A4 – reading a simple trend)
3.2   Answer: 14 min ✓   (A3 – compare information from two graphs)
3.3   Answer: Most likely description is *"As sugar a is being degraded* ✓ *in the reactor by an enzyme, sugars b and c are being released* "✓ or "*As the concentration of sugar a decreases with time, the concentrations of b and c increase*"   (the sum of the final concentrations of b and c appear to equal the initial concentration of sugar a).
        To achieve this level E task (extracting qualitative data from quantitative data), students need to read and understand all the concentration and time information presented graphically and then critically evaluate the relationships between the data sets.
3.4   Answer: Most likely conclusion is *"Sugar a is composed of sugar b and sugar c* ✓ *with twice as much b as c."* ✓
        To achieve this level F task, students need to read and understand all the concentration and time information presented graphically, critically evaluate the relationships between the data sets, and then draw on their own knowledge related to biochemistry to finally supply a suitable explanation for the observed trend in the graph.
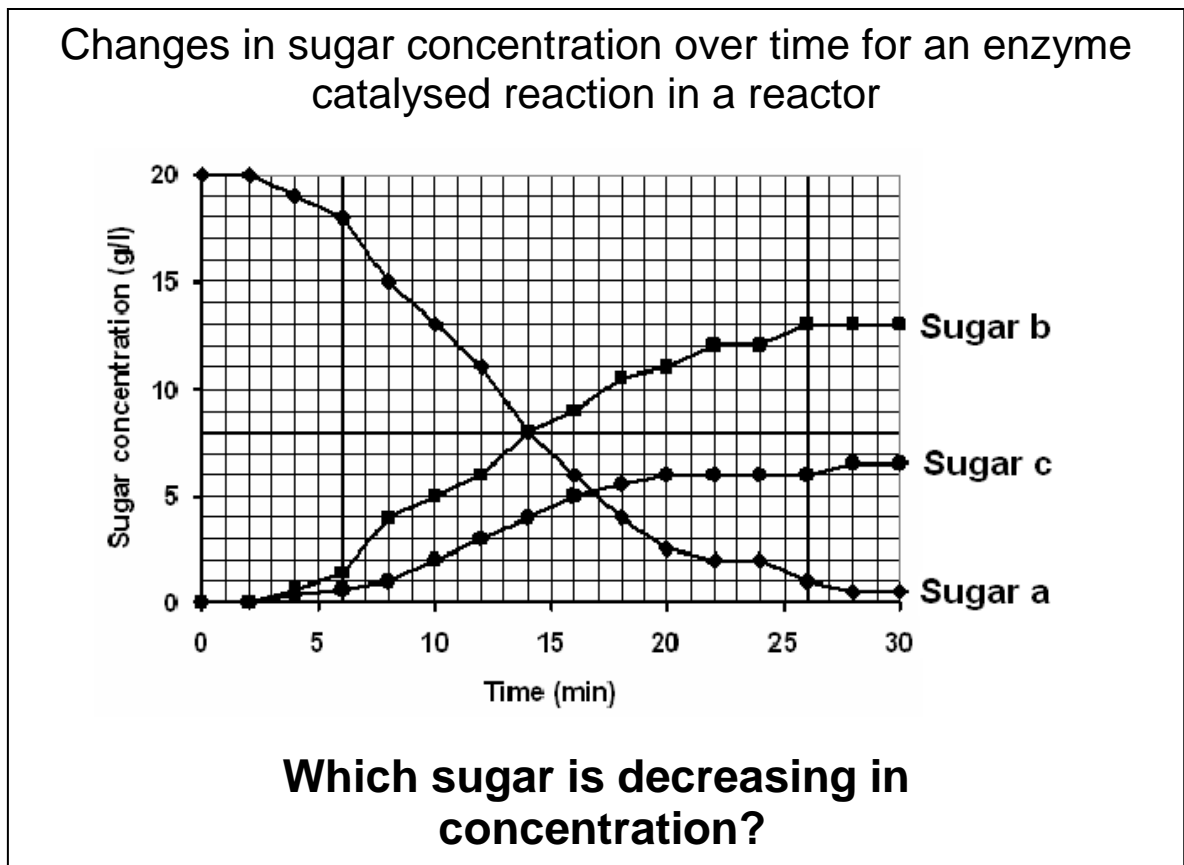

Figure 7. Sample slide from eye track test Task 3 showing question 3.1.

3.1	Which sugar is decreasing in concentration?
3.2	After what time interval are the concentrations of sugar a and sugar b equal?
3.3	Describe the observed results.
3.4	If possible, describe the composition of sugar a in detail.

Figure 8. Questions used on the four slides comprising eye track test **Task 3**.

A summary of task level identifiers according to Kimura as well as the corresponding descriptors of Curcio for Questionnaire #2 are tabulated in Table 5.

Table 5. Summary of task levels in Questionnaire #2

TASK	QU	Curcio's levels						
		Read the data	Read between the data	Reading beyond the data				
		Kimura's levels						
		A1	A2	A3	A4	D	E	F
1	1.1	X						
	1.2		X					
	1.3			X				
	1.4						X	
	1.5					X		
	1.6							X
2	2.1			X				
	2.2					X		
3	3.1				X			
	3.2			X				
	3.3						X	
	3.4							X
TOTALS: Kimura		1	1	3	1	2	2	2
TOTALS: Curcio		1	4		7			

3.5 Data collection procedures

Phase 1:

Questionnaire #1 was administered to the whole class in a test situation. Students were asked to complete 16 tasks related to the interpretation of graphical data. Questionnaires were returned to the researcher and evaluated as per section 3.4.2.1.

Phase 2:

The eye tracker employed in this phase of the project was a Tobii T120 Eye Tracker (see section 2.7 for description) with a 60 Hz data sampling rate made available by the Department of Computer Science and Informatics, UFS. Participants were seated approximately 70 cm from the display.

Test subjects were subjected individually to a sequence of electronic stimuli in the form of a slide show consisting of graphic representations linked to questions (questionnaire #2) shown to them on the 17" TFT (thin-film transistor technology) display of the Tobii T120 Eye Tracker using Tobii Studio 1.58 software.

For each trial, a graph would be presented with a question below it on the eye tracker display (set up in *Tobii Studio 1.58* with graphs imported as JPEG files) calibrated and set up with the aid of Prof. P. Blignaut, Department of Computer Science and Informatics, UFS. Participants were instructed to answer the question appearing on each slide verbally as soon as they had an answer. The researcher recorded the participant's responses manually as well as digitally using a voice recorder (Olympus VN-6500PC, 220h recording time in WMA format). By pushing the spacebar on the keyboard the researcher could load the next graph and question slide on the display. Response times were automatically recorded between spacebar strokes. By using a remote viewer the researcher could follow and control the collection of eye gaze data in real time. This method was devised so that participants in the eye tracking study would not have to take their eyes away from the display in order to enter answers, as would be the case if using the keyboard or writing. Before starting the experiment, participants were requested to answer the questions as rapidly and as accurately as possible and to only ask for guidance if they really did not understand the question. The participants' verbal responses (digitally recorded) were then transcribed by the author and evaluated as per section 3.4.2.2. A WMV file with eye movement scan paths plus an integrated video of the participant was also obtained for each participant.

3.6 Methods of data analysis

Phase 1:

Descriptive statistics:

Arithmetic means (hereafter referred to simply as 'mean'), standard deviations and frequency distributions were calculated using Microsoft Excel functions. Graphs were

drawn using the Microsoft Excel Chart Wizard. Scores obtained from the graphical literacy test (questionnaire #1) were analysed according to the hierarchy of Kimura (1999) for each level.

The overall score (in the form of a percentage) obtained for the graphical literacy test by the students was correlated, using bivariate analysis, with the average results obtained by them for the six biochemistry modules that they have completed to date. Pearson's r was computed and correlation noted as follows: $0,1 \leq r < 0,3$ is weak; $0,3 \leq r < 0,5$ is moderate and $r > 0,5$ is strong. This data will be used to confirm or reject the hypothesis that students' graphical literacy level has an impact on their academic performance (Research Question 2). This analysis was repeated in group context (gender and ethnicity).

Inferential statistics:

For comparison between the different groups of students (black vs white; male vs female, etc) a two-sample student's t-Test (Two-Sample Assuming Unequal Variances) was performed using the data module of Microsoft Excel.

Phase 2:

Means and standard deviations were calculated using Microsoft Excel functions. All graphs were drawn using the Microsoft Excel Chart Wizard. Scores obtained from the graphical literacy test (questionnaire #2) were analysed according to the hierarchy of Kimura (1999) for each level.

For comparison between the two different groups of students (top 10 vs bottom 10) a two-sample student's t-Test (Two-Sample Assuming Unequal Variances) was performed using the data module of Microsoft Excel.

Tobii Studio™ 1.58 analysis software was used for the recording and analysis of eye movements (e.g. fixation duration) in an attempt to identify possible sources of confusion on the graphs while students were answering questionnaire #2 as well as investigating any possible differences between the two groups of students (top 10 vs bottom 10). More specifically, heatmaps were used to visualise differences in fixation length on different areas of the graph (maximum duration in any area is known as a hotspot) while fixation duration was used as a quantitative measure in AOI's. The time taken per question (response time) as well as the overall time to complete the test were also used to compare the two groups.

CHAPTER 4. DATA ANALYSIS, RESULTS AND DISCUSSION

4.1 Research question 1

The question: *"What are the graphical literacy levels of undergraduate biochemistry students at the UFS?"* was answered by analysing the individual task responses, the overall test mean score, test data in terms of percent correct responses analysed as per the descriptors of Kimura and Curcio, respectively, and, finally, in terms of individual achievements at Kimura's Levels and Curcio's descriptors.

Test mean

Overall, the biochemistry students displayed very poor graphical literacy capabilities with a test mean score of 48,3% ($\pm$12,1) being obtained. This is disturbingly low considering that the test subjects are senior B.Sc. students in their final year of study and majoring in biochemistry. The percent correct responses per question recorded by students for Questionnaire #1 is shown in **Appendix B**. The lowest score was 19,2% and the highest score was only 78,8%. However, testing theory suggests that "ideal" tests have a mean near 50%, maximising the spread of scores (Beichner, 1994), an observation confirmed by the frequency analysis result shown in Fig. 9. In addition, the relatively small standard error of the mean (25% of the mean) indicates that there is not much uncertainty in the 48,3% mean score value. The graphical literacy test designed by the researcher can, therefore, be regarded as being close to the ideal. The low test mean score obtained is consistent with previous findings by, for example, Lake (1999) and Knuth (2000), who reported that secondary and tertiary students in general experience difficulty reading and interpreting graphs.
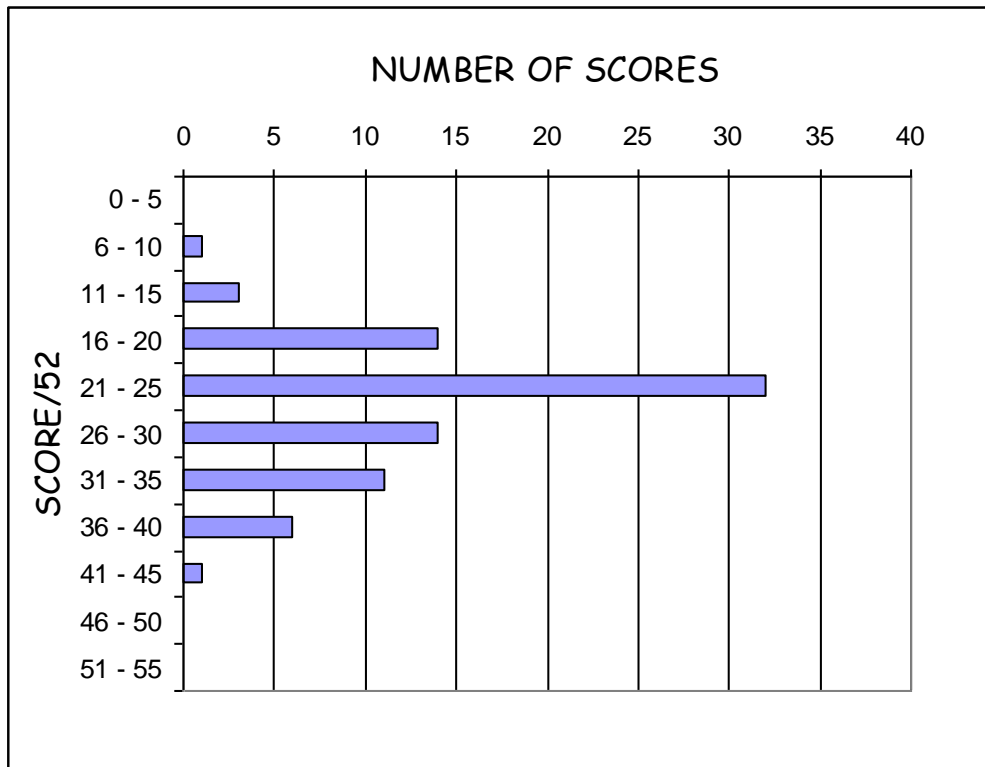
Figure 9. Frequency distribution of scores obtained by students completing graphical literacy questionnaire #1.

Analysis of tasks

A detailed analysis of the 16 individual tasks follows.

Only one question (2.1, a Level A2 task) was answered correctly by all 82 students.

Task #1, comprising three Level A1 tasks (locating values or reading the data), that should have been answered correctly by 100% of the students, resulted in correct responses ranging from 59,8% to 89%.   Only 44 students (53,7% of the class) correctly answered all three questions (i.e. are operating at the most basic level of graphical literacy) while two students gave incorrect responses to all three questions. Qu 1.3 confused 20 students (24,4%) who then only gave one point (A) as an answer instead of two points (A and C), clearly not noticing that points A and C share a common y-coordinate value of 100. Male students outperformed female students on Qu 1.1 (91,4% of the male students and 87,2% of the female students giving correct responses), Qu 1.2 (91,4% vs 85,1%), and Qu 1.3 (65,7% vs 55,3%).

Task #2, comprising four Level A2 tasks and one Level A3 task (reading between the data), that should also have been answered correctly by 100% of the students, resulted in correct responses ranging from 64,6% to 100%. Qu 2.1, a Level A2 task, was answered correctly by all 82 students (the 100% response). However, 22 students (26,8%) could not identify velocity as the dependent axis (Qu 2.4) confirming our observations that students are confused between the terms dependent and independent. Qu 2.3 caused similar confusion to Qu 1.3 in that 24 students (29,3%) did not note that half the maximum velocity (0,5) occurs at *two* pH values, namely 7 *and* 10. Finding ratios on a graph was a problem for 10 students (12,2% of the class) who could not find the correct ratio between the velocity obtained at pH 6 and the velocity obtained at pH 7 and incorrectly answered ¼, the ratio between the velocity value at pH 6 (0,25) and the highest velocity (1,0) obtained at pH 8,5. Only 33 students (40,2%) correctly answered all five questions (i.e. can read between the data locating maximum values or ratios). Male students outperformed female students on Qu 2.2 (100% of the male students and 97,9% of the female students giving correct responses), Qu 2.3 (74,3% vs 57,5%), and Qu 2.4 (82,9% vs 66,0%). .However, females (89,4%) outperformed the males (82,9%) on Qu 2.5.

Task #3, identifying features on a substrate-dependence graph, examined students' responses at three levels (A1, A2 and A3) and analysed whether they could read the data or between the data. Qu 3.1 (locating an axis and reading off the units), a very simple Level A1 task, that should also have been answered correctly by 100% of the students was only answered correctly ("*mM, the same as [S]*") by 40 students (48,8%). The option chosen most often (by 17 students) was *b*, a totally meaningless term (v/[S]) unrelated to any units of measurement. Clearly, these students assumed, erroneously, that the slope of the graph (v/[S]) was somehow related to the units of a point on the x-axis. Qu 3.2, a simple Level A2 task of finding twice the K_m value (0,2) and then reading off the resultant y-axis value (70) was answered correctly by 45 students (54,9%). Simply doubling the ½V_{max} value of 50 to 100, presumably because the K_m value of 0,1 has been doubled to 0,2, was chosen as a solution by seven students (8,5%). The other 21 students gave a range of incorrect values. Qu 3.3 (Level A3 - comparing features on a graph) was answered correctly by only 14 students (17,1%) while the other 68 students (82,9%) were clearly unaware that the value of ½V_{max} is a characteristic of the graph and does not change. The most popular answer, given by 20 students, was to simply report the velocity value of about 78 obtained at a [S] value of 0,3, i.e. three times the K_m value of 0,1. Male students outperformed female students on Qu 3.1 (54,3% of the male students and 44,7% of

the female students giving correct responses), Qu 3.2 (62,9% vs 48,9,5%), and Qu 3.3 (22,9% vs 12,8%).

Task #4, identifying features on an oxygen binding curve, examined students' responses at two levels (A3 and E) and analysed for the first time whether they could read beyond the data. For the two Level A3 tasks (Qu 4.1 and 4,2), 52 students correctly identified the independent variable and 49 of the same students (63,4%) correctly identified the dependent variable. Thirty students (36,6%) confused the two variables. Qu 4.3, a Level E task (extracting qualitative information from quantitative information and requiring deeper thinking) was answered correctly by 40 students (48,8%). The most popular incorrect answer, chosen by 25 students (30,5%), was option *a* (increases). These students assumed, incorrectly, that since the graph is moving to the right, *something* (the binding affinity) must be increasing because the values on the x-axis were increasing. They did not realise that at a lower pH a higher partial pressure of oxygen was required in order to achieve 50% binding. Male students outperformed female students on Qu 4.1 (71,4% of the male students and 57,5% of the female students giving correct responses) and Qu 4.2 (68,6% vs 53,2%). Females (48,9%) marginally outperformed males (48,6%) on Qu 4.3.

Task #5 examined students' responses at two levels (A3 and D) and also assessed for the presence of the classic slope/height confusion exhibited by physics students interpreting kinematics graphs.   Qu 5.1 (Level A3 – comparing slopes) was answered correctly by only 19 students (23,2%) with only 15 (18,3%) giving the correct explanation (at 20 sec the velocity of enzyme A is *greater* than that of enzyme B because the slope of A is greater than that of B) for their answer (i.e. they didn't guess). Instead of looking for changes in slope, the balance of 63 students (76,8% of the class) focused on the more perceptually obvious change in the height of the graph (at t = 20 sec, line B lies *above* line A). Qu 5.2 (Level D – reading a trend and noting that the slopes are not the same) was cognitively even more demanding than Qu 5.1 and was answered correctly by only seven students (8,5%) with only five (6,1%) giving the correct explanation for their answer (i.e. they didn't guess).   The 75 students (91,5%) who responded incorrectly to Qu 5.2 (by answering *Yes, at t = 40 sec*) did not realise that the two enzymes will not have the same velocity in the time frame of 0 to 100 seconds because the slopes of lines A and B are not the same. Instead, they chose t = 40 sec, the intersection point at which the graph lines have the same height, at a time when the *absorbances* are equal.   It seems that the majority of the Biochemistry students are concentrating on the wrong features of the graph in order to formulate their answers. Since these questions were formulated in a similar manner to

those used for probing physics students' problems with kinematics (see, for example, McDermott *et al.*, 1987) it is clear that the classic slope/height confusion is very persistent and even more prevalent than this researcher expected it to be. Male students outperformed female students on Qu 5.1 (25,7% of the male students and 21,3% of the female students giving correct responses) and Qu 5.2 (14,3% vs 4,3%).

The **sixth task** examined students' responses at three levels (A2 and D) and also required that they note the slope of the graphs. Qu 6.1 (Level A2 – identify steepest slope: a – between 0 and 40 sec) was answered correctly by 43 students (52,4%). The most popular incorrect answer, chosen by 14 students (17,1%), was option *d* (between 0 and 120 sec). These students clearly did not understand that they were required to look at the slope (the velocity) of the graph in order to answer the question and simply chose the safest option – the whole graph. Option c (between 80 and 120 sec) was chosen by 13 students (15,9%) who inexplicably decided that the part of the graph with the lowest slope indicated maximum velocity, perhaps because the maximum product concentration is shown at 120 sec. The 39 students (47,6%) who were completely inaccurate in their interpretation of the slopes of the two graph lines clearly did not know which features of the graph they should "read" in order to answer the question. Qu 6.2 (Level E – extracting qualitative information from quantitative information) was answered correctly by 60 students (73,2%). This was relatively high considering that this question was a Level E task, however, because a multiple choice format was used it could be surmised that recognition, not reasoning, played a significant role in the students' choice of answer. Questions 6.3 and 6.4 (both Level D tasks) resulted in similar responses with 44 students (53,7%) and 43 students (52,4%), respectively, giving correct answers. These two questions eliminated guessing by requiring students to give a reason for their answer. Of the 44 students who correctly answered Qu 6.3, five could have been guessing since they gave no explanation. Of the 43 students who correctly answered Qu 6.4, 11 could have been guessing since they gave no, or a completely inaccurate or garbled, explanation. The balance of students gave the correct explanation and could use the trend in a graph to predict a future outcome. The majority of incorrect explanations indicated that the students were experiencing difficulties in expressing themselves in English. An example of this (for 6.4): "*When the concentration drops the number of product cannot be affected by anything unless reagents are reduced*". Male students outperformed female students on Qu 6.1 (65,7% of the male students and 42,6% of the female students giving correct responses), Qu 6.2 (88,6% vs 61,7%), Qu 6.3 (54,3% vs 53,2%) and Qu 6.4 (60,0% vs 46,8%).

Task 7, a Level A4 task, required students to match the information in a narrative passage to a graphical representation. The majority of the students, 67 (81,7%), did not have a problem identifying the graph with a direct relationship between [E] and rate (graph 2: a straight line) but only 47 of the 67 students gave the correct explanation for their choice, an indication that the other 20 students were guessing. The most popular incorrect answer, chosen by eight students (9,7%), was option *a*. These students seemed to understand "double" (in the narrative description) as meaning "exponential" and chose the graph corresponding to an exponential trend. This was confirmed by the number of explanations using this terminology. An example of another incorrect explanation (justification for choosing option *c* - graph 3) was: *"As the enzyme catalyses the reaction the velocity will increase, but a catalyst only lasts for a certain period of time so as it gets finished the velocity decreases and becomes constant".* This explanation is completely at odds with the information given in the statement at the beginning of the task and is also biochemically incorrect. One rather garbled justification, for choosing option *a* (graph 1), was: *"The velocity increases as the curve becomes steeper from the increase of [E]".* Repeated reading of this response does little to clarify exactly what the student was intending to convey, other than a cry for help with respect to English language proficiency. Female students outperformed male students on this task (83,0% of the female students and 80,0% of the male students giving correct responses),

Task 8 examined students' responses at three levels (A2, D and E) and analysed whether they could read the data, read between the data or read beyond the data. This task was answered correctly by the most students with 77 students (93,9%), 80 students (97,6%) and 75 students (91,5%) correctly answering Qu 8.1, 8.2 and 8.3, respectively. The most likely reason for this high positive response could once again be the multiple choice format of the task :– recognition (or educated guessing) played a significant role in helping the students answer the questions. One could also argue at this point that the Level E task (Qu 8.2) was not really a Level E task since higher order thinking is not really a characteristic of multiple choice test items. Male students outperformed female students on Qu 8.1 (97,1% of the male students and 91.5% of the female students giving correct responses) and Qu 8.3 (97,1% vs 87,2%). Females (97,9%) marginally outperformed males (97,1%) on Qu 8.2.

Task 9 examined students' responses at two levels (A3 and D) and analysed whether they could compare slopes at different points on the fluctuating graph. Identifying the dependent and independent variables was also required. For the Level A3 task (Qu 9.1), 35 students (42,7%) correctly identified the dependent and independent variables, however, only 15 of these students could motivate their choice of answer. The balance of 20 students either gave no explanation or wrote something along the lines of *"The pH of amino acids is independent of the equivalents of OH⁻"* , endowing amino acids with an apparently new and hitherto unknown characteristic!. The 47 students (57,3%) who gave incorrect answers chose options that indicated an extensive misunderstanding of the terms "dependent" and "independent". Unsurprisingly the majority of these students chose option b (Both 'pH' and 'Equivalents of OH⁻' are dependent on each other), the safest option. This response was rather disturbing, however, since anyone with a basic chemistry background would know that the concentration of hydrogen or hydroxyl ions will determine the pH, not the converse. Surprisingly, only three students (3,7%) correctly identified the three points (B, D and F) on the graph (Qu 9.2 – Level D) with the lowest slope (corresponding to a small increase in pH with addition of OH); the worst response in the entire questionnaire. ⁻ If the students answer "A" (or "A to B" or "A to C"), the lowest point on the graph, then they have confused the slope of the graph with the height of the graph. i.e. slow = low = lowest part of graph. This was the case for 48 students who gave answers covering the bottom half of the graph (points A to D). The explanation by the 41 students who actually attempted to motivate their incorrect answer were mostly vague references to "*steepyness (sic) of slope*", "*long graphs*", "*constant change*", "*flattened graph*" and "*linear slope*". Point A was chosen by seven students who clearly confused slope with height in looking for an answer. Another 19 students chose the upper part of the graph and gave answers covering all of the points from D to G. The 13 students who gave an explanation for their incorrect answer generally referred to the idea that the pH was slower between D and E because the reaction was nearing completion or 'end point' or the height between the points was the least. What is clear is that most of the students did not understand the question and only gave one or two points or a range. Students performed marginally better with qu 9.3 (also a Level D task) where 10 students (12,2%) correctly identified the two points (C and E) on the graph with the steepest slope (corresponding to a large increase in pH with addition of OH⁻). However, 52 students gave ranges between A and E as their answer, 42 giving explanations that ranged from "*the reaction starts here*", "*…the optimum point of pH increase*" to "*pH increases with 2 units from B – C*". One student decided that point A signified the fastest change in pH because *"the*

solution was still more concentrated because not much OH was added yet". The rest of the students answered G or F-G, confusing slope with height. One response for answering 'F to G' was unique: "*The distance between F and G is lesser than other points".* This student inexplicably interpreted "change in pH the fastest" as signifying the shortest distance between two points and not the slope of the graph. From these responses, particularly the choice of points based on differences in height (distance in units), it is clear that the majority of students were focussing on the more perceptually obvious changes in height rather than slope. This again confirmed the finding of task 5 that the overwhelming majority of the students confuse the slope and height of a graph. However, the exceptionally low score for this question is an indication that the students find it more difficult to interpret a curved graph than a straight-line graph. The titration graph in question involves changes in slope as well as changes in height and the changes in slope seem to be not as perceptually obvious to the students as changes in height. This type of graph requires more careful examination before information can be extracted from the graph. Fluency in English again appears to be a problem with more than a few students offering *"the line is very steepy towards the top of the graph"* as an explanation for Qu 9.3. Male students outperformed female students on Qu 9.1 (51,4% of the male students and 36.2% of the female students giving correct responses), Qu 9.2 (8,6% vs 0%) and Qu 9.3 (22,9% vs 4,3%).

The **tenth task** required students to match the quantitative information in a pH vs time graph with a logarithmic x-axis scale to a different graphical representation, examined students' responses at Kimura's level A4 but also required that they be able to recognise different graphing practices with regard to scale in order to choose the correct graph. Our observations that a large number of students experience difficulties with scale were confirmed by the finding that only half of the students (41) correctly chose graph 1 (option *a*) as their answer. The other 41 students were incapable of interpreting the x-axis scale used in the test item graph. The most popular answer, chosen by 36 students (36,9%) was option *c* (graph 3), a straight-line graph, which would appear completely different if the x-axis was drawn with a logarithmic scale. Graph 2 (option *b*) was chosen by three students and two students didn't know which graph to choose. Female students outperformed male students on this task (55,3% of the female students and 42,9% of the male students giving correct responses).

Task 11 examined students' responses at Kimura's level A2 but also analysed for the first time the extent of 'zoom graph confusion' (see section 2.5.1). Surprisingly, only 38 students (46,3%) noticed (in Qu 11.1) that the y-axis of the graph started at nine and not at zero. However, since the correct option was *d* (none) and the answer most

likely to be chosen by students unsure of the answer, the result for Qu 11.2 is perhaps more revealing. Only 30 students (36,6%) correctly chose option *c*, an exact figure and not a potential guess. This implies that 52 students did not notice the zoom effect used in the graph. These students extracted information from the graph without first critically examining the y-axis, revealing a preference for 'looking at the picture' rather than 'reading the y-axis'. Male students outperformed female students on Qu 11.1 (62,9% of the male students and 34,0% of the female students giving correct responses) and Qu 11.2 (42,9% vs 31,9%).

Task 12 examined students' responses at four levels (A1, A2, A3 and E). A second y-axis (which does not commence at zero) was introduced for the first time and students had to discriminate between the two graphs, a potential source of confusion. For **Qu 12.1** (Level A3), only 34 students (41,5%) correctly read the 2nd y-axis to identify the range of the NaCl gradient and chose option *c* (2.5 - 5 mM). The other 48 students became confused on being confronted with a secondary y-axis and chose option *a* (20 students), option *b* (15 students), option *d* (12 students) and no options (one student). Students who chose option *a* incorrectly read the primary y-axis (A_{280}) and recognised the values given in option *a*. The 15 students who chose option *b* erroneously assumed that since the NaCl gradient starts on the primary y-axis and ends on the secondary y-axis the gradient must start at 0,1 mM (primary y-axis) and end at 5 mM (secondary y-axis). Option *d* was the "I don't know" option. Qu 12.2, a simple Level A2 task locating a maximum value, was answered correctly by 65 students (79,3%), a disappointing result as this question should have resulted in a 100% correct response. The other 17 students gave answers as a range of fractions or listed single fractions such as number 8. Inexplicably, six students listed fraction number 20 as containing the most protein, even though this fraction contains almost no protein (evident if you read the primary y-axis). Locating corresponding points on two graphs and extrapolating to the x- and y-axes was not a problem for 30 students (36,6%) who correctly answered Qu 12.3 with 3,5 mM. The other 52 students, however, gave a range of values from 0,3 to 20 with 0,3 being the most popular (given by 10 students). The 0,3 value was read off the primary y-axis, the corresponding point on the secondary y-axis actually being the correct answer! Clearly, these students are unfamiliar with use of secondary y-axes.

Describing the graph (with some help in the form of multiple choice options) for Qu 12.4 (Level E) was possible for 43 students (52,4%) who correctly chose option *a*. The rest of the students chose option *b* (11 students), option *c* (17 students), option *d* (nine students) and none chosen (two students). The students who chose option c did

not read the labels on the graphs and did not notice that the NaCl is increasing in concentration, not the protein. The rest of the students experienced problems with two graphs on one set of axes with two y-axes and could not extract the necessary information required to answer the question. Apparently, these students have difficulty contending with multiple sources of information which may include keys or legends, axes and labels.

Male students outperformed female students on Qu 12.1 (54,3% of the male students and 31,9% of the female students giving correct responses), Qu 12.2 (82,9% vs 76,6%), Qu 12.3 (51,4% vs 25,5%) and Qu 12.4 (62,9% vs 44,7%).

Task 13 examined students' responses at three levels (A2, A3 and A4). The task was complicated, however, by the reciprocal nature of the data. Answers therefore required a simple calculation (inversion and negative inversion) in order to choose the correct option. This was not a problem for 31 students (37,8%) answering Qu 13.1; 32 students (39,0%) answering Qu 13.2; and 56 students (68,3%) answering Qu 13.3. Answers given by the remainder of the students indicated a serious problem working with reciprocal values of data. Interestingly, only 10 students (12,2%) were able to correctly answer Qu 13.4. Although enzyme B has higher velocities at all substrate concentrations, 72 students made the mistake of looking for an answer in the negative quadrant (extrapolated portion), where the graphs actually intersect, and erroneously concluded that the enzymes had the same velocity at a negative substrate concentration. Male students outperformed female students only on Qu 13.1 (42,9% of the male students and 34,0% of the female students giving correct responses). Females outperformed the males on Qu 13.2 (females 40,4% vs males 37,1%), Qu 13.3 (females 70,2% vs males 65,7%) and Qu 13.4 (females 14,9% vs males 8,6%).

The **fourteenth task** examined students' responses at one level (D - reading a trend from fluctuating phenomena) but was complicated by the fluctuating nature of the slope of the graph which can be represented in a different format. Relating one type of graph to another did not prove a challenge for 24 students (29,3%) who correctly chose graph 3. The other 58 students were unable to translate from a G graph to a ΔG graph simply by using the slope of the G vs reaction progress as the height of a ΔG vs reaction progress graph. In fact, the majority of the students (46 :- 56%) chose graph 2, a graph with a similar shape to the original graph. Thus, in trying to construct a new graph from data embedded in another graph, students seem unable to ignore the shape of the original graph and were relying solely on the visual, or iconic, appearance of the graph and ignoring specifiers and referents such as the axis labels.

The 'don't know' option was chosen by two students and graph 1 was chosen by 10 students who clearly had no idea how to visualise change in slopes. Male students outperformed female students on this task (34,3% of the male students and 25,5% of the female students giving correct responses).

Task 15 examined students' responses at three levels (A3, E and F) and also assessed for the first time whether the students could think critically and posit a likely explanation for the presence of two pH optimum peaks in the graph. Qu 15.1, a very easy percentage calculation (at Level A3), was answered correctly by 77 students (93,9%). Giving a description for qu 15.2 proved a challenging Level E task with only 32 students (39%) managing to satisfactorily describe the features on a simple pH optimum graph. The remaining 50 responses were vague narrative sentences largely unrelated to the graphed information. A number of students decided that the graph resembled an elution profile obtained after chromatography (a technique totally unrelated to determining the pH optimum of an enzyme reaction). Typical responses were: *"A graph showing/comparing two peaks of eg amino acids at what pH is the maximum velocity of each"; "The elution of proteins; which has eluted first"* and *"Elution graph of an amino acid or protein"*. Reading and understanding all the velocity and pH information presented graphically and then drawing on their own enzymology knowledge to finally supply, for qu 15.3 (a Level F task), a suitable scenario (two enzymes, each with their own pH optimum) to explain the graph was possible for only four students (4,9%). The capacity to provide an explanation that goes 'beyond the data' appears to be sorely lacking in the senior undergraduate biochemistry students at the UFS. Examples of off-the-mark explanations (quoted verbatim) are: *"Proteins with higher molecular weight elutes last and those with small weight elutes first"; "Because of chemicals produced at pH 8, that resulted to increase in temperature"; "As the concentration of the protein decreases, the velocity decreases"*, and *"…as the velocity increases the acidity looses it's value, and as it decreases the basic gains it's alkalinity"*. Male students outperformed female students on Qu 15.1 (97,1% of the male students and 91.5% of the female students giving correct responses) and Qu 15.3 (5,7% vs 4,3%). Females (46,8%) outperformed males (28,6%) on Qu 15.2.

The final task, **Task 16**, examined students' responses at level F. Students were required to offer a likely explanation for the presence of the dip in velocity. Even though there are two likely scenarios for the observed results, only six students (7,3%) were able to give a satisfactory explanation with four of the six students scoring two points for a full explanation. Four of the six students were male and two were female. The rest of the students either did not attempt the task, wrote completely unintelligible

gibberish or gave responses exposing a profound lack of abstract or critical thinking. A few of these included the 'elution' response found in task 15 as well as some microbiological techniques such as a dilution series. One rather confused response was: *"The steps between 15 – 20 could have been carried out without going through step 5 – 14, because they seem more faster to save time."* Another confusing response: "*The researcher used spectrophotometer he started from those with light colour to the darker ones.*" One student even thought that the hapless researcher had a shortage of laboratory equipment responding with *"It possible that he had only 19 cuvette holders and measured the 1ˢᵗ and then started with the last ones that he used."* Two students cast doubts on the researcher's laboratory skills with *"He did not clean the cuvettes or put the wrong values of mixtures in"* and *"did not clean the cuvettes or must have contaminated his results somehow" .* The results for this final task were disappointing to say the least. Students in the natural sciences should be able to operate at an abstract level such as Kimura's Level F, especially if they intend continuing with post-graduate study.

Graphical literacy levels

The average percentage correct responses given by students for Questionnaire #1 were grouped according to level and calculated using Kimura's framework. These results are visualised graphically in Fig. 10.

It is clear from Fig. 10 that the percentage of positive responses declines as the difficulty level increases, an expected result. However, Level E tasks gave an unexpectedly high positive response, probably because of the multiple-choice nature of these questions. These overall scores per Kimura's levels are relatively low, considering the seniority and age of the students. This observation is confirmed when responses at level A1 are analysed: only 64,4% of responses at this level were correct. Third-year Biochemistry students should have answered 100% of all the tasks set at Level A1. Only 8,5% of the Level F responses were correct, an indication that the overwhelming majority of the Biochemistry students were incapable of answering tasks set at this level.
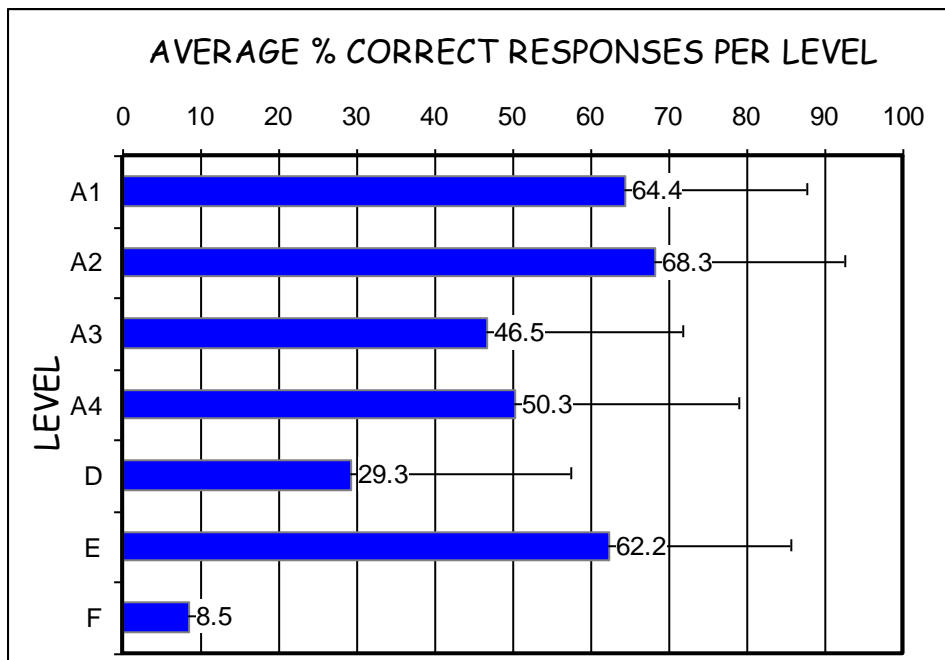
Figure 10. Average percentage correct responses for Questionnaire #1 grouped and calculated as per Kimura's levels of statistical ability.

When the same scores per task question were grouped and calculated according to Curcio's levels of graph comprehension (see Table 4) a clearer picture emerged regarding the biochemistry students' graphical literacy skills. Overall scores for Curcio's three levels are visualised in Fig. 11 and indicate, once again, a disturbingly low response at the highest level of graphical comprehension: 'Reading Beyond the Data'.
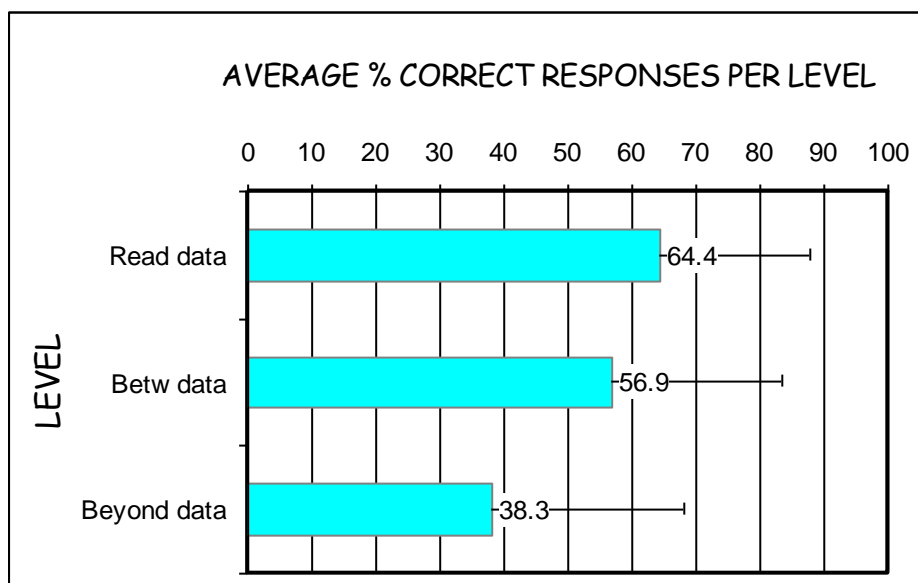
Figure 11. Average percentage correct responses for Questionnaire #1 grouped and calculated as per Curcio's levels of graphical comprehension.

In order to draw conclusions regarding students' performance at different levels of graphical (or statistical) ability, it is necessary to analyse student's individual performances at each level (i.e. a micro-level or sub-skill level) and derive a graphical literacy "profile" of each student. The results are presented graphically in detail in **Appendix C** (Kimura's levels) and **Appendix D** (Curcio's levels) and are anonymous (student's names and student numbers have been replaced by a number but results are still arranged numerically according to their student number). Disturbingly, only 10 students out of the 82 in the class performed 100% at Level A1 (Kimura). Only four students were operating at Level A2 and no students operated 100% at Level A3. Level A4 was achieved 100% by five students while no students achieved 100% at Level D. Level E was achieved by three students and two students achieved at Level F. Only three students operated 100% at two levels. The most graphically literate student (using Kimura's levels) was student #23 while the most graphically 'illiterate' student was student #1.

When individual performance was analysed according to Curcio's levels of graphical comprehension (**Appendix D**), the lack of graphical literacy amongst biochemistry students appeared slightly better. Ten students out of the 82 could be regarded as being able to 'read the data', four students were able to 'read between the data' and ten students were able to read beyond the data (if only the level where 100% was scored was taken into account). No individual students operated at all three levels. The most graphically literate student (using Curcio's levels) was student #49. Student #1 retained her position as least graphically literate.

4.2 Research question 2

Answering the question: *"Are the graphical literacy levels of undergraduate biochemistry students' reflected in their academic performance?"* was attempted using bivariate analysis. .

An analysis (bivariate) using Pearson's correlation coefficient indicated a moderate positive linear relationship between the mean of the students' biochemistry subject module grades (obtained by the students over two years of study) and their test mean score obtained for Questionnaire #1: $r(82)=0,433$. For these data, the mean ($\pm$SD) for the Biochemistry subject module grades was 57,1% ($\pm$7,7) and the mean test score for Questionnaire #1 was 48,3% ($\pm$12,1). The graph obtained with this data is shown in

Fig. 12 and implies that a low score on the graphical literacy test is a relatively good indicator of poor academic performance in biochemistry, and vice versa, but does not imply causality.
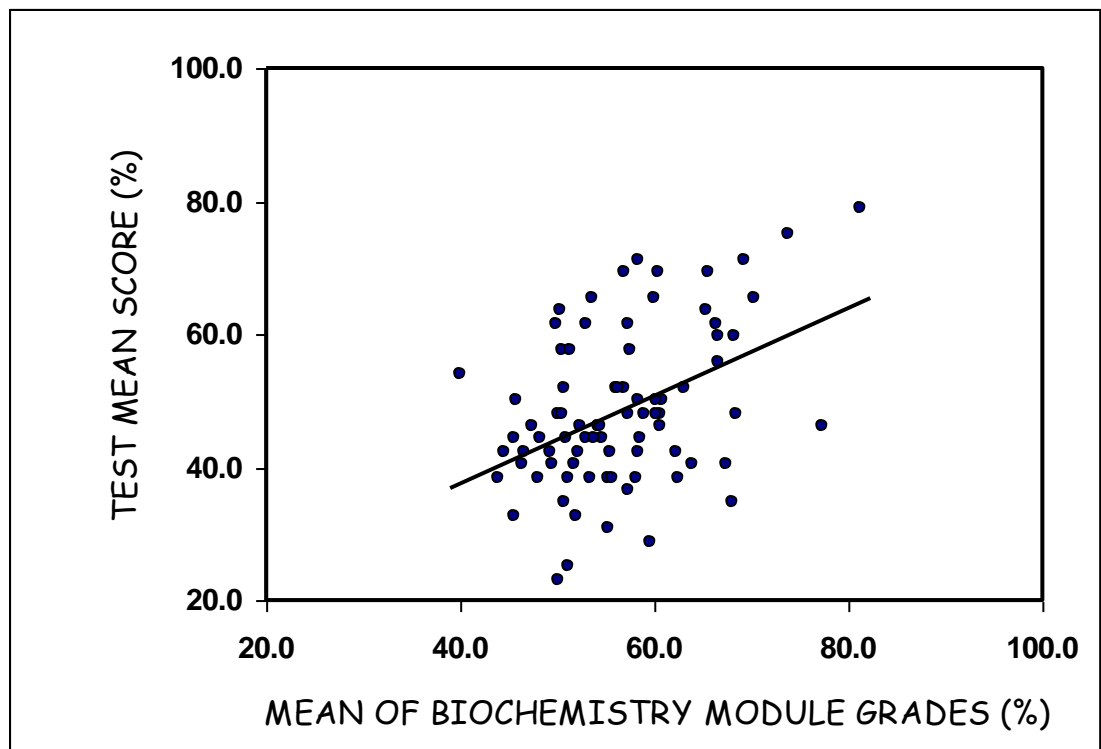

Figure 12. Bivariate analysis of test mean scores and Biochemistry subject module means obtained by students.

In an attempt to identify the source of the widespread scatter in Fig. 12, and resultant lower r-value, the bivariate analysis was repeated in a group context. Results are shown in Fig. 13.
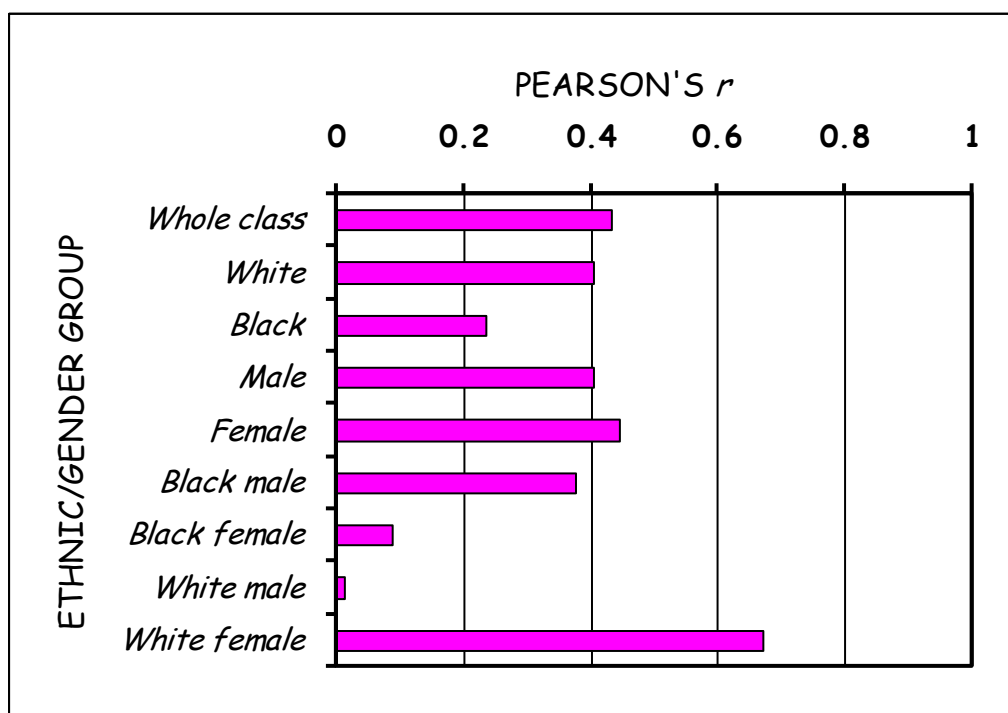
Figure 13. Summary of bivariate analyses results in group context.

It is clear from Fig. 13 that the source of the scatter in Fig. 12 is the data for white males and black females. White males showed the worst correlation (r=0,012) between test mean scores and the grades obtained by them for undergraduate biochemistry subject modules. White females showed the best correlation (r=0,671) between test mean scores and the grades obtained by them for undergraduate Biochemistry subject modules. However, it should be noted that these two groups also comprise the smallest groups with eight students and nine students per group, respectively. It would be foolhardy to draw any conclusions whatsoever regarding possible reasons for these conflicting correlations.

The finding that the graphical literacy levels of undergraduate biochemistry students' are reflected in their academic performance does not imply causality but rather confirms that the understanding of biochemistry and the interpretation of biochemical data requires well-developed logical and analytical cognitive skills. Although both the theory and practice of decoding and encoding graphs is a component of the syllabus of almost every undergraduate biochemistry course, the weighting in terms of allocation of points during assessment tasks is usually not enough to favour the student who has well-developed graph encoding or decoding skills, a fact that further confirms the conclusion that it is well-developed logical and analytical cognitive skills that are required for success in biochemistry and not merely good graphing skills.

4.3 Research question 3

The potentially contentious research question*: "Is gender or race related to graphical literacy levels of undergraduate biochemistry students?"* will hopefully be answered after extensive comparison of overall responses as well as individual achievements at different levels of graphical literacy.

Test mean

The group of students was not homogeneous with respect to the graphical literacy test score mean. When the graphical literacy test mean was calculated per gender or ethnic group it became evident that there are group differences and that some of these are significant. Statistical analysis (summarised in Table 6) revealed that the test mean obtained by white students was significantly higher than that obtained by black students. The test mean obtained by male students was also significantly higher than that obtained by female students. Although both black and white males scored more than their gender opposites only the black males scored significantly higher than their gender opposite, the gender effect being insignificant w.r.t. white students. Black females scored the lowest test mean and can be considered a risk group worthy of intervention in terms of graphical literacy training. White males obtained the highest overall test mean and can, therefore, be considered the most graphically literate group. However, white males also comprise the smallest group (eight individuals), thus relegating this conclusion to merely tentative until confirmed using a larger group.

In terms of individual responses, males outperformed females in 41 of the 51 questions.

Table 6. Tests of significance between graphical literacy test mean scores obtained by gender and ethnic groupings.

Mean scores		*T-Vakue*	*df*	*P-Value*
Ethnicity				
White	**Black**	-4,94	25	0,000021**
59,7% (±10,7)	45,3% (±10,7)			
Gender				
Male	**Female**	3,11	73	0,0013*
52,9% (±11,6)	44,9% (±11,5)			
Gender ethnicity				
White male	**White female**	0,57	14	0,29
61,3% (±8,6)	58,3% (±12,6)			
Black male	**Black female**	3,37	46	0,00076**
50,4% (±11,3)	41,7% (±8,6)			
White male	**Black male**	2,92	15	0,0053*
61,3% (±8,6)	50,4% (±11,3)			
White female	**Black female**	3,74	10	0,0019*
58,3% (±12,6)	41,7% (±8,6)			
White male	**Black female**	5,87	10	0,000078**
61,3% (±8,6)	41,7% (±8,6)			
Black male	**White female**	-1,67	13	0,059
50,4% (±11,3)	58,3% (±12,6)			

* p<0,05
** p<0,001

The extent of the differences between the groups become more apparent when visualised graphically (see Fig. 14).
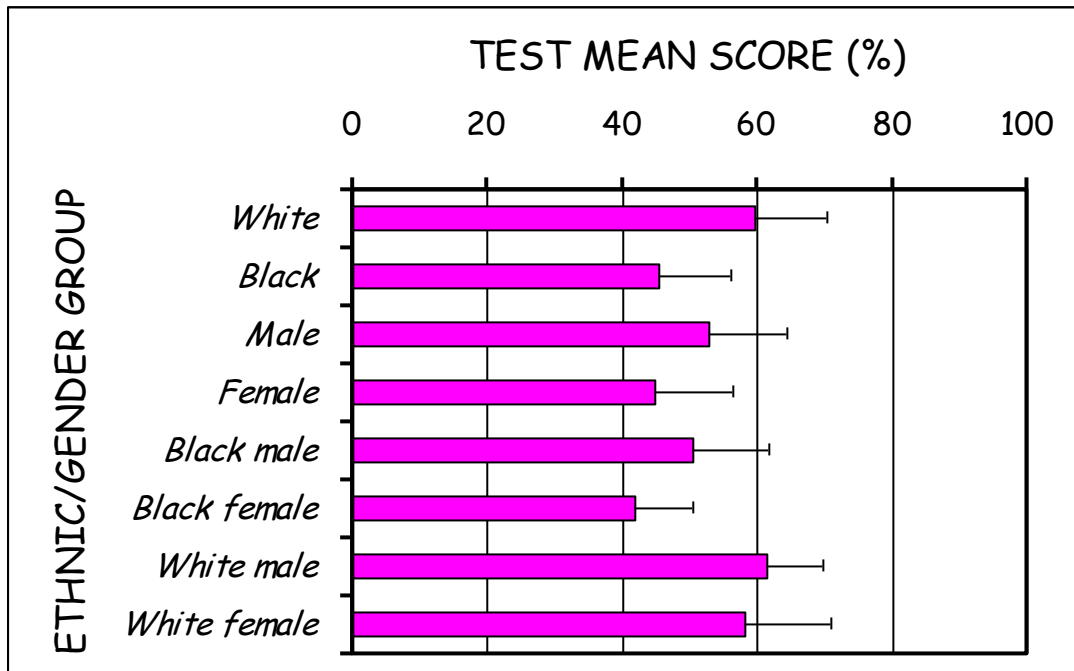
Figure 14. Graphical comparison of test mean score calculated per gender or ethnic group.

These gender differences in decoding graphical tasks are consistent with those reported by Lowrie & Diezmann (2011), who found that boys aged 9-12 years outperformed girls on graphical languages, and Spelke (2005), who found that boys tend to perform better on tasks that require mental rotations. Possible reasons could be the type of task question – short – that tends to advantage males (Lokan *et al.*, 2001) and the different life experiences of females (Tracey, 1990).

Graphical literacy levels

The mean percentage correct responses given by students for Questionnaire #1 were grouped according to level and calculated using Kimura's framework. These results were then grouped according to gender and ethnicity. These results are visualised graphically in Fig. 15.

White males outperformed the other gender/ethnic groups in five of the seven levels of Kimura. Only the white females were able to outperform the white males in Level A4 and Level F tasks. Black females performed the worst at all levels (except for Level A4 tasks where black males performed the worst) and can be considered as possessing, at best, rudimentary graphical literacy skills.
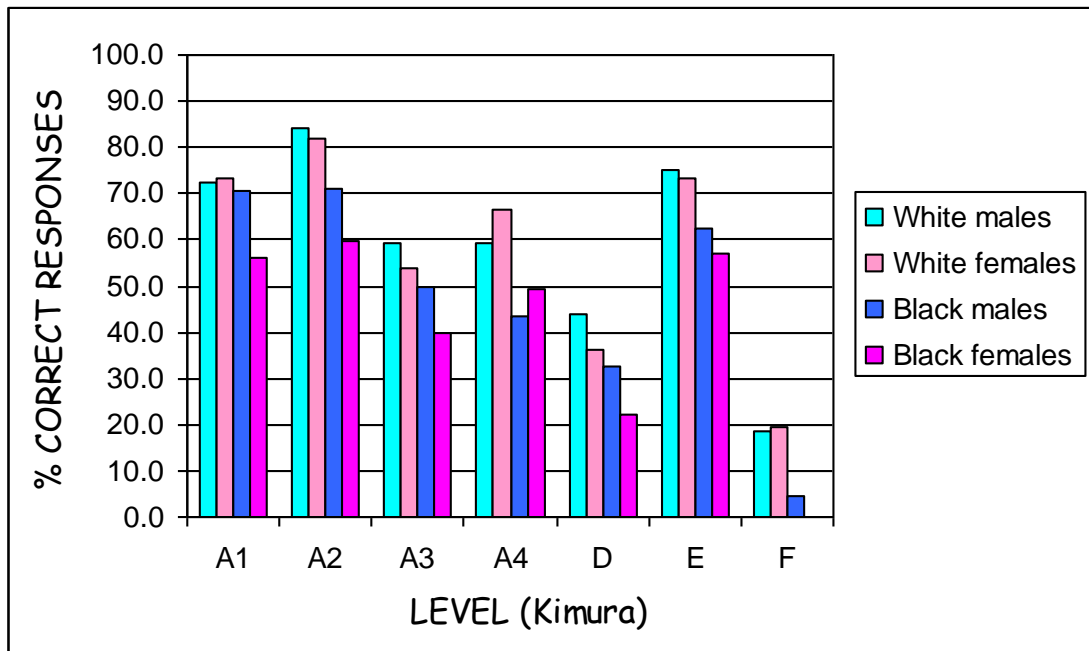
Figure 15.   Graphical comparison of correct responses at Kimura's levels achieved by students after grouping in terms of gender and ethnicity.

There appears to be a dearth of literature regarding ethnic differences in graphical literacy performance (too contentious, no doubt). The overwhelmingly obvious factor contributing to ethnic differences in decoding abilities could be the sometimes vast discrepancies in the quality of education in mathematics and science received by white and black students. Black students' lower levels of graph-visual literacy could be a result of a deficiency in prior experience, use of and exposure to graphs and graphing practices within a scientific educational environment – a typical scenario in the majority of South African schools.

Given the sample composition and size, particularly w.r.t. white males, the conclusions reached for this section can, at best, be described as preliminary and tentative. Having said this, there is still some indication of a possible relationship between aspects of graphical literacy abilities and the gender or ethnicity of the test participants.

4.4 Research question 4

Research question 4 (*Is it possible, using eye gaze data, to discern a difference in the way students with high graphical literacy and students with low graphical literacy interpret graphs?*) was addressed using eye tracking and two groups of students – a

TOP 10 group and a BOTTOM 10 group. Results were first analysed by assessing students' responses to the 12 tasks in Questionnaire #2 after which eye tracking data was used to compare the two groups.

The test mean score for individual students as well as the individual graphical literacy 'profiles' (levels as per Kimura, 1999) of students obtained in Phase 1 of this study (see **APPENDIX C**) were used to select 10 students who scored the lowest and 10 students who scored the highest for Questionnaire #1. These students were then grouped as "Bottom 10" and "Top 10" (individual graphical literacy profiles are shown in Fig. 16). The mean of the test mean scores (in Questionnaire #1) of the students in each group were 32,1% (±6,7) and 63,5% (±20,0) for "Bottom 10" and "Top 10", respectively. The worst student was student #1 and the best student was student #23 (the original list of students was ordered in terms of university student number and numbered 1- 82).
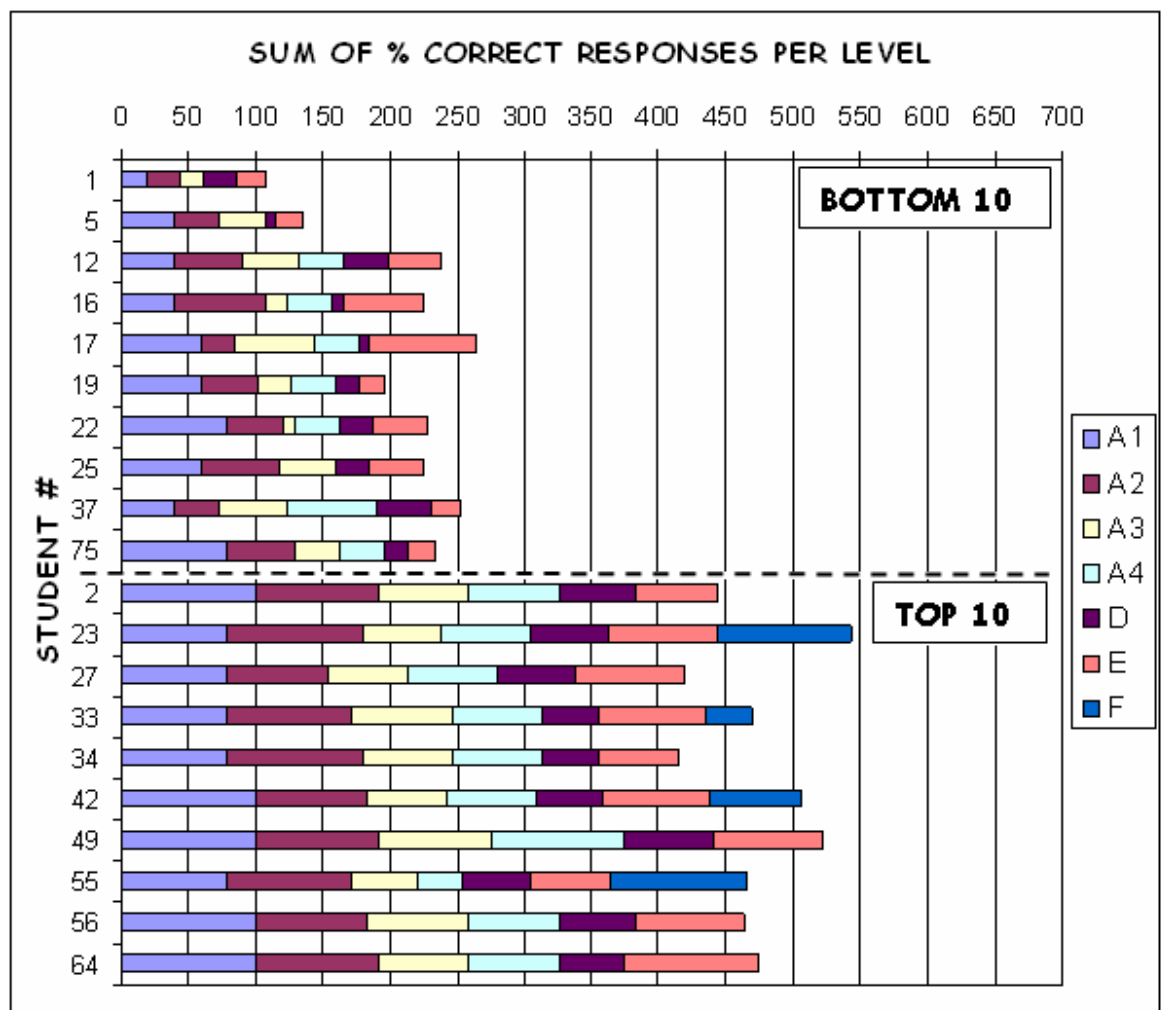
F
Figure 16. Individual graphical literacy profiles (using Kimura's level descriptors and Questionnaire #1) of students selected for the "Top 10" and "Bottom 10" groups for Phase 2 of the study.

Test mean

Overall, the combined (Top 10 + Bottom 10) group of students confirmed their poor graphical literacy capabilities by obtaining a test mean score of 48,5% (±8,29) for Questionnaire #2 (displayed on a monitor and answers recorded). This score was identical to the test mean score obtained with Questionnaire #1! The percent correct responses per question recorded by students for Questionnaire #2 is shown in Fig. 17. The lowest score was 6/16 (37,5%) and the highest score was only 10/16 (62,5%). All of the students were able to correctly answer Qu 3.1, a Level A4 task, but no students were able to give a satisfactory explanation required for Qu 2.1.
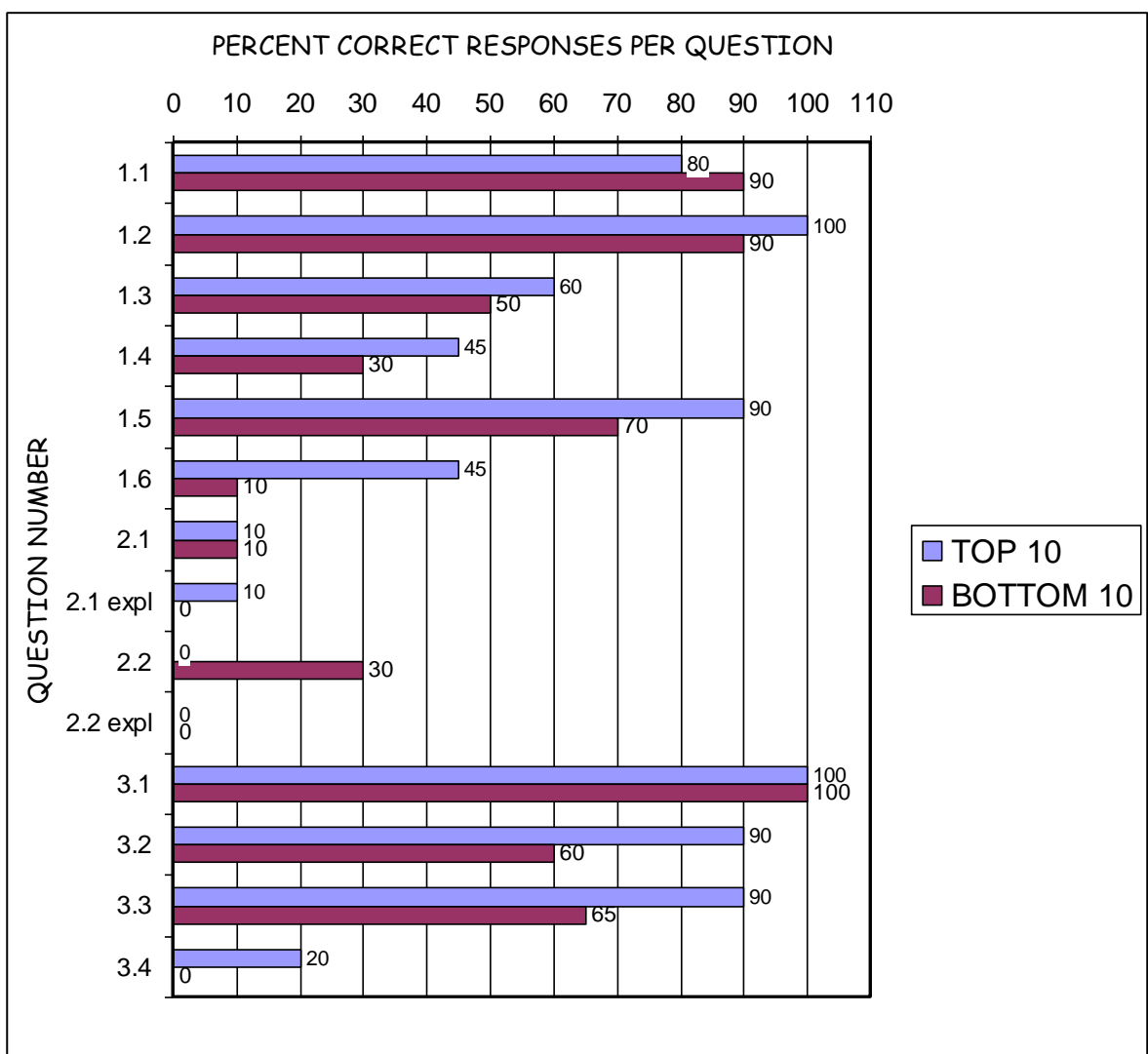
Figure 17. Correct responses recorded per question for Questionnaire #2.

As expected, the performance on the test differed significantly (p < 0,0001). between the two groups of students with the Top 10 group scoring 58,8% (±4,37), 14,4 percentage points more than the Bottom 10 group's 44,4% (±5,47) test mean score.

Response time

The total time taken to complete the test differed between the two groups. The Top 10 group took a significantly ($p < 0.005$) shorter mean time of 293,7 (±44,6) seconds to complete the test than the Bottom 10 group who took a mean time of 458,6 (±138,3) seconds to complete the test. The Bottom 10 group took, therefore, 56% longer than the Top 10 group to complete the 12 questions. Student #42 took the shortest time of 227,5 seconds and student #5 took the longest time of 619,7 seconds to complete the test. In general, the more cognitively demanding a task was the longer the response time was, for example Task 2 and the last two questions of Task 3 took the longest for students to answer.

Mean response times per question for each group are summarised in Fig. 18.
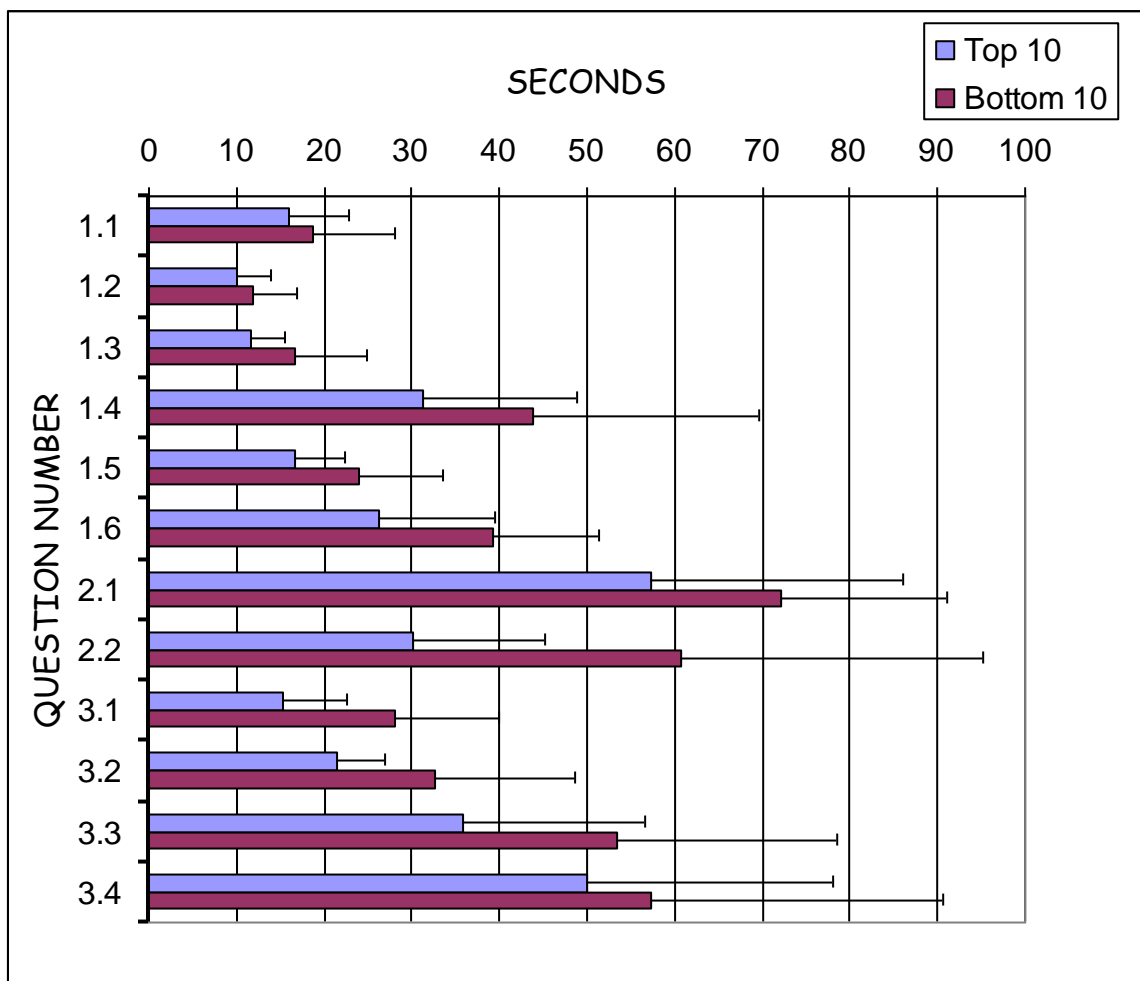
Figure 18. Mean response times per question for each group for Questionnaire #2.

Bivariate analysis using Pearson's correlation coefficient indicated a strong negative linear relationship between the total response times and the test score obtained by

students for Questionnaire #2: $r(20) = -0,73$. The graph obtained with this data is shown in Fig. 19 and implies that the longer a student takes to complete Questionnaire #2 the lower the student's test score will be and vice versa.
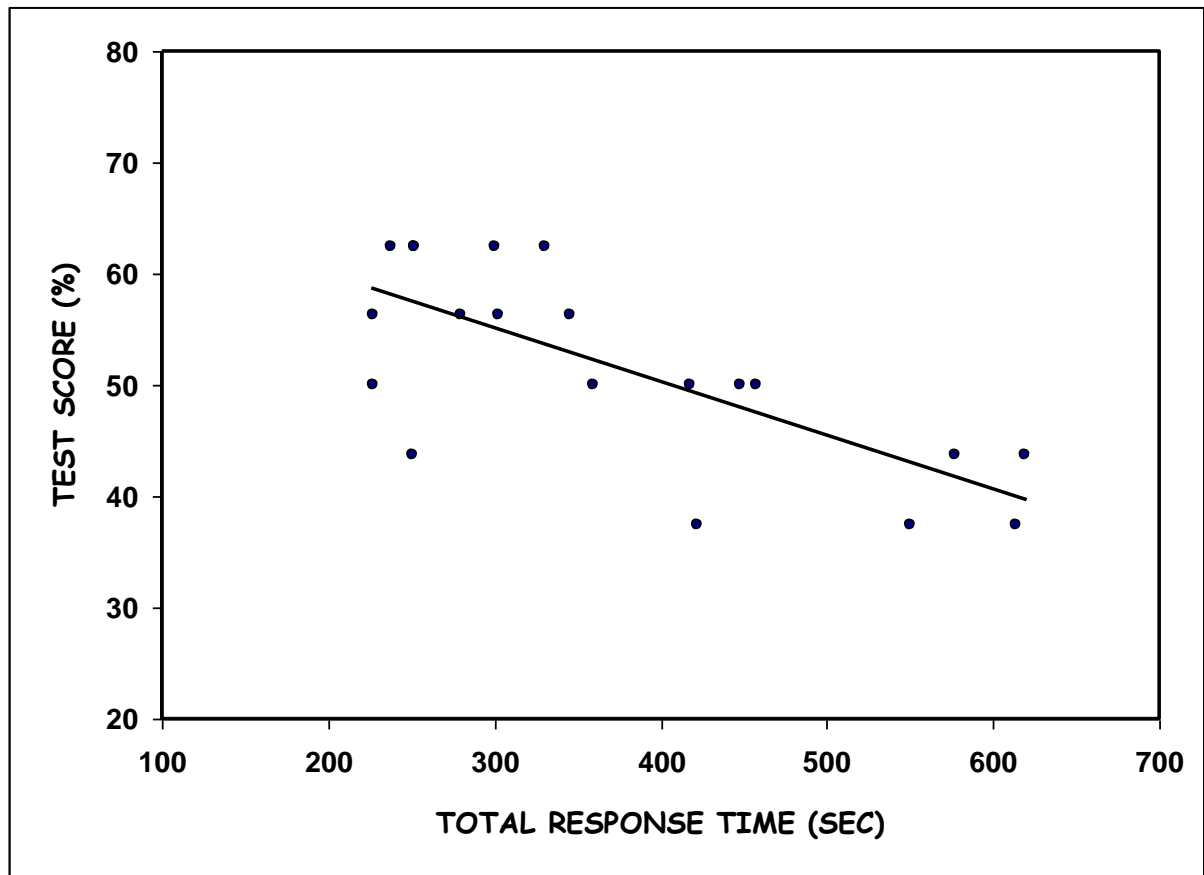

Figure 19. Bivariate analysis of test scores and total response time.

Levels of graphical literacy

The mean percentage correct responses given by the Top 10 and Bottom 10 groups of students for Questionnaire #2 were grouped according to level and calculated using Kimura's framework. These results are visualised graphically in Fig. 20. Although the sample size and number of questions at each level is too small to draw any concrete conclusions, the comparison does serve to highlight the fact that there is not much difference between the two groups w.r.t. cognitively undemanding tasks at the lower levels of Kimura's hierarchy (i.e. Level A1 – D). However, it also clear that the Top 10 group of students outperformed the Bottom 10 group students at Level E and Level F tasks, a result expected from the profile of student selected for this phase of the study.
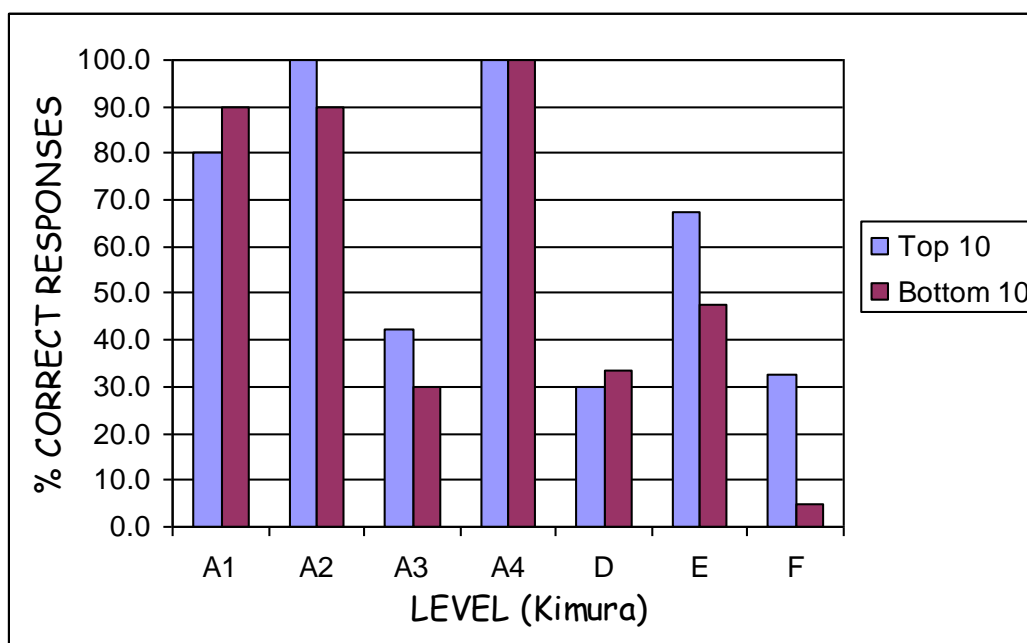
Figure 20. Graphical comparison of mean correct responses for Questionnaire #2 at Kimura's levels achieved by Top 10 and Bottom 10 students.

Analysis of individual tasks

A brief analysis of the 3 tasks follows.

Task 1, an expanded version of Task 8 in Questionnaire #1 (identifying features on a temperature-dependence graph), examined students' responses at six levels (A1, A2, A3, D, E and F) and analysed whether they could read the data, read between the data and read beyond the data. Responses for this task are summarised in Table 7.

Table 7. Correct responses for Task 1 of Questionnaire #2.

GROUP	Correct responses per question					
	1.1	**1.2**	**1.3**	**1.4**	**1.5**	**1.6**
TOP 10	8	10	6	9	9	9
BOTTOM 10	9	9	5	6	7	2

The correct answer for **Qu. 1.1** (*What velocity is observed at a temperature of 55°C?*) was 20 μM/min (Level A1 - locating values; read the data). This type of task was chosen for the experiment because it is so widely performed: using a graph to elicit the value of one variable corresponding to a given value of another, and should have resulted in a 100% positive response. Although 17 of the students could answer this

question without difficulty, three students gave incorrect answers. Two of the three students answered "45" with no units and one student answered "10 μM/min". The most likely explanation for these answers is that the students misread the x-axis (50 and 60) and extrapolated onto the incorrect point on the graph before reading off the corresponding value on the y-axis (45 and 10, respectively). Interestingly, all of the Top 10 responses, bar one, included the units (μM/min) in their answer whereas all of the Bottom 10 students, bar one, did not include the units in their answer.

The correct answer for **Qu. 1.2** (*What is the maximum reaction velocity?*) was 90 μM/min (Level A2 – identify max value, read off y-axis). Nineteen of the 20 students could identify the maximum y-value. One student gave the incorrect answer of 45 by reading the x-axis even though the y-axis was clearly labelled "Velocity". Students (exactly the same individuals) included or omitted units in their answer in an identical manner to that observed for Qu 1.1.

The correct answer for **Qu. 1.3** (*Which axis is the dependent axis?*) was the 'Velocity' or y-axis (Level A3 – compare two features on a graph: the dependent and independent axes). Eleven (six from Top 10 group and five from Bottom 10 group) of the 20 students (i.e. 55%) were able to correctly identify the dependent axis, a similar result to that obtained for Questionnaire #1. It appears that, irrespective of the type of graph or which group of students are being tested, approximately 40% of the students seem to have a problem identifying the dependent or independent axes, either because of a lack of understanding of the concept of dependency, or because of a vocabulary or language comprehension problem (unfamiliarity with the term).

The correct response for **Qu. 1.4** (*How would you describe the trends in the graph?*), a Level E task (extracting qualitative data from quantitative data), was something in the vein of "*The velocity of an enzyme-catalysed reaction increases initially to a maximum and then decreases with an increase in temperature*". Nine of the Top 10 group and six of the Bottom 10 group were able to give a satisfactory description of the trends in the graph, although their command of English was tenuous to say the least. The term "trends" seemed to be poorly understood and eight of the students requested the term to be explained to them before they could formulate an answer. The remaining five students gave answers as divergent as "*It exponential*" and "*It goes up and down*". One student said that the graph *"…resembles a casein graph..",* a totally unrelated type of graph.

The correct answer for **Qu. 1.5** (*What do you think will happen at a temperature of 90°C?*), a Level D task (prediction made on the basis of a trend), was "*Velocity will reach or be close to zero*" OR "*eventually no activity will be detectable*". Nine of the Top 10 group and seven of the Bottom 10 group were able to give a satisfactory prediction by extrapolating the graph beyond 70°C. The remaining four students displayed a marked lack of trend prediction by guessing: "*Optimum activity*"; "*The velocity will remain constant.*"; "*They will no longer have substrate like the enzyme no you don't have substrates anymore.*"; and "*..the graph has reached it's peak and it can't go further.*"

The best explanation required to answer **Qu. 1.6** (*Why do you think the graph reaches a maximum and then decreases to a minimum?*), a Level F task (drawing on own knowledge to formulate an explanation), would include references to thermodynamics (increase in temperature = increase in reaction velocity) as well as protein denaturation and subsequent loss of enzyme activity. Nine of the Top 10 group and only two of the Bottom 10 group were able to give a satisfactory explanation for the observed trend. This confirms that the Bottom 10 group of students really are representative of students who operate only at the basic level of graph comprehension – reading the data – and are incapable of any higher order cognitive processing or abstract reasoning 'beyond the data'. The nine students who gave unsatisfactory explanations resorted to describing the trend again along the lines of: "*I think the enzyme is, is no longer active as it reaches the minimum*"; or "*Basically it has reached the optimum point*"; or "*It shows the like optimum velocity of the enzyme and then after that it can't reach any other maximum and then slowly decreases*". One student referred to substrate depletion and another student referred to pH in different organs of the body,

Task 2 (which is Task 5 from Quiz #1) examined students' responses at two levels (A3 and D) and also assessed the classic slope/height confusion exhibited by physics students interpreting kinematics graphs. **Qu 2.1** (Level A3 – comparing slopes) was answered correctly by only two students, one from the Top 10 group (who also gave the correct explanation:- *at 20 sec the velocity of enzyme A is greater than that of enzyme B because the slope of A is greater than that of B*) and one from the Bottom 10 group. This result was surprising since four of the 19 students who correctly answered Qu 5.1 in Questionnaire #1 were included in this study (one in the Top 10 group and three in the Bottom 10 group) and not one of them (including the two who correctly answered both Qu 5.1 *and* 5.2) repeated their achievement from

Questionnaire #1! Conversely, the two students who gave the correct answer in this task could not do so in the first questionnaire. It appears as if none of the students recognised Qu 2.1 as being identical to Qu 5.1 in Questionnaire #1 and it might just as well have been an entirely new question, perhaps because the manner of presentation and testing environment differed markedly from that of Questionnaire #1. The balance of the students (except for two who answered 'I don't know') all exhibited the classic slope/height confusion detailed in the analysis of Qu 5.1 (Questionnaire #1) by answering *at t = 20 sec A is less than B because line B lies above line A*.

Qu 2.2 (Level D – reading a trend and noting that the slopes are not the same) was answered (*No, because slopes of lines A and B are not the same*) half-correctly by only three students, all from the Bottom 10 group, who all gave the incorrect explanation or gave no explanation. The other 17 students answered '*Yes, at t = 40 sec because the lines cross*' and did not realise that the two enzymes will not have the same velocity in the time frame of 0 to 100 seconds because the slopes of lines A and B are not the same. Instead, they chose t = 40 sec, the intersection point at which the graph lines have the same height, at a time when the *absorbances* are equal. The results of this task confirms the persistent presence of the by-now-classic slope/height confusion.

Task 3 examined students' responses at four levels (A3, A4, E and F) and analysed whether they could read between the data and read beyond the data. The correct answer for Qu. 3.1 (*Which sugar is decreasing in concentration?*) was "sugar a" (Level A4 – reading a simple trend). All of the students answered this question 100% correctly, a not unexpected result. The next question, Qu. 3.2 (*After what time interval are the concentrations of sugar a and sugar b equal?*), was on a lower level (Level A3 – compare information from two graphs) with an equally simple and direct answer (14 minutes) and should have been relatively easy for the students. Indeed, nine of the Top 10 group and six of the Bottom 10 group were able to identify the intercept of the two graphs. The one student from the Top 10 group who answered incorrectly identified the intercept as occurring at 17 min, the intercept of the graphs for sugar a and sugar c. The four incorrect responses by the Bottom 10 group were "*15 minutes*" (twice); "*28 minutes*"; "*30 seconds*" and "*at a concentration of 5*". The students who answered 30 seconds or 28 minutes were looking at the start and end of the graph and it could be surmised that they interpreted these zero-slope segments as signifying 'equal concentrations". It is unclear why the one student referred to 'a

concentration of 5' when the question asked for a time-related answer and there is no intercept of any graphs occurring at 5 g/l.

Qu. 3.3 (*Describe the observed results*) was a Level E task (extracting qualitative data from quantitative data) that did not appear to pose a problem for the students. Most of them (nine from the Top 10 group and six from the Bottom 10 group) were able to articulate a description along the lines of *"As sugar a is being degraded in the reactor by an enzyme, sugars b and c are being released "* simply by reading the information for the task and then critically comparing the data sets (the sugar concentrations). The one student from the Top 10 group who gave an incorrect description came up with this gem of obfuscation: *"Sugar a decreases um um with um with um um, how do you describe this, um um like um, is it a sigma graph, the one resembling like one that drops gradually at the end"*. It is clear that he was struggling to express himself in English, typical of most of the students. The other four students became confused with enzyme activity (*"The enzyme activity decreases with a decrease in sugar concentration over an increase in time."*) or gave long rambling explanations (*"Sugar a started at a higher concentration of b and c at zero minutes until it reaches a new concentration ……"* - for a total of 1 minute) or became totally confused (*"Sugar a concentration increased when time was increasing but as time gone by it stopped…"*) or regarded the results as individual experiments (*"I think that, um, the catalyst, the enzyme was not added at zero and five and then they started adding the enzyme whereby the reaction at b was faster than the reaction at c,"*).

The last question, **Qu. 3.4** (*If possible, describe the composition of sugar a in detail*), the most cognitively demanding task of the test, was a Level F task (the creation of a scenario to account for the data) with the most likely scenario being *"Sugar a (a trisaccharide or polymer) is composed of sugar b and sugar c with twice as much b as c."* This task should have been easy for senior undergraduate biochemistry students who only had to read and understand all the concentration and time information presented graphically, critically evaluate the relationships between the data sets, and then draw on their own knowledge related to carbohydrate biochemistry to articulate a plausible explanation for the observed trend. On the contrary, this task proved beyond the higher order cognitive abilities of 17 of the 20 students. Of the three students who responded positively (all from the Top 10 group) only one scored two points by correctly discerning that for every molecule of the monomeric sugar c there would be two molecules of sugar b combining to form polymeric *sugar a*. The other two students only observed that sugar a was composed of sugars b and c and did not

interpret the data further. Of the 17 students who failed to master this Level F task and give an appropriate answer, 10 answered in the negative ("*No, It's impossible to say*" or "*I don't know*") even after requesting that the question be reframed by the researcher. Three of the remaining seven students concluded that sugar a was a disaccharide (lactose, galactose or sucrose). One student resorted to a convoluted description of the graph: "*I think so. Sugar a is the only sugar which has the ok, the same concentrations after some time as the other two sugars, after 14 minutes it was similar to, it had the same concentration as sugar b so from sugar b and sugar c I think we can.*" The last three students could only hint at possible scenarios, all of them implausible, by responding "*I think sugar a, I don't know what to say about it, but I think it kinda like didn't have the enzyme at first, like the way I see it, it's like the reaction was faster at first but it declined as time went by so I think there was no catalyst.*"; or "*A does collide with b and c at different concentrations which means at those levels they must show some level of similarity of some sort.*"; or "*I think sugar a is constructed in, it's not constructed in a in a complex way, it's constructed in a simple way. I think it's the one that is deeper than b and c. It's more heavy than b and c.*" This last response was indicative of extremely deficient chemistry or biochemistry literacy (or even English literacy) as well as abstract thinking skills. .

A disturbing observation: only one student (#64) out of the 20 students, on the basis of his sophisticated answer as well as his correct response to Qu 1.6, can be said to be operating at an abstract level or at Kimura's Level F. This ratio is similar to that observed for responses at Level F for Questionnaire #1. Interestingly, student #64 was unable to correctly answer any Level F tasks in Questionnaire #1.

Eye tracking: Areas of interest

Based on heatmaps and gaze plots in real time as students were responding to on-sceen stimuli, it was decided to define the following Areas Of Interest (AOI's) and extract the mean gaze time (mean of total time of all fixations in each AOI) for each AOI:

      CAPTION – the title of the graph
      QUESTION – the task question at the bottom
      X CAPTION – the x-axis title
      X UNITS – the x-axis units
      Y CAPTION – the y-axis title

Y UNITS – the y-axis units

TL to BR – plot area divided into nine blocks ranging form TOP LEFT to BOTTOM RIGHT. Only data from the relevant blocks were used.

The results of the AOI analysis were used to compare the groups w.r.t. mean time spent by students extracting information from the different areas of the graphs, this approach being deemed more meaningful when assessing cognitive load than total number of fixations (Renshaw *et al.*, 2003).

Task 1

The AOIs identified for task 1 are depicted in Fig. 21 using Qu 1.4 as an example.
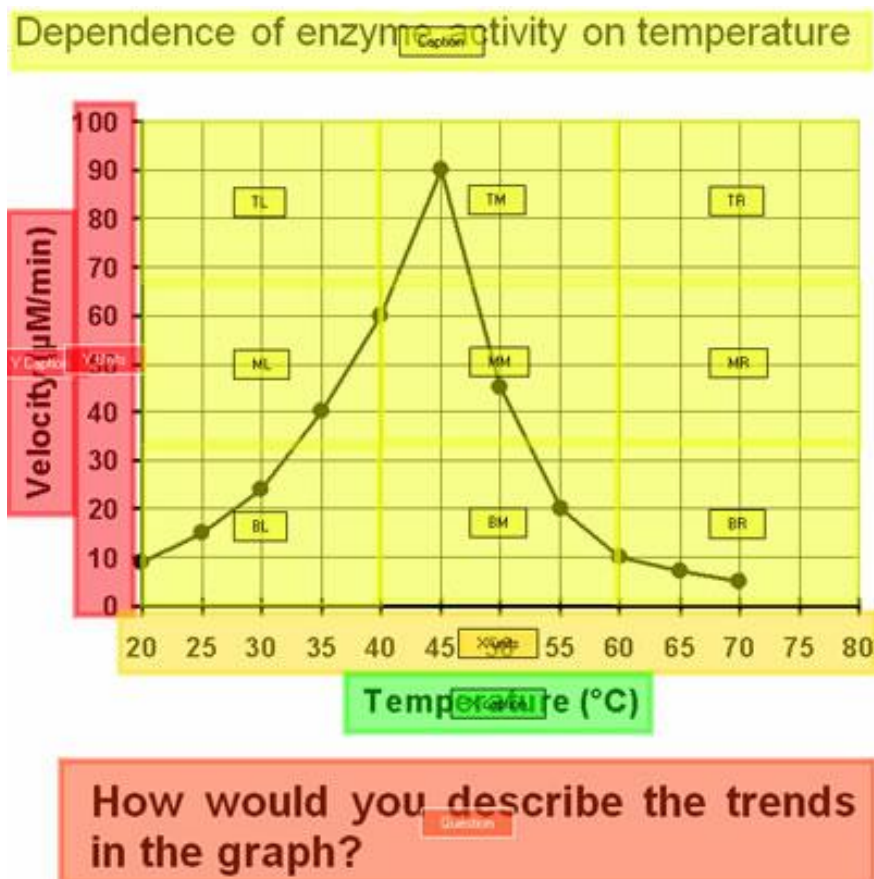
Figure 21. Areas Of Interest identified for Task 1 of Questionnaire #2.

The result of the AOI analysis for task 1 are presented graphically in Fig. 22. A brief analysis of each question in Fig. 22 follows:

Qu 1.1   More time was spent reading the question, almost equal for both groups, and this mean time was almost twice as much as that spent on the graph caption.

The Top 10 group spent more time on average on the Y-AXIS CAPTION AOI because they included the velocity units (μM/min) in their answer (the Bottom 10 group did not give units).

Mean time spent on the BL (Bottom Left) and BM (Bottom Middle) blocks (the sub-areas of the graph where students are most likely to look for an answer) for both graphs was almost equal since most students gave the correct answer
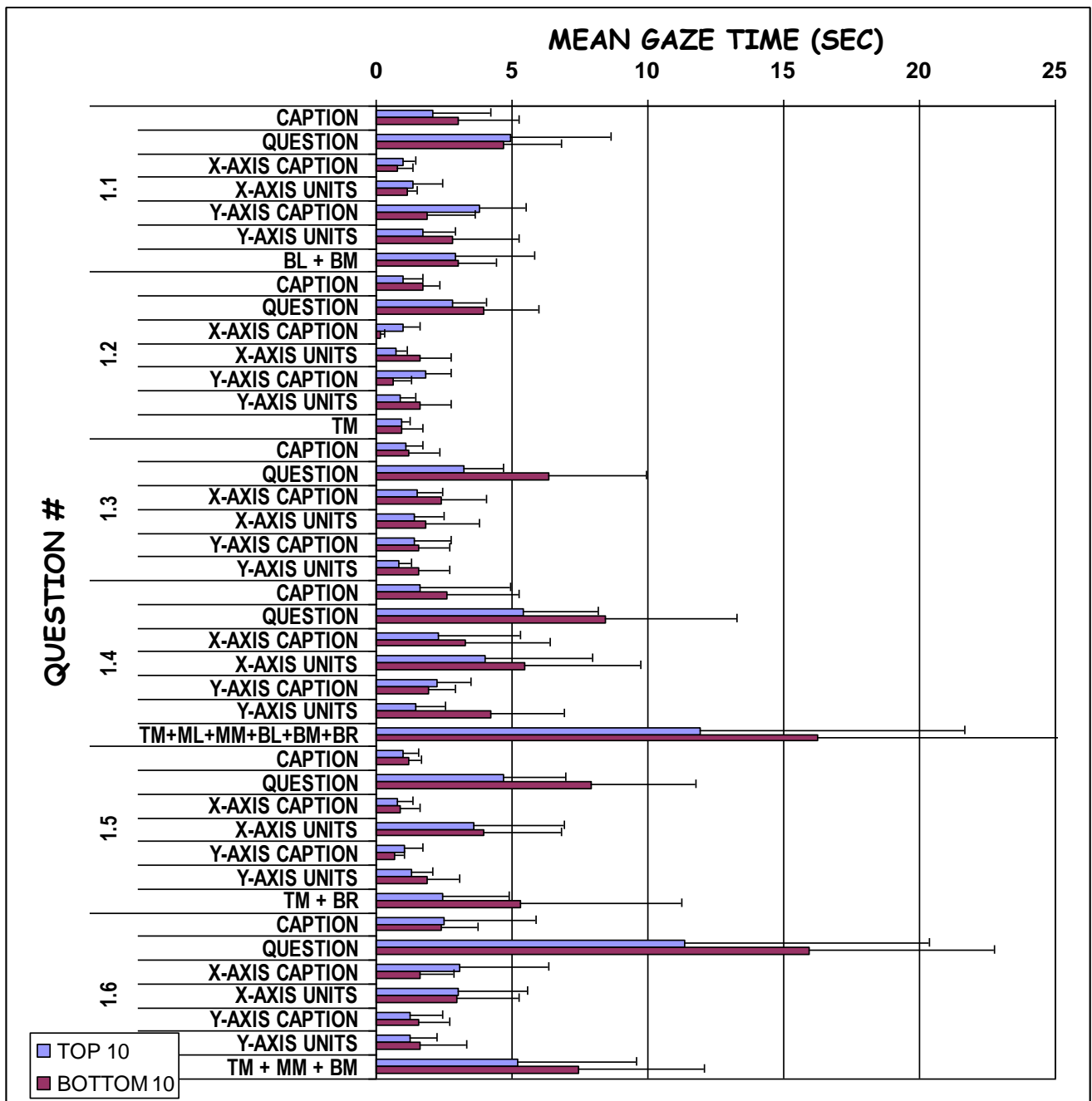
Figure 22. Mean gaze times per AOI and question number for Task 1.

Qu 1.2   Overall, the least amount of time was spent on this question (the easiest in this task).

More time was spent reading the question than the caption but the Top 10 group spent less time on both than the Bottom 10 group, perhaps because they understood the question and knew what was expected of them.

The Top 10 group spent more time on the Y-AXIS CAPTION AOI because they included the velocity units (μM/min) in their answer, the Bottom 10 group did not give units (same students in each case as in Qu 1.1);

Time spent on the TM (Top Middle) block (the sub-area of the graph where students are most likely to look for an answer) for both graphs was almost equal since most students gave the correct answer.

Qu 1.3   More time was spent reading the question than the caption for both groups with both groups spending an almost equal amount of time on the CAPTION AOI.

The Bottom 10 group spent twice as much time on the QUESTION AOI than the Top 10 group even though correct responses were almost equal for the two groups.

Qu 1.4   Considerably more time was spent reading the question than the caption for both groups with the Bottom 10 group spending about a third more time than the Top 10 group on the QUESTION AOI. Both groups struggled with the term 'trends' and were not initially sure of what was expected of them.

The Bottom 10 group spent about a third more time on the relevant graph area than the Top 10 group trying to find a trend.

Qu 1.5   Considerably more time was spent reading the question than the caption for both groups with the Bottom 10 group spending about 75% more time than the Top 10 group on the QUESTION AOI.

Both groups spent almost equal amount of time on X-AXIS UNITS AOI looking at the area around 70 - 80°C before predicting a trend towards 90°C.

The Bottom 10 group spent about twice as much time as the Top 10 group on the TM (Top Middle) and BR (Bottom Right) blocks (areas of the graph identified from the heatmaps) in order to predict a trend..

Qu 1.6   This considerably more challenging question resulted in the most time being spent on the QUESTION AOI than other AOIs (also the most of all of the questions in this task) with the Bottom 10 group spending about a third more time than the Top 10 group trying to decipher the question.

The Bottom 10 group spent about a third more time than the Top 10 group on the TM (Top Middle), MM (Middle Middle) and BM (Bottom Middle) blocks (areas of the graph identified from the heatmaps), looking for clues even though this is an abstract Level F. In fact, raw data showed that the Bottom 10 group spent twice as much time on the TM (Top Middle) block AOI (the highest part/maximum of the graph) than the Top 10 group, a result reflected in their incorrect answers referring to a description of 'maximum' or 'optimum' temperature.

Task 2

The AOIs identified for task 2 are depicted in Fig. 23 using Qu 2.2 as an example.
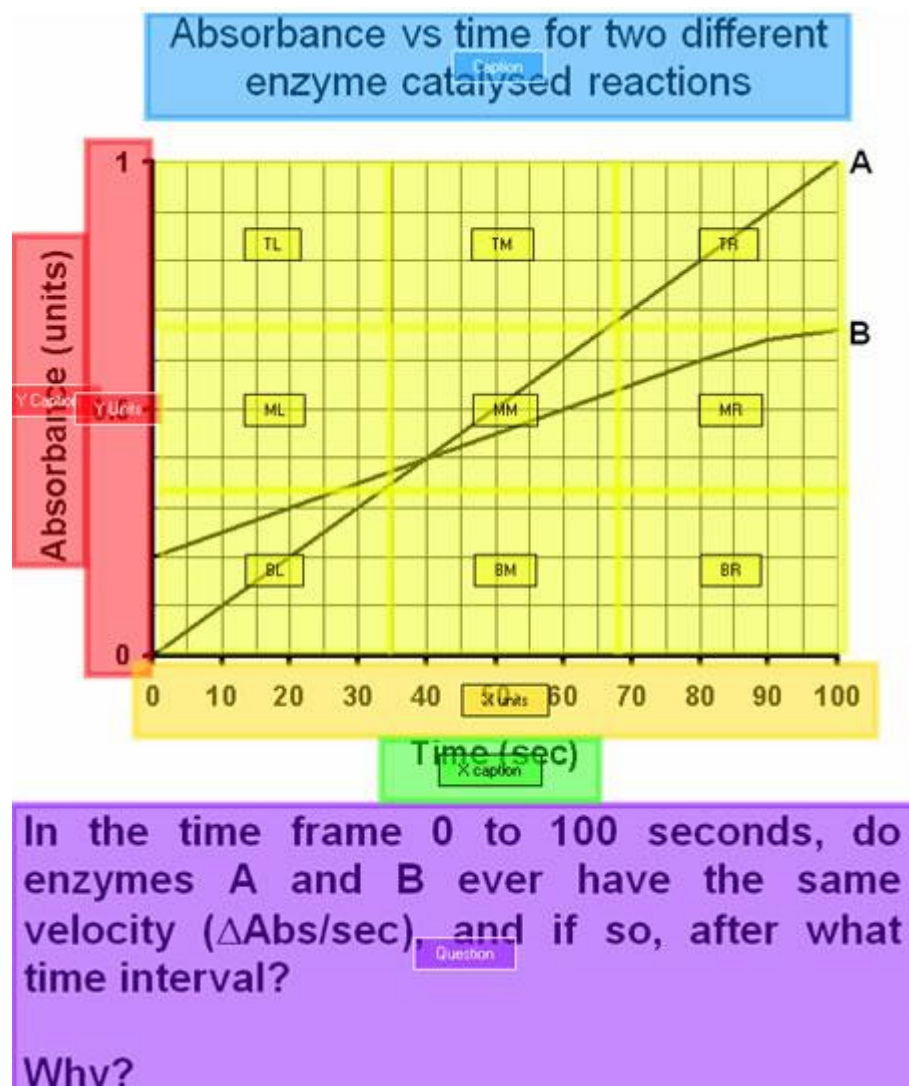
Figure 23. Areas Of Interest identified for Task 2 of Questionnaire #2.

The result of the AOI analysis for task 2 are presented graphically in Fig. 24 followed by a brief analysis of each question in Fig. 24.
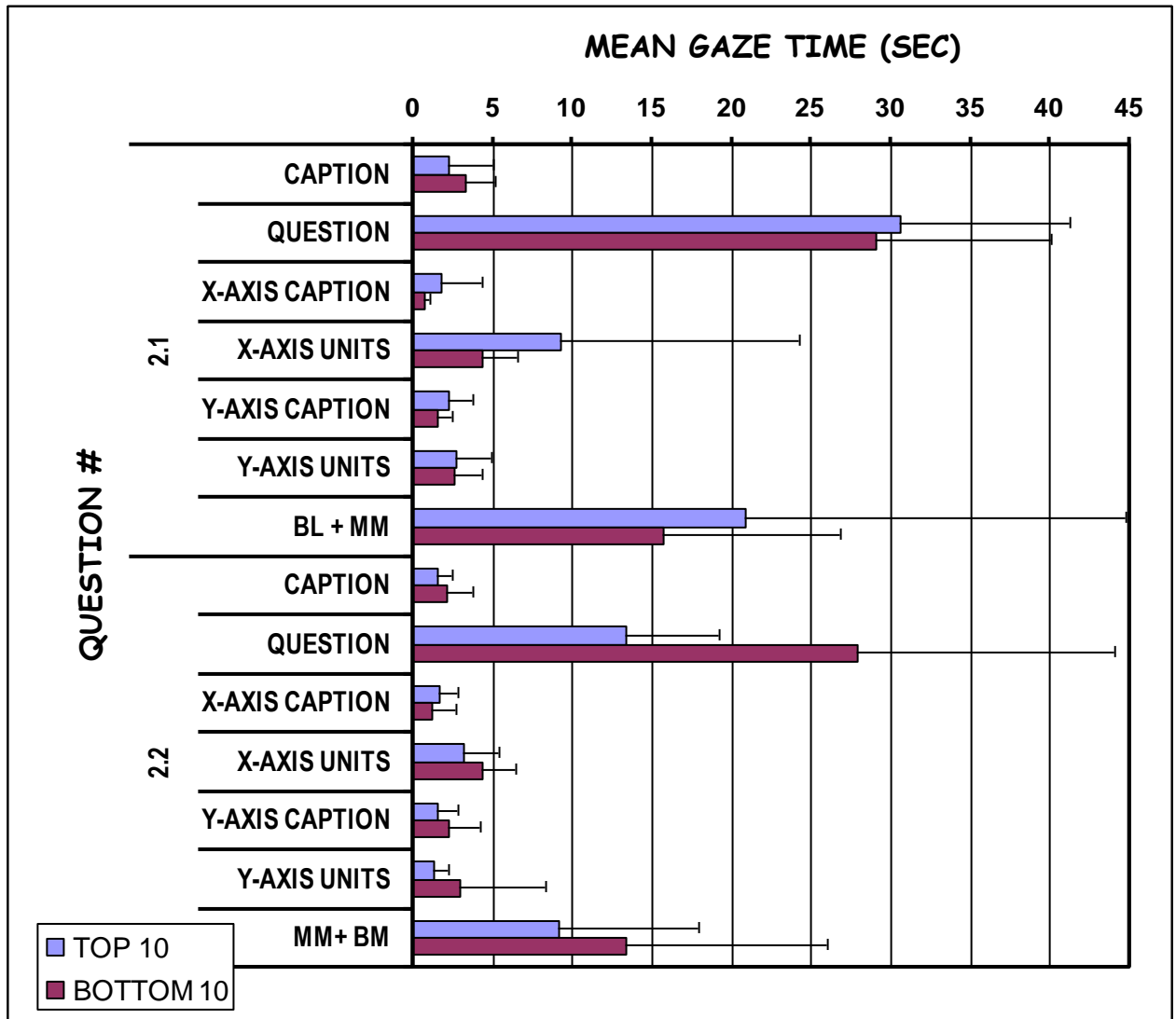
Figure 24. Mean gaze times per AOI and question number for Task 2.

Qu 2.1   More time on average was spent reading the question than looking at any other area of the graph, an indication that the students from both groups were unsure of what was expected of them. In contrast to the other tasks and questions, the Top 10 group spent about 40% more time on the QUESTION AOI than the Bottom 10 group for this question, although this might be as a result of one student taking much longer than the other students.

The Bottom 10 group spent marginally more time reading the caption than the Top 10 group.

The Top 10 group spent about 30% more time on the MM (Middle Middle – where graphs intersect) and BL (Bottom Left - where 20 seconds is found) block AOIs than the Bottom 10 group.

The Top 10 group spent about twice as much time on the X-AXIS UNITS AOI (where 20 sec is found on the x-axis) than the Bottom 10 group.

On average, the Top 10 group spent more time on the AOIs (except for the CAPTION AOI) than the Bottom 10 group but still delivered the same number of incorrect responses as the Bottom 10 group.

Qu 2.2   On average, the Bottom 10 group spent more time on all of the AOIs (except for the low X-AXIS CAPTION AOI) than the Top 10 group but delivered the only three semi-correct responses.

Similar to Qu 2.1, more time on average was spent reading the question than looking at any other area of the graph, an indication once again that the students from both groups were unsure of what was expected of them.

About twice as much time was spent by the Bottom 10 group on the QUESTION AOI than that spent by the Top 10 group.

The Bottom 10 group spent about 50% more time on the MM (Middle Middle) and BM (Bottom Middle) block AOIs than the Bottom 10 group. These blocks include the intersect of the two graph lines as well as the area connecting the 40 second value on the x-axis and the intersect.

Very little time was spent on other areas of the graph, for example MR (Middle Right) and TR (Top Right), where the slopes of the graph lines could be verified as differing: - a confirmation of slope/height confusion.

Task 3

The AOIs identified for task 3 are depicted in Fig. 25 using Qu 3.1 as an example. Note that an additional AOI, 'SUGARS', to include the annotations for each graph line has been added to the right of the graph.
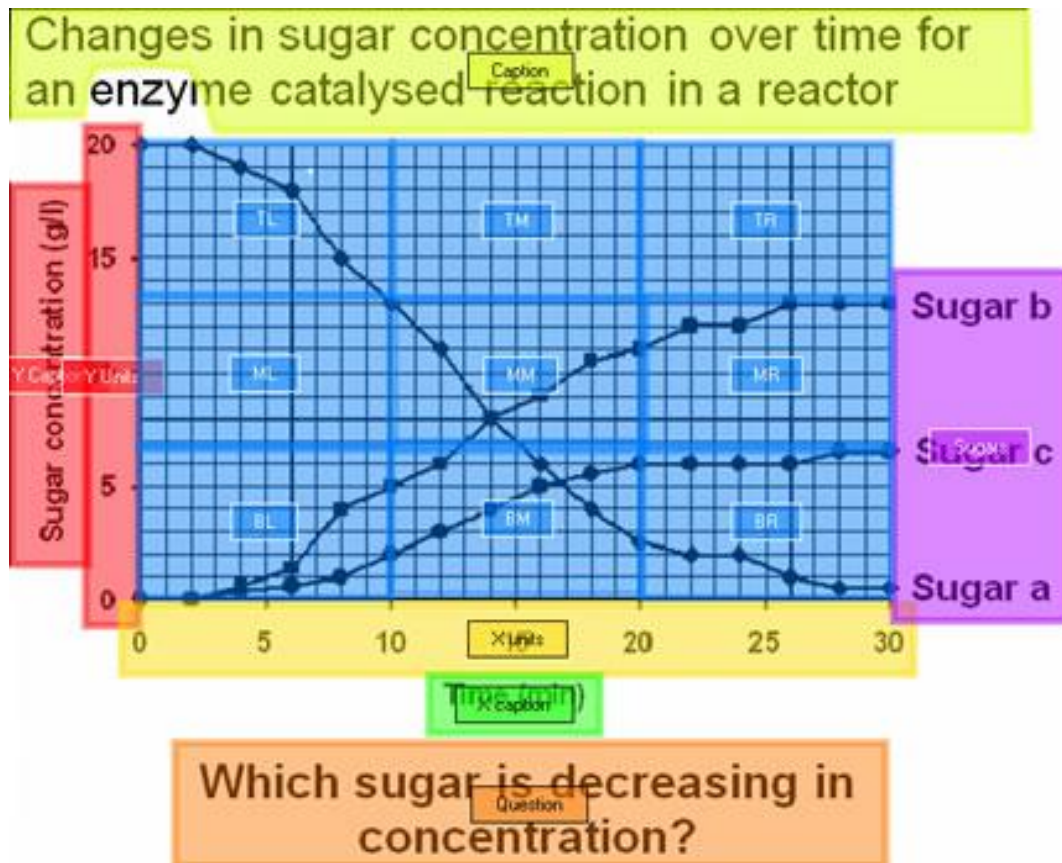
Figure 25. Areas Of Interest identified for Task 3 of Questionnaire #2.

The result of the AOI analysis for task 3 are presented graphically in Fig. 26 followed by a brief analysis of each question in Fig. 26.

Qu 3.1   The QUESTION AOI once again received the most attention by the students with similar times being recorded by both groups

The Bottom 10 group spent about twice as much time as the Top 10 group on the CAPTION AOI, perhaps to orientate their thinking patterns.

Identifying the graph line with a negative slope (Sugar a) requires looking at the top left and bottom right (TL + BR) of the graph including the annotations. The Bottom 10 group spent about 50% more time looking in these areas than the Top 10 group.

Since this question was relatively easy and the only question that resulted in 100% positive responses it is not surprising that the mean gaze times were relatively lower than for the rest of the questions in this task.
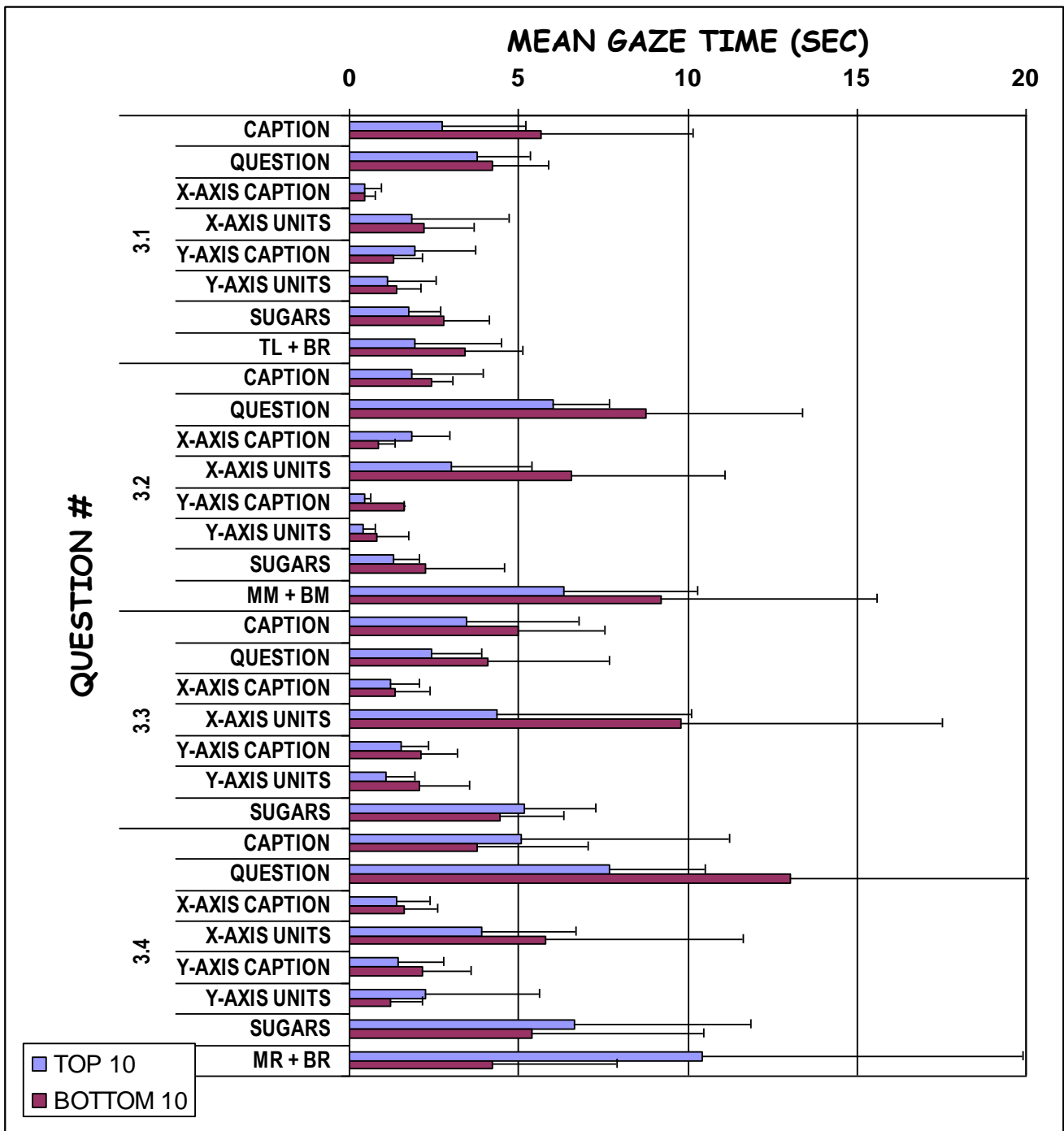
Figure 26. Mean gaze times per AOI and question number for Task 3.

Qu 3.2   The Bottom 10 group spent about 50% more time on reading the question than the Top 10 group. Since the Bottom 10 group fared worse than the Top 10 group with this question (six versus nine correct responses, respectively), it is not surprising that they spent more time on the QUESTION AOI.

The highest mean gaze time was calculated for the MM (Middle Middle) and BM (Bottom Middle) block AOIs since these blocks include the

intersect of the sugar a and sugar b graph lines as well as the area connecting the 14 minute value on the x-axis and the intersect.

Qu 3.3 The Bottom 10 group spent marginally more time on the graph title (CAPTION AOI) and question than the top 10 group and as a whole fared slightly worse on this description question.

The Bottom 10 group seemed to spend more time looking at the x-axis (X-AXIS UNITS AOI) pinpointing the start and end of the graph lines while formulating a description of the graph.

Both groups spent similar times looking at the SUGARS AOI while they identified the three graph lines.

Graph area block AOIs were not calculated for this question since students' gaze covered the whole graph area as they tracked the graph lines.

Qu 3.4 This, the most challenging question in Questionnaire #2, proved to be very revealing. Both groups spent a lot of time reading and re-reading the question, the most time of any question in Task 3, with the Bottom 10 group spending almost twice as long as the Top 10 group on this AOI (a confirmation of their confusion).

Although both groups spent similar times on the SUGAR AOI, with the Top 10 group spending more time, the time spent on the MR + BR (Middle Right + Bottom Right) block AOIs is more revealing. None of the Bottom 10 group were able to give a satisfactory response to this question and they did not seem to take into account the final concentrations of sugars *b* and *c* in their explanations, borne out by the fact that the time spent on the MR + BR blocks was less than half that spent by the Top 10 group (three of whom gave satisfactory responses) on these blocks. In fact, raw data showed that the Top 10 group spent four times as much time on the MR block AOI (where the graph line for sugar *b* ends) than the Bottom 10 group and were using this data in formulating their answers.

If just the QUESTION AOIs are compared graphically, as in Fig. 27, then it becomes even more apparent that the Bottom 10 group of students are struggling to comprehend the task question.
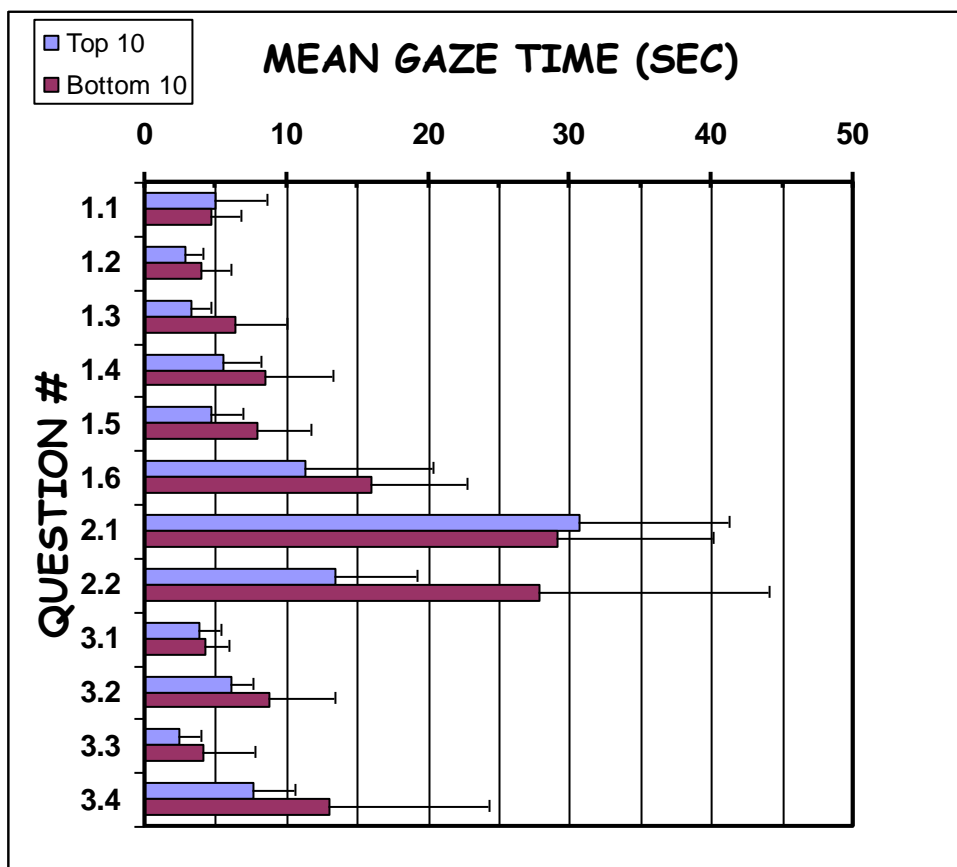
Figure 27. Mean gaze times per QUESTION AOI for Questionnaire #2.

Interestingly, mean gaze times for the CAPTION AOI were, with the exception of Qu 3.3, always less than the QUESTION AOI. This was unexpected as the CAPTION AOI contains the title or heading of the graph and should anchor viewers in a data set.

All of the preceding quantitative data analysis of the AOIs can be visualised very elegantly in the form of heatmaps, the final subsection in this section.

Eye tracking: Heatmaps

Heatmaps visualise AOI data and can be regarded as an aggregate representation showing what areas of visual attention a group of people fixated on when viewing a diagram. Heatmaps show maximum fixations as red '"hotspots" with a gradient decreasing to green as fixations decrease. Heatmaps for each task and a discussion follows:

The heatmaps for all of the questions in **Task 1** are shown in Figs. 28 and 29.

Inspection of the heatmaps for **Qu 1.1** reveal very little difference between the two groups in fixations on the question and one can conclude that both groups of students understood the question. The only real difference is the y-axis title where the Top 10 group spent more time looking at the actual velocity units (they gave units as part of their answer while the Bottom 10 group of students omitted the units). The keyword "55°C" in the question was a 'hotspot' that received the most fixations.

The phrase and keywords "maximum reaction velocity" received the longest fixations in the QUESTION AOI for both groups. Once again, the Top 10 group spent more time looking at the velocity units in the y-axis title than the Bottom 10 group (who omitted the units as part of their answer). The Bottom 10 group spent more time (and fixations) on the maximum (the peak) of the graph line and on the corresponding y-axis value (90) than the Top 10 group before giving an answer, possibly a result of lack of confidence.

It is clear from the heatmap for **Qu 1.3** that the concept of dependency of data is a big problem for the Bottom 10 group. These students did not spend a lot of time looking at the graph, instead they spent a lot of time reading the keyword "dependent" in the question and also the keyword "Temperature" in the x-axis title. In addition, there were a lot more fixations on other areas of the graph, for example the axes and start of the axes, than the Top 10 group. It appears as if the Bottom 10 students, after reading and rereading the question and keyword, were looking for clues as to the identity of the dependent elsewhere on the graph. The Top 10 students spent little time on the graph and answered this question from memory. The possibility also exists that the Top 10 group of students were able to retain all three question items (velocity, temperature, 55°C) while solving the problem and did not have to return to the question to refresh their memories. This is an indication that their might be short-term memory (or working memory) processing differences between the two groups of students with the Bottom 10 group exhibiting working memory limitations, an observation confirming the findings of Peebles & Cheng (2002).
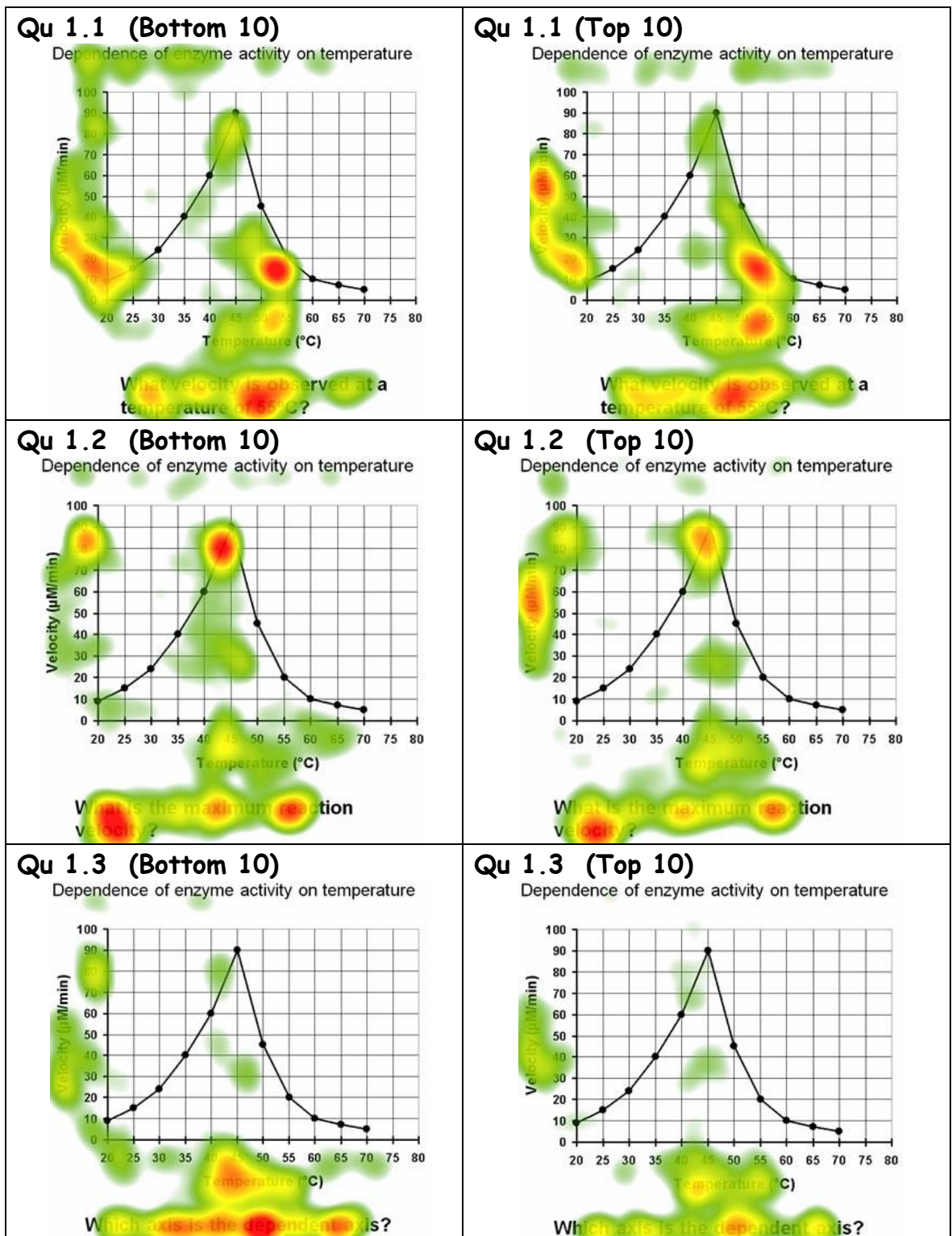
Figure 28. Heatmaps for questions 1.1 to 1.3 of Task 1.

For **Qu 1.4** (Fig. 29), the Bottom 10 students spent a lot more time than the Top10 group on the question (including the keyword "trends"), the x-axis title, the starting point of the graph line and the maximum point (peak) of the graph line. The heatmap suggests that the Top 10 students cursorily scanned the graph, found the information that they were looking for and then articulated an answer.
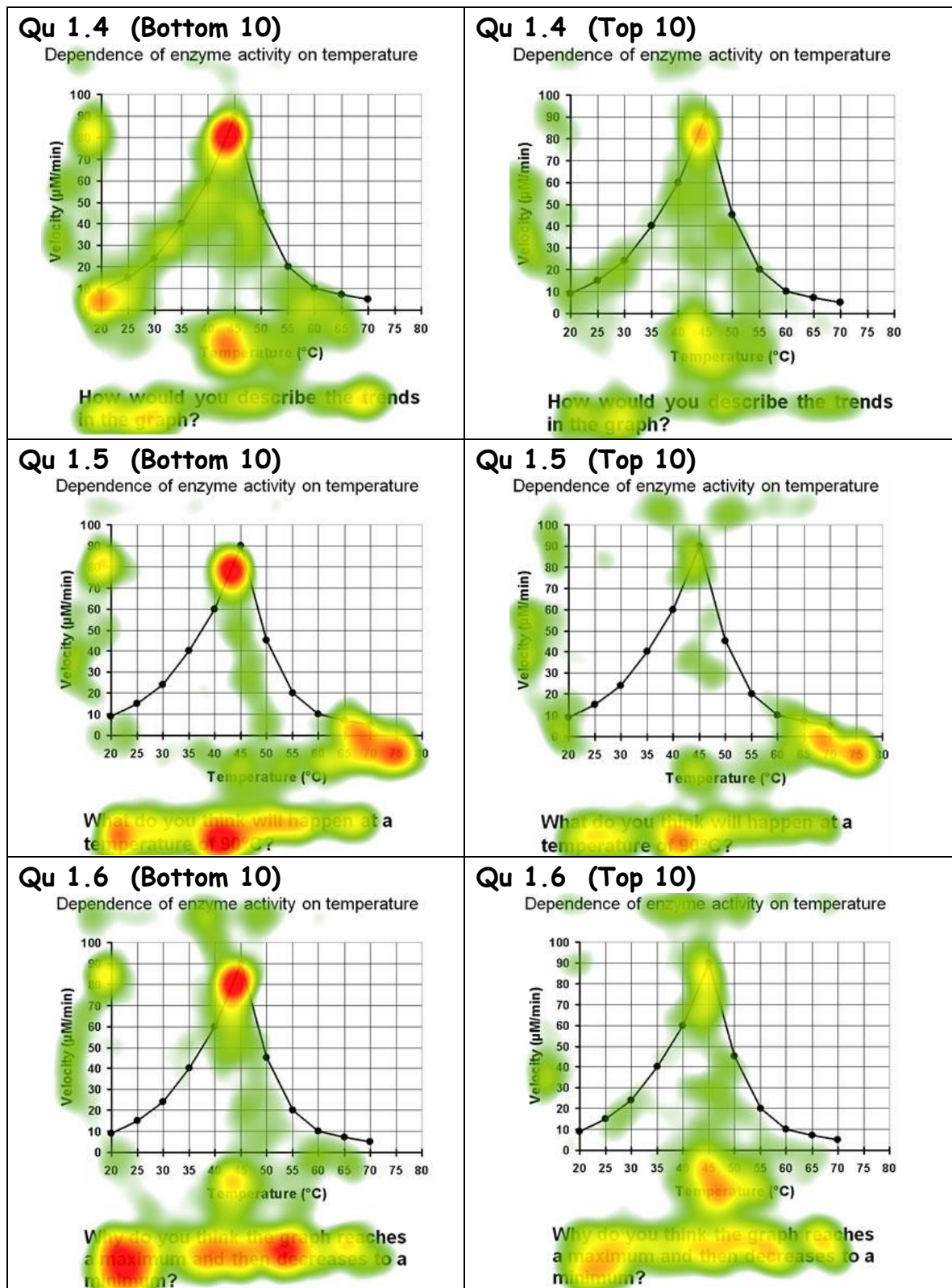
Figure 29. Heatmaps for questions 1.4 to 1.6 of Task 1.

The heatmap for **Qu 1.5** (Fig. 29) once again shows the Bottom 10 students concentrating on keywords (temperature, 90°C) in the question in an effort to clarify the task (make a prediction). Their heatmap also differs considerably from that of the top 10 group of students in the sense that they appear to have become confused on reading "90°C" and quickly scanned the graph axes for a "90", found it on the y-axis,

extrapolated across to the graph line peak, and then, one could surmise, realising that this was not the answer, tracked back down to the x-axis and then to the right where they scanned the temperatures as well as the graph line before formulating a prediction.

The heatmap for **Qu 1.6**, the Level F task, showed some interesting features. Nine of the Top 10 group and only two of the Bottom 10 group were able to give a satisfactory explanation for the observed trend and this lack of abstract thinking is evident from the heatmap. The Bottom 10 group of students spent a lot of time reading and rereading the question keywords 'maximum', 'decreases' and 'minimum' (an apparent effort to understand what was required of them) and also spent a lot of time looking at the peak of the graph line in search of an answer even though the answer, being a Level F task, is not apparent in the data and must be formulated from their own understanding and prior knowledge. The Top 10 group spent a lot less time on the question (although they read the same keywords), a little more time on the x-axis title (temperature) and, based on what they observed, were able to formulate an explanation. This is another confirmation that the Bottom 10 group of students are incapable of any higher order cognitive processing or abstract reasoning 'beyond the data'.

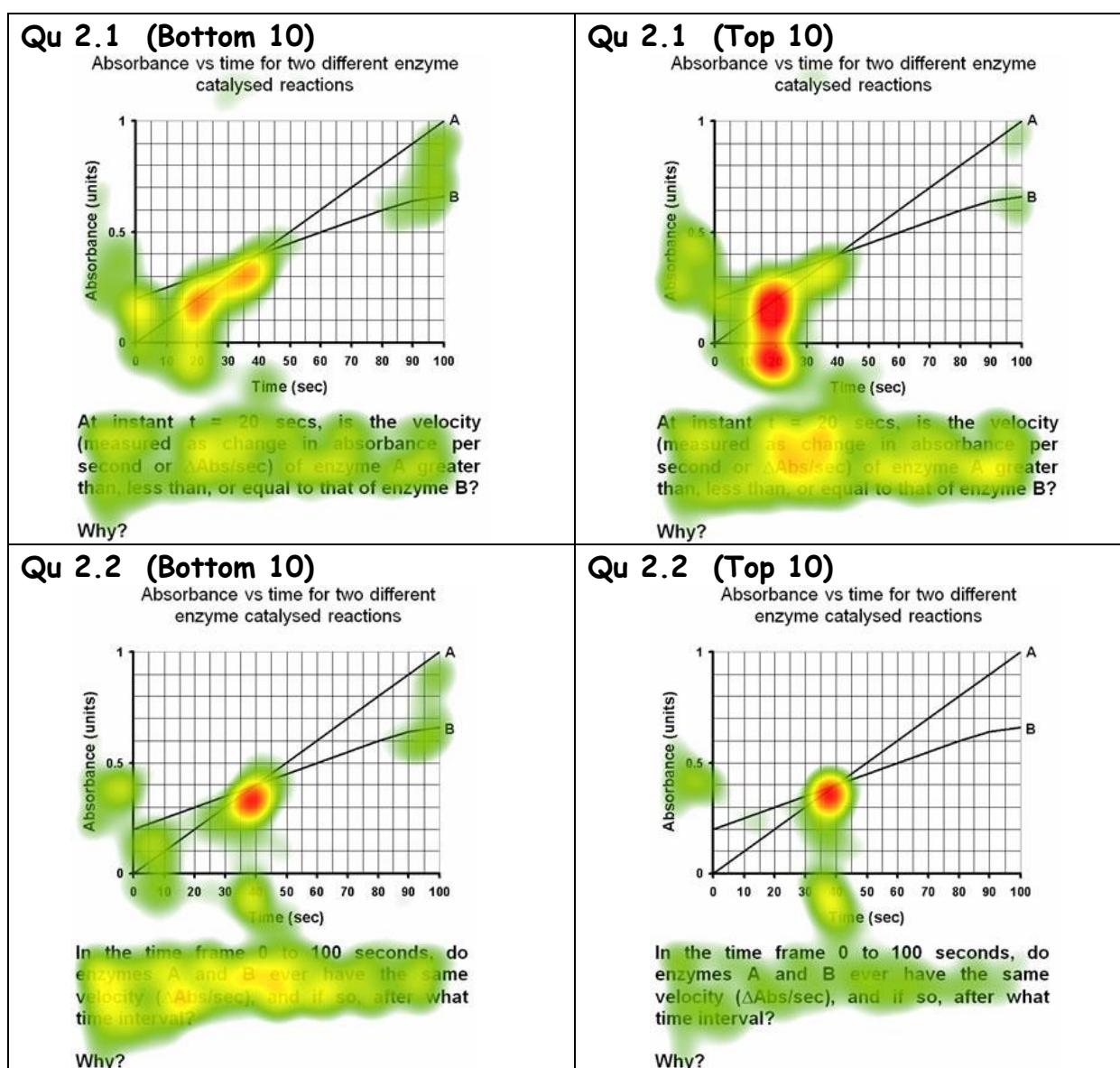The heatmaps for the questions in **Task 2** are shown in Fig. 30.

Figure 30. Heatmaps for Task 2.

For **Qu 2.1**, both groups spent a lot of time reading the question and hardly any at all reading the title of the graph. The keywords that received the most attention from both groups were 'change' and 'ΔAbs/sec', both referring to slope. Some time was spent on the x- and y-axis titles and a small amount of time was spent on the end of the graph lines A and B (the difference between which clearly indicates a difference in slope). The most prominent hotspot, however, was centred over the segment of graph line A corresponding to the x-axis value of 20 sec and the area connecting the two with the Top 10 group spending a lot more time on this area than the Bottom 10 group. Although the heatmaps indicate that time was spent following the first half of both graph lines, it is clear from the answers given that it was not the slope being observed

by the students but rather the height difference between the two graph lines that was making an impression.

The heatmaps for **Qu 2.2** could be regarded as a classic confirmation of slope/height confusion without ever resorting to questionnaires or AOI analysis. The intense hotspot over the intercept of the two graph lines and the second hotspot over the 40 sec on the x-axis says is all that is required to draw this conclusion No time was spent following the graph lines to compare their slopes and the gazes of the students went straight to the intercept at 40 seconds. Interestingly, although the Bottom 10 group seemed to spend more time reading the question (popular keywords were 'B ever' and 'ΔAbs') than the Top 10 group, it was the Bottom 10 group students that were able to deliver the only three correct responses. This might explain the 'cooler' hotspots around the end of graph lines A and B in the top right area of the graph.

The heatmaps for the questions in **Task 3** are shown in Fig. 31.

Since **Qu 3.1** was relatively easy and the only question that resulted in a 100% positive response for both groups of students, it is interesting to see that differences exist between the two heatmaps. The most obvious difference is in the intensity of the hotspots, which although in the same areas, are much more intense for the Bottom 10 group. This is an indication of longer fixation time in these areas than the Top 10 group and can be confirmed by consulting the AOI data. Both groups spent more time on the sugar a graph line than the other two graph lines. The Bottom 10 group took longer to answer this question and the reason is obvious from the heatmap: they scanned almost the entire graph area, including the sugar *b* and sugar *c* graph lines, decoding all available data before formulating an answer. The Top 10 group concentrated on the sugar *a* graph line with only minimal attention paid to the other graph lines; they knew what was expected of them, located and identified the necessary data, and formulated their answers.

Since the Bottom 10 group fared worse with question **3.2** than the Top 10 group (six versus nine correct responses, respectively), it is not surprising that they spent more time reading the question, in particular the keywords '*sugar a*' and '*sugar b*'. A noticeable difference between the two heatmaps is in the relative size and intensity of the one major hotspot over the intercept of the graph lines. In the case of the Top 10 group the hotspot is smaller and only lies over the intercept of the sugar *a* and sugar *b* graph lines, not the sugar *a* and sugar *c* intercept. The Bottom 10 group of students

seemed to spend an equal amount of time on both intercepts (hence one large hotspot) and also seemed to go back and forth between the intercepts and the x-axis, perhaps confirming their answer before articulating it.
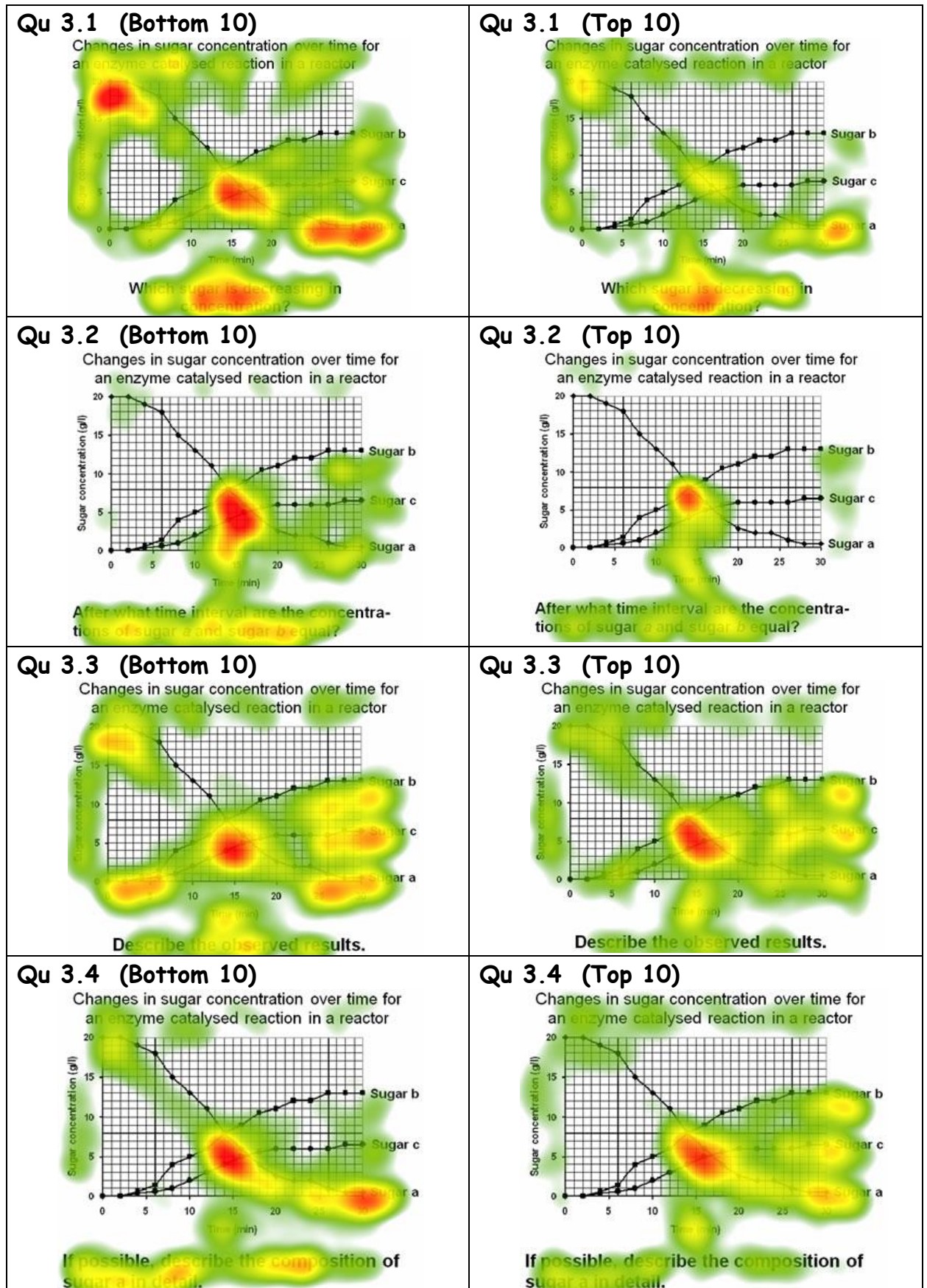

Figure 31. Heatmaps for Task 3.

The heatmaps for **Qu 3.3** showed that both groups scanned the start and end of all three graph lines as well as their intercepts in order to formulate a description of the graph. The Bottom 10 group spent more time looking at the question and the start of the sugar graph lines than the Top 10 group who, in turn spent more time looking at the intercepts of the three graph lines.

Qu 3.4 was the most challenging question in Questionnaire #2 and the heatmaps for this question proved to be very revealing. Both groups spent a lot of time reading and re-reading the question, in particular the question elements (or keywords) '*describe*', '*composition*' and '*sugar a*'. However, AOI analysis indicated that the Bottom 10 group spent almost twice as long as the Top 10 group on this AOI and these hotspots therefore confirm a suspicion that the Bottom 10 group were confused as to what was required of them (a comprehension problem) or required constant refreshing of short-term memory (which was fading as processing time lengthened). None of the Bottom 10 group was able to give a satisfactory response to this question and they did not seem to take into account the final concentrations of sugar *b* and *c* in their explanations. This observation was confirmed by the fact that the Bottom 10 group barely looked at the sugar *b* and sugar *c* graph lines (and their final concentrations) and instead concentrated on the sugar *a* graph line. On the contrary, the Top 10 group of students did look at the SUGARS AOI and all of the graph lines and were able to formulate answers containing some or other reference to sugars b and c, even though only three students gave a satisfactory explanation for sugar a's composition, The Top 10 group also seemed to spend more time reading the title of the graph than the Bottom 10 group. The most likely explanation for these results is that the Bottom 10 group of students are exhibiting a lack of comprehension (at an abstract level) of what is required of them to solve the task problem. These students spend more time examining text than interpreting the visual (or graphic) representation of the data. Alternatively, the additional integration task imposed by asking the students to describe the composition of sugar a (applying prior knowledge) has imposed a load on working memory which resulted in frequent, time-absorbing, revisits to both the data and question areas.

The clearly visible differences in both location and average duration of fixations illustrated by heatmaps have proven to be a valuable tool with respect to analysing differences in graph comprehension between the two groups of students. Overall, all of the students spent the majority of their time reading and re-reading information from

the axes and question areas of the graph and not looking at the pattern of the graph lines. Comprehension of these visual displays is thus easier when complex cognitive processing is avoided and emphasis placed on simple pattern identification processes, a conclusion previously reported by Carpenter and Shah (1998).

Eye tracking: Gaze plots

Gaze plots (or an eye movement pattern) for individual students can be used to confirm the observations discussed in the previous section on heat maps and, when analysed on an individual basis (by using the time-line feature), gave very revealing insights into the cognitive processes of the students. However, this study resulted in 20 x 12 = 240 gaze plots which could be used to confirm the optimal sequence of eye movements and fixations for retrieving the required information for solving a graphical problem – a cognitive science approach that was beyond the scope of this project.

Two gaze plots for individual students were retrieved, however, to illustrate striking differences in graph-decoding strategies used by students in answering Qu. 1.5. The gaze plots (shown in Fig. 32) for a weak (Bottom 10 group) student (#17), who was unable to provide an answer (response time = 27 seconds), and a Top 10 group student (#55), who made the correct prediction (response time = 14 seconds), showed that student #17 (orange circles – size relative to time spent or fixation duration in that area) returned to re-read the question 6 times (each time for 3 seconds) while student #55 (blue circles) only read the question twice for less than 2 seconds each time. In fact, student #55 spent more time on the graph area and the caption (reading four times) than reading the question (student #17 read the caption once). Fig. 32 also shows that student #55 followed the trend of the graph line before making a prediction (the last fixation is in the lower right quadrant of the graph). Student #17 spent a lot of time scanning the axes of the graph searching for an answer and her last fixations were on the x-axis labels and the question before she gave up and guessed incorrectly "*You don't have substrates anymore…*" The scan path for student #17 is characterised by almost double the number of saccades than that of student #55, a possible indication of high cognitive load for this student. This pattern seemed to be followed by all of the students in the two groups and is confirmed by the hotspots (derived from an aggregate of gaze plots) in Fig. 29: the hotspots for the Top 10 group consisted only of the question keyword (temperature 90°C) and the graph line plus x-axis values of 70 - 80°C while the hotspots for the Bottom 10 group were concentrated on the graph maximum, the question and the lower right quadrant of the graph (where

90°C is expected to be). The relatively larger fixations of student #17 (characterised by larger, orange circles) confirms the cognitive processing argument of Goldberg & Kotval (1998) which states that if an object is difficult to encode, then the fixation will last longer (due to processing taking longer) than if the object is easy to encode.
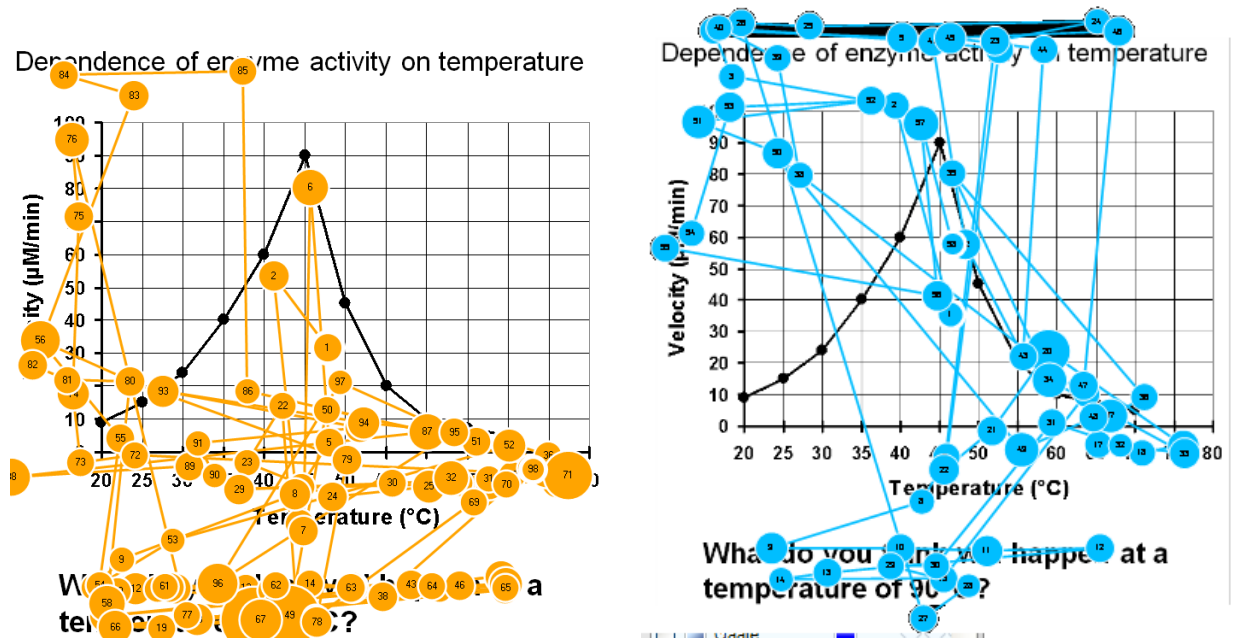

Figure 32. Gaze plots obtained for student #17 (orange) and student #55 (blue) while engaged in answering question 1.5 of Task 1.

Simply recording response times and visual scan paths of individuals does not take important cognitive factors such as working memory limitations or strategic decisions into account (Peebles & Cheng, 2002). For example, it is likely that, during the time course of the complex graph-based reasoning problem posed by question 1.5, certain key information (such as the question variables) may be forgotten and have to be rescanned. In addition, given that student #17 is aware that information is available for rescanning at all times, it is possible that she may trade off additional saccades for a reduction in working memory load. This becomes a distinct possibility when one compares the gaze plots of student #17 and student #55.

Since the possibility has now been raised that the Bottom 10 group of students, and student #17 in particular, could be experiencing graph-decoding problems as a result of working memory limitations, it was decided to compare the gaze plots and scan paths obtained by one Bottom 10 group student and one Top 10 group student for a lower level task that both students answered correctly. The task chosen was question 1.1 (a level A1 task) and the gaze plots for student #1 (the student with the lowest graphical literacy) and student #34 (Top 10 group) are shown in Fig. 33.
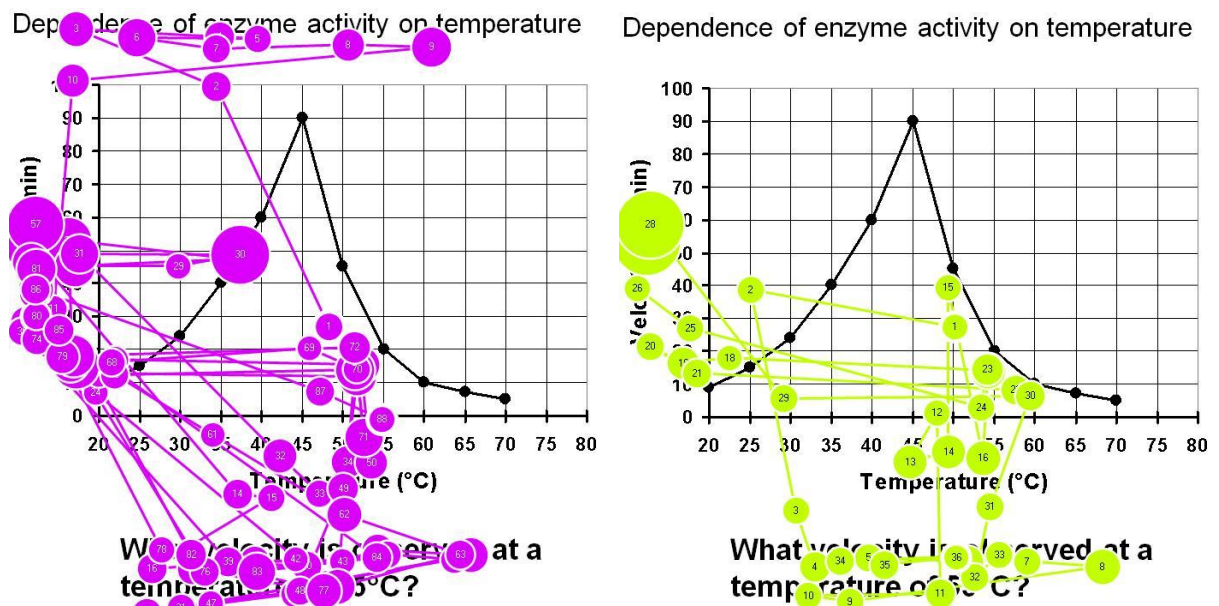
Figure 33. Gaze plots obtained for student #1 (mauve) and student #34 (green) while engaged in answering question 1.1 of Task 1.

The scan path for student #1 was: caption → y-axis → y-axis caption → x-axis → question (read fully) → y-axis value 55 → graph → y-axis → x-axis → graph → y-axis (correct answer of 20) →graph → question → graph → y-axis → x-axis → y-axis → question → graph → question → y-axis → question → y-axis → x-axis; (answered question and ended at this point after 27 seconds).

The scan path for student #34 was: middle of graph → question → x-axis → orientate on x-axis to 55 → graph → y-axis (value of 20) → graph → y-axis → y-axis units → y-axis → graph → question (for confirmation); (answered question and ended at this point after 11,5 seconds). The caption information was not accessed at all.

It is clear that student #34 followed an optimal sequence of fixations to reach the *given location* in the graph representing the *given value* (55°C) of the *given variable* (temperature) and then from there to the *target location* on the y-axis representing the corresponding value (20) of the *required variable* (velocity). This sequence was followed by a sequence to confirm the answer. This suggests that student #34 was able to retain all three question items while solving the problem.

Student #1 followed a sequence of fixations that required five rescans of the question, in particular the required variable 'velocity' (as evidenced by the large number of fixations on this area of the question). This non-optimal sequence requiring recall of

the required variable (the last variable in the question and the one most likely to be forgotten) may be interpreted as an indication of the effect of working memory limitations (according to the GBR model of Peebles & Cheng, 2002).

The same strategy was followed by student #1 and student #34 in question 1.2, a level A2 task answered correctly by both students. Gaze plots are shown in Fig. 34 and confirm the conclusion reached with question 1.1 regarding working memory limitations exhibited by student #1 who had to rescan the question twice before answering '90' (the maximum reaction velocity). These two tasks are at a level (A1 and A2) that require no prior subject-specific knowledge or any abstract reasoning; only the most basic 'graph sense' is required and it is therefore reasonably safe to speculate that the only variable accounting for the difference in fixation sequences between the two students is, in fact, working memory capacity.
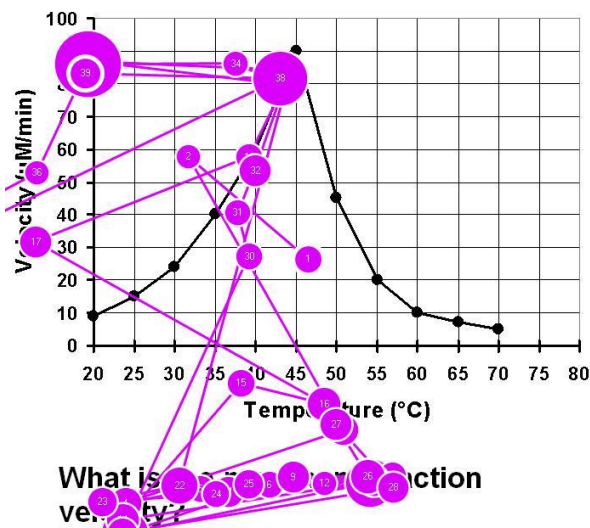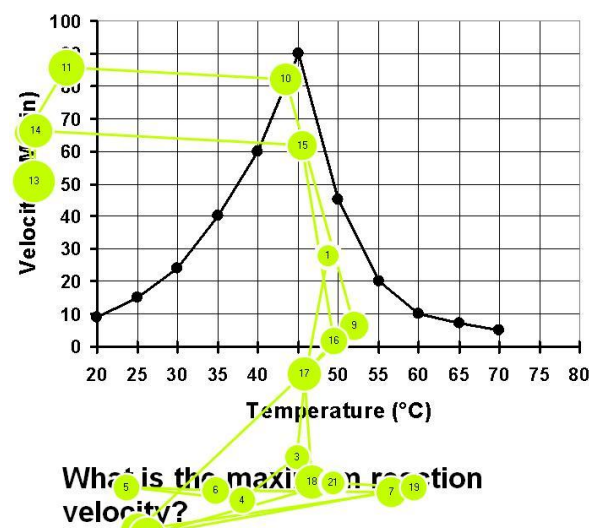

Figure 34. Gaze plots obtained for student #1 (mauve) and student #34 (green) while engaged in answering question 1.1 of Task 1.

Individual scan paths of other students in the two groups revealed similar patterns as those obtained with student #1 and #34, respectively. This implies that it is highly likely that the Bottom 10 group of students are, in fact, exhibiting working memory limitations while attempting to solve simple as well as complex graph-based reasoning problems. Working memory capacity can be tested by a variety of tasks (for example, the dual-task 'reading span' of Daneman & Carpenter, 1980) which would provide valuable data supporting eye movement data as a measure of graph comprehension. In fact, measures of working memory capacity are believed to be strongly related to performance in other complex cognitive tasks such as reading comprehension,

problem solving, and with any measures of the intelligence quotient (Conway, Kane & Engle, 2003). Evidence that working memory is linked to key learning outcomes in literacy and numeracy is extensive (Cowan & Alloway, 2008). A longitudinal study has confirmed that a child's working memory at 5 years old is a better predictor of academic success than IQ (Alloway & Alloway, 2010), implying that working memory impairments are associated with low learning outcomes and can be regarded as a high risk factor for educational underachievement for children. However, strategies for improving working memory capacity are available to educators (Alloway, 2010).

4.5 Test of the hypotheses

The **first hypothesis** required that the students score a mean of less than 50% on a subject-specific graphical literacy test incorporating examples of common graph decoding tasks as well a selection of biochemistry concepts in Cartesian graphs. An additional requirement was that students perform poorly, if at all, at Kimura's Level F (an abstract level of graph interpretation requiring the creation of dimensionally new information 'beyond the data'). The data obtained (a test mean score of 48,3%, only one question being answered correctly by all 82 students and only two students operating at Level F) supported the hypothesis. The null hypothesis is therefore rejected in favour of the alternative hypothesis.

The **second hypothesis** requires a moderate to strong positive correlation between the graphical literacy test mean scores obtained by the students and the arithmetic mean of the students' grades in biochemistry modules. The data obtained for the whole class (r=0,433) supported the hypothesis but does not imply causality. The strongest correlation was obtained for white females (r=0,671). The null hypothesis is therefore rejected in favour of the alternative hypothesis.

The **third hypothesis** requires gender and ethnicity differences (in favour of white males) in the mean scores of the graphical literacy test as well as graphical literacy levels of senior undergraduate biochemistry students at the UFS. White males did indeed score higher on the graphical literacy test (61,3 ±8,6%) than the other gender/ethnic groups. Overall, white students (59,7±10,7%) scored significantly (p=0,000021) higher than black students (45,3±10,7%), and male students (52,9±11,6%) scored significantly (p=0,0013) higher than female students (44,9±11,5%). In addition, the white males outperformed the other gender/ethnic

groups in five of the seven levels of Kimura's hierarchy. Black females performed the worst in six of the seven levels of Kimura's hierachy. The data obtained supported the hypothesis and the null hypothesis is therefore rejected in favour of the alternative hypothesis. However, it should be borne in mind that any conclusions reached regarding this hypothesis can at best be regarded as tentative and preliminary, given the sample composition and size.

The **fourth hypothesis** requires eye gaze data obtained using subject-specific graphs to be used to discern differences in the way senior undergraduate biochemistry students with high graphical literacy and students with low graphical literacy interpret graphs. Eye gaze data revealed significant differences ($p<0,05$) in task response times (Top 10 group of 293,7±44,6 sec vs Bottom 10 group of 458,6±138,3 sec) and mean gaze times in predefined areas of interest. There were also obvious visual differences in the location and distribution of hotspots on the heatmaps obtained for each of the twelve tasks. Gaze plots and scan paths of students in the two groups differed considerably and raised the possibility of working memory limitations as a source of low graphical literacy. The data obtained supported the hypothesis and the null hypothesis is therefore rejected in favour of the alternative hypothesis.

CHAPTER 5.   CONCLUSIONS AND RECOMMENDATIONS

5.1 Summary of the study

This study has attempted to identify the levels and type of problem associated with graph comprehension exhibited by senior biochemistry students at the UFS. Two experiments were used in this study to quantify the levels of graphical literacy amongst biochemistry undergraduate students and to identify reasons for low graphical literacy.

The main results are:

1.   The graphical literacy test devised by the researcher can be regarded as being reliable.

2.   Overall, the biochemistry students displayed notable deficits in graphical literacy capabilities.

3.   The biochemistry students struggled to operate at even the most basic level of graphical literacy, correctly answering only elementary-level questions for the most part, while only two students could be regarded as operating at an advanced abstract level.

4.   A low score on the graphical literacy test is a relatively good indicator of poor academic performance in biochemistry, and vice versa.

5.   Gender and ethnicity appears to play a role in graphical literacy with white males being tentatively identified as the most graphical literate and black females being the most graphical illiterate groups. This conclusion can at best be regarded as tentative because of the sample composition and size as well as confounding factors such as the test participants' facility with the language of instruction and test questions.

6.   Problems with graph comprehension identified were: finding ratios between data; identifying dependent and independent variables; interpreting slope and height changes on a curved graph; identifying and interpreting scale; using a second y-axis and multiple sources of information; working with reciprocal values of data; and extracting information from a graph without first critically examining the y-axis ('zoom graph' confusion). Students also did not incorporate their own prior knowledge or understanding into the construction of reasonable responses, instead they stated relationships in the data or incorporated creativity (fiction).

7. A major difficulty experienced by the students was not knowing whether to extract desired information from the slope or the height of a graph (the classic 'slope/height confusion').

8. Eye tracking data revealed subtle and not-so-subtle differences in the way students with low graphical literacy and students with high graphical literacy process information on graphs and is a valuable tool for assessing graphical literacy as well as cognitive aspects of graph comprehension, in particular, working memory.

The confirmation of previously held 'suspicions' regarding students' lack of critical thinking at an abstract level were confirmed by this study. The visual representation of data (or a concept) has to be integrated at some point with the abstract higher-level knowledge (subject or other) of the meaning of the graphical elements in order to achieve comprehension. If prior learned knowledge and the visual properties of the graphic representation are incompatible (or prior knowledge is absent or deficient) then integration will fail. This failure of integration and subsequent failure of comprehension is clearly evident from the results obtained in this study. Since a well-developed graphical literacy is essential for success at a post-graduate level in the natural sciences, the finding that a very small number of senior undergraduate students are operating at a 'beyond the data' abstract level is very disturbing and has implications for the future of biochemistry and the natural sciences in general in South Africa.

The eye tracking study reported here shows that it is possible to use eye gaze data to identify students experiencing problems with graph comprehension and that these problems could be linked to working memory limitations, lack of abstract thinking skills or deficiencies in language comprehension. The eye movement data is also consistent with the more conventional questionnaire, or survey, performance measure and might be of more value in graphical literacy data generation.

The integration of prior knowledge and the visual properties of the graph could not be tested but an indication that it could play a role is confirmed by correlation studies that showed a moderate positive correlation between undergraduate academic performance of the students and their score on the graphical literacy tests used in this study.

5.2 Recommendations

Simply identifying the fact that students cannot use graphs as "fluently" as they should is not enough. The results of this study have highlighted problem areas that have implications for the teaching of biochemistry at undergraduate level as follows:

(i) **Awareness.** The first step is for teachers of biochemistry to become aware of the problem. Knowing that the students cannot use graphs at the same level as the teacher means that in-class discussions of graphs found in biochemistry instructional material must include thorough explanations of all the information each one relates. Increased emphasis, therefore, needs to be placed on identifying information embedded in the framework of the graph and its referents such as axis labels, the scale used for measurement and what type of information is being presented.

(ii) **Scale.** Students need to develop a better understanding of scale and of making interpolations and extrapolations, particularly where inverse values are used. Students should be aware of the advantages and disadvantages of "Zoom graphs".

(iii) **Slope and height.** The very prevalent slope/height confusion should be addressed at the commencement of the introductory biochemistry course and parallels drawn between the students' physics background in kinematics and velocity graphs in enzyme kinetics. Class exercises and assessments should compel students to recognise their own incidence of slope/height confusion and correct it.

(iv) **Abstract concepts.** Teachers of biochemistry need to focus more on as well include more assessments related to higher cognitive levels in interpreting graphic data such as experimental results that support abstract and advanced concepts, determining trends or interpreting data from two or more variables concurrently. This last example, where two or even three y-axes are used, appears more often in research articles than in textbooks and such examples should be presented to students as an example of the 'real scientific world'.

(v) **Translating between representations.** Design activities that allow students to translate from one representation to another in an effort to link visual feature information to relevant quantitative information.

(vi) **Metacognitive aspects.** Train students to think of graph reading as an evaluation and interpretation task rather than a mere fact or data retrieval task. Students should be warned against being misled by prior expectations.

(vii) **Tutorials.** A tutorial on basic graphing skills should be presented either during the first-year cell biology course module or during the first practical session of the second-year introductory biochemistry course module. An example of such a tutorial (available online) is "*Basic Graphing Skills*" by Jennifer M. Wenner (2009).

Of concern is the distinct possibility that comprehension problems are a direct result of poor English language skills (the students didn't understand the question) or working memory limitations (the students forget what they are looking for on the graph). Although beyond the scope of this study, teachers of biochemistry should take cognisance of this and take more time explaining, in basic English, graph descriptions and legends as well as encouraging students to develop their working memory.

5.3 Contributions to knowledge

The project is innovative in the following respects:

◆ The application of graphical literacy theory and methodology for the first time in the field of biochemistry teaching and learning.

◆ The application of a technologically advanced remote eye tracking device in collecting quantitative data in a field typically dominated by the gathering of qualitative data.

5.4 Limitations of the study

A few limitations of the study design and sample entail caution when interpreting the data. Briefly, these limitations are:

(i) The sample was drawn from one cohort (3rd year students) at one educational institution following the same curriculum in biochemistry;

(ii) The gender and ethnicity profile of the sample was skewed overwhelmingly in favour of black females who comprised almost half of the sample for both Phase 1 and Phase 2 of the study. The fact that the

black females exhibited the lowest graphical literacy 'profiles' of all of the gender/ethnic groups should be taken into account when interpreting means and overall graphical literacy levels for the sample.

(iii) Only five students in the sample were native English speakers (identified on their UFS admission as English home language) meaning that the remaining 77 students were reading the test questions and articulating their answers in their second or third language. This factor might influence the comprehension of tasks. Unfortunately no English literacy test results are available for the cohort of students from which this sample was drawn (the UFS only started literacy testing at the beginning of 2010) but it might be advantageous for this type of study to administer a language literacy (or proficiency) test together with a graphical literacy test in order to correlate these particular literacies. The eye tracking phase of the study allowed the students the opportunity to request an elucidation of a task question and a relatively large number took advantage of this offer.

(iv) The majority of questions in the various tasks comprising Questionnaire #1 were level A tasks (A1 – A4) focussing mainly on extracting information from the graphs and looking for relationships among the data. The number (14) of higher level (D, E and F) questions such as prediction, extrapolation and making inferences was almost half of the lower level questions (29). This oversight was corrected in the case of Questionnaire #2 where the number of level A questions almost equalled the number of higher level questions. The reader of this dissertation should take into account that higher level tasks (such as those at Level F) are the most difficult to formulate and assess.

(v) Because of the large number of lower level questions included in Questionnaire #1, the test eventually comprised a rather unwieldy 43 questions. The time taken to complete the test exceeded 40 minutes and a very real possibility exists that the students were tired, both physically and mentally, by the time they were confronted with the cognitively demanding level E and Level F tasks. This could result in an underperformance at these levels being observed. However, this was not a factor in Questionnaire #2 (only 12 questions taking a total of about seven minutes) and the proportion of students achieving at Level F was similar to the figure obtained with Questionnaire #2.

(vi) Level B and C tasks were not included in the study since they are concerned more with arranging data and drawing graphs, i.e. encoding, not decoding. However, one could argue that assessing 'graph sense' should be part of the graphical literacy 'profile' of a sample group or individual and then include a few 'drawing graphs' exercises in a graphical literacy test.

5.5 Suggestions for further studies

Clearly, this study has identified certain aspects of graphical literacy in biochemistry that need to be studied further. These are:

(i) **The question of validity**. The test (Questionnaire #1) should be administered to senior biochemistry students at a minimum of two other tertiary institutions in order to establish validity. One obstacle is that the current test is too long and, taking into account point (v) in section 5.4 above, should be shortened. This entails administering the new test to a new cohort of biochemistry students at the UFS before establishing validity.

(ii) **The challenge of Kimura's Level F.** The results of this study points to a definite need to gather more data on the challenge of Kimura's Level F (or Curcio's 'Reading Beyond The Data') for senior biochemistry students. Questions that arise concern the conceptual structure of Level F: Are the performances at Level F strictly statistical or do they relate to more general critical thinking skills? What about the influence of biochemistry subject knowledge? At what point do Level E tasks become Level F tasks, where are the boundaries? Is it possible to set an upper boundary for Level F?

(iii) **Language:** Because of a possible language comprehension problem, it might prove valuable to study the value of including translations of the questions in the students' home language.

(iv) **Cross-disciplinary:** Since graphical literacy is a requirement for all biological sciences, it might prove valuable to modify the graph examples and administer the test to microbiology students, for example.

(v) **Participant age:** A three- or even four-year longitudinal study designed to examine the decoding performance of students' solving of graph-

related problems over time would enhance understanding of the development of students' ability to decode graphs. Another possibility would be to administer a more generic graphical literacy test to first-year cell biology students as well as second-, third- and fourth-year undergraduate biochemistry students in order to compare their different levels of performance.

(vi)   <u>**Participant gender:**</u>   In order to gain insights into explanations for the dramatic performance differences in graphical literacy exhibited by males and females in this study, the identification and addition of more items that reveal performance differences would strengthen the present study.

(vii)   <u>**Working memory:**</u>   Testing of working memory capacity in conjunction with graphical literacy testing would result in a clearer understanding of the sources of low graphical literacy and would also direct any interventions with respect to at-risk students.

REFERENCES

Alloway, T.P. (2010). *Improving working memory: supporting students' learning.* London: SAGE Publications.

Anderson, T.R. & Grayson, D.J. (1994). Improving students' understanding of carbohydrate metabolism in first-year biochemistry at tertiary level. *Research in Science Education*, 24, 1-10.

Aoyama, K. & Stephens, M. (2003). Graph interpretation aspects of statistical literacy: a Japanese perspective. *Mathematics Education Research Journal,* 15(3), 207-225.

Avgerinou, M. & Ericson, J. (1997). A review of the concept of visual literacy. *British Journal of Educational Technology,* 28(4), 280-291.

Beichner, R.J. (1994). Testing student interpretation of kinematics graphs. *American Journal of Physics,* 62, 750-768.

Biggs, J.B. & Collis, K.F. (1982). *Evaluating the quality of learning: The SOLO taxonomy.* New York: Academic Press.

Biggs, J.B. & Collis, K.F. (1991). *Multimodal learning and the quality of intelligent behavior.* In: Intelligence: reconceptualization and measurement (Rowe, H., Ed.), 57-76. Hillsdale, NJ: Erlbaum.

Biggs, N, Lloyd, E.K. & Wilson, R.J. (1976). Graph theory, 1736-1936, 66. New York, Oxford University Press.

Bracey, G.W. (2004). Putting things graphically. *Phi Delta Kappa.* 86(3), 252-253.
　　(Available at: http://www.goliath.ecnext.com/coms2/gi_0199-3409774/RESEARCH-Putting-Things-Graphically.html)

Braden, R.A. & Hortin, J.A. (1982). Identifying the theoretical foundations of visual literacy. *Journal of Visual/Verbal Languaging*, 2(2), 37-42.

Brasell, H.M. & Rowe, M.B. (1993). Graphing skills among high school physics students. *School Science and Mathematics*, 93(2), 63-70.

Carpenter, P.A. & Shah, P. (1998). A model of the perceptual and conceptual processes in graph comprehension. *Journal of Experimental Psychology: Applied*, 4, 75-100.

Carswell, C.M., Frankenberger, S. & Bernhard, D. (1991). Graphing in depth: Perspectives on the use of three-dimensional graphs to represent lower-dimensional data. *Behaviour and Information Technology*, 10, 459-474.

Chambers, J., Cleveland, W., Kleiner, B. & Tukey, P. (1983). *Graphical methods for data analysis.* Belmont CA: Wadsworth International Group.

Clark, R. C. & Mayer, R. E. (2002). *E-learning and the science of instruction: proven guidelines for consumers and designers of multimedia learning.* San Francisco: Jossey-Bass Pfeiffer.

Conway, A.R., Kane, M.J. & Engle, R.W. (2003). Working memory capacity and its relation to general intelligence. *Trends in Cognitive Sciences,* 7(12), 547-552.

Cortina, J.M., (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology*, 78, 98-104.

Cowan, N. & Alloway, T.P. (2008). *The development of working memory.* In: Development of memory in childhood (Cowan, N., Ed.), 303-342. Hove, England: Psychology Press.

Cowen, L. (2001). *An eye movement analysis of web-page usability.* Unpublished masters thesis, Lancaster University.

(Available at: http://www.psych.lancs.ac.uk/people/uploads/LindenBall20031001T094007.pdf)

Cox, P. (2010). *Glossary of mathematical mistakes.* Accessed January 2010 at http://www.mathmistakes.com.

Cronbach, L.J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika,* 16(3), 297-334.

Curcio, F.R. (1981). *The effect of prior knowledge, reading and mathematics achievement, and sex on comprehending mathematical relationships expressed in graphs* (Final Report). Brooklyn, NY: St. Francis College. (ERIC Document Reproduction Service No. ED 210 185).

Curcio, F.R. (1987). Comprehension of mathematical relationships expressed in graphs. *Journal for Research in Mathematics Education, 18*(5), 382-393.

Daneman, M. & Carpenter, P.A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning & Verbal Behaviour,* 19(4), 450-466.

Dondis, D.A. (1973). *A Primer of Visual Literacy.* Cambridge, Massachusetts: MIT Press.

Ewing, K. (2005). *Studying web pages using eye tracking.* Whitepaper.
        (Available at: http://www. alt64.com/tobiiftp/Studying_web_pages_using_eyetracking.pdf)

Few, S. (2004). Elegance through simplicity. *Intelligent Enterprise,* 7(15), 35-39.

Fischer, M.H. (2000). Do irrelevant depth cues affect the comprehension of bar graphs? *Applied Cognitive Psychology,* 14(2), 151-162.

Forgasz, H.J., Leder, G.C. & Kauer, B. (2001). Who can(t) do maths – boys/girls? Beliefs of Australian and Singaporean secondary school students. *Asia Pacific Journal of Education,* 21(2), 106-116.

Forgasz, H.J., Leder, G.C. & Kloosterman, P. (2004). New perspectives on the gender stereotyping of mathematics. *Mathematics Thinking and Learning*, 6, 389-420.

Friel, S., Curcio, F.R. & Bright, G.W. (2001). Making sense of graphs: Critical factors influencing comprehension and instructional implications. *Journal for Research in Mathematics Education*, 32(2), 124-159.

Gattis, M. & Holyoak, K.J. (1996). Mapping conceptual to spatial relations in visual reasoning. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 22, 231-239.

Gardner, H. (1999). *Intelligence reframed: Multiple intelligence for the 21st century.* New York: Basic Books.

Garnett, P.J., Garnett, P.J. & Hackling, M.W. (1995). Students' alternative conceptions in chemistry: a review of research implications for teaching and learning. *Studies in Science Education*, 25, 69-95.

George, D., & Mallery, P. (2003). *SPSS for Windows step by step: A simple guide and reference. 11.0 update* (4th ed.). Boston: Allyn & Bacon.

Goldberg, J.H. & Kotval, X.P. (1998). *Eye movement-based evaluation of the computer interface.* In: Advances in Occupational Ergonomics and Safety (Kumar, S.K. ed.), 529-532. Amsterdam, IOS Press.

Grayson, D.J., Anderson, T.R. & Crossley, L.G. (2001). A four-level framework for identifying and classifying student conceptual and reasoning difficulties. *International Journal of Science Education*, 23, 611-622.

Gulliver, S.R. & Ghinea, G. (2004). *Region of interest displays: Addressing a perceptual problem?* Multimedia software engineering 2004. Proceedings of the IEEE sixth international symposium. pp. 2-9. (Available at: http://www.ieeexplore.ieee.org/iel5/9470/30047/01376634.pdf)

Guthrie, J.T., Weber, S., & Kimmerly, N. (1993). Searching documents: Cognitive processes and deficits in understanding graphs, tables and illustrations. *Contemporary Educational Psychology. Searching for information,* 18(2), 186-221.

Hale, P. (2000). Kinematics and Graphs: Students' difficulties and CBLS. *Mathematics Teacher,* 5 (93), 414-418.

Hadjidemetriou, C. & Williams, J. (2000). *Assessing graphical literacy in year 10 mathematics pupils.* British Educational Research Association Student Symposium.
    (Available at: http://www.lta.education.manchester.ac.uk/ch/ch1.pdf)

Hansen, D.W., Hansen, J.P., Nielsen, M., Johansen, A.S. & Stegmann, M.B. (2002). *Eye typing using markov and active appearance models.* Applications of computer vision 2002. Proceedings of the sixth IEEE workshop, 132-136.
    (Available: http://www.ieeexplore.ieee.org/iel5/8422/26538/01182170.pdf?arnumber=1182170)

Hanson, R.W. (1989). The role of ATP in metabolism. *Biochemistry Education*, 17 (2), 86-92.

Harrison, A.G., Grayson, D.J. & Treagust, D.F. (1999). Investigating a grade 11 student's evolving conceptions of heat and temperature. *Journal of Research in Science Teaching*, 36 (1), 55-87.

Hill, D. (1988). Misleading illustrations. *Research in Science Education*, 18, 290-297.

Hink, J.K., Eustace, J.K. & Wogalter, M.S. (1998). Do grables enable the extraction of quantitative information better than pure graphs and tables? *International Journal of Industrial Ergonomics*, 22, 439-447.

Holsanova, J. (2004). *Tracking multimodal interaction with new media.* A position paper at the Workshop On The Citizens Use And Comprehension Of Information On The Internet.
    (Available at: http://www.lucs.lu.se/People/Jana.Holsanova/PDF/Holsanova.2004.pdf)

Jacob, R.J.K. & Karn, K.S. (2003). *Commentary on section 4: Eye tracking in human-computer interaction and usability research: Ready to deliver promises.*
    (Available at: http://www.alexpoole.info/academic/Poole&Ball%20EyeTracking.pdf)

Kimura, S. (1999). *Toukeizyouhoukyouikuno Karikyuramuto 5-dankaino Toukeitekitankyu Purosesu [Curriculum of statistics education and five phases of statistical inquiry process]* (in Japanese). In: Toukeizyouhoukyouikuno Rironto Zyugyouzissenno Tenkai (Zentouken, Ed.), 33-46. Tsukuba, Syuppankai.

Knuth, E.J. (2000). Understanding connections between equations and graphs. *Mathematics Teacher,* 1 (93), p 48-54.

Kumar, N. & Benbasat, I. (2004). The effect of relationship encoding, task type, and complexity of information representation: An empirical evaluation of 2D and 3D line graphs. *MIS Quarterly,* 28(2).

Lake, D. (1999). Helping students to go Solo: teaching critical numeracy in the biological sciences. *Journal of Biological Education,* 33(4), 191-199.

Lazarowitz, R. & Penso, S. (1992). High school students' difficulties in learning biology concepts. *Journal of Biological Education,* 26(3), 215-223.

Leedy, P.D. (1989). *Practical Research: Planning and Design.* New York: Macmillan.

Linn, M. & Petersen, A. (1985). Emergence and characterization of sex differences in spatial ability: a meta-analysis. *Child development*, 56, 1479-1498.

Lohse, G., Walker, N., Biolsi, K. & Rueter, K.H. (1991). Classifying graphical information. *Behaviour and Information Technology*, 10, 419-436.

Lokan, J., Greenwood, L. & Cresswell, J. (2001). *15-up and counting, reading, writing, reasoning… how literate are Australia's students? The PISA 2000 survey of students reading, mathematical and scientific literacy skills.* Camberwell, Australia: Australian Council for Educational Research.

Lowrie, T. & Diezmann, C.M. (2011). Solving graphics tasks: Gender differences in middle-school students. *Learning and Instruction*, 21(1), 109-125.

Mayer, R.E. (1993). *Illustrations that instruct*. In: Advances in Instructional Psychology (Glaser, R. ed.), 253-384. Vol. 4. Hillsdale: Lawrence Erlbaum Associates.

Mayer, R. E., & Gallini, J. K. (1990). When is an Illustration Worth Ten Thousand Words? *Journal of Educational Psychology*, 82(4), 715-726.

McDermott, L.C., Rosenquist, M.L., & van Zee, E.H. (1987). Student difficulties in connecting graphs and physics: Examples from kinematics. *American Journal of Physics*, 55(6), 503-513.

McDonald, R.P. (1999). *Test theory: A unified treatment*. Psychology Press. ISBN 0805830758.

McKenzie, D.L. & Padilla, M.J. (1986). The construction and validation of the Test of Graphing in Science (TOGS). *Journal of Research in Science Teaching*, 23, 571-579.

Messaris, P. (1994). *Visual Literacy: Image, Mind, and Reality*. Boulder: Westview Press.

Milekic, S. (2003). *The more you look the more you get: Intension-based interface using gaze tracking*. Museums & Web 2003. Selected papers. Proceedings from an international conference. (Available at: http://www.uarts.edu/faculty/smilekic/paper/milekicMW03.pdf)

Moore, D.M., & Dwyer, F.M., (eds.) (1994). *Visual Literacy. A Spectrum of Visual Learning.* Englewood Cliffs, NJ: Educational Technology Publications.

Morimoto, C.H., Amir, A. & Flickner, M. (2002). *Detecting eye position and gaze from a single camera and 2 light sources*. 16th International Conference on Pattern Recognition (ICPR'02) proceedings, 4, 40314. (Available at: http//www.ieeecomputersociety.org)

Morimoto, C.H. & Mimica, M.R.M. (2004). Eye gaze tracking techniques for interactive applications. *Computer Vision and Image Understanding,* 98(1), 4-24.

Paivio, A. (1991). Dual coding theory: retrospect and current status. *Canadian Journal of Psychology,* 45, 255-87

Peebles, D. & Cheng, P.C-H. (2002). Extending task analytic models of graph-based reasoning: A cognitive model of problem solving with Cartesian graphs in ACT-R/PM. *Cognitive Systems Research*, 3, 77-86.

Peeck, J. (1993). Increasing picture effects in learning from illustrated text. *Learning and Instruction*, 3, 227-238.

Pettersson, R. (1989). *Visuals for information: research and practice.* Educational Technology Publications, Englewood Cliffs, New Jersey, USA.

Pfundt, H. & Duit, R. (1998). Bibliography: Students' alternative frameworks and science education. (Kiel, Germany: IPN).

Postigo, Y. & Pozo, J.I. (2004). On the road to graphicacy: The learning of graphical representation systems. *Educational Psychology,* 24(5), 623-644.

Qvarfordt, P. & Zhai, S. (2005). *Conversing with the user based on eye-gaze patterns* . Paper presented at the Conference on Human Factors in Computing Systems. Proceedings of the SIGCHI conference on human factors in computer systems. pp. 221-230.
(Available at: http:// portal.acm.org/citation.cfm?id=1054972.1055004)

Readence, J., Bean, T., & Baldwin, S. (2004). *Content area literacy: an integrated approach,* 8th Ed., 260. Dubuque, IA: Kendall-Hunt.

Renshaw, J.A., Finlay, J.E., Tyfa, D. & Ward, R.D. (2003). *Designing for visual influence: An eye tracking study of the usability of graphical management information.* Human computer interaction. INTERACT'03. (Available at:
http://www.idemployee.id.tue.nl/g.w.m.rauterberg/conferences/INTERACT2003/INTERACT20 03-p144.PDF)

Renshaw, J.A., Finlay, J.E., Tyfa, D. & Ward, R.D. (2004). Understanding visual influence in graph design through temporal and spatial eye movement characteristics. *Interacting with Computers,* 16(3), 557-578.

Rieber, L.P. (1995). A historical review of visualisation in human cognition. *Educational Technology, Research and Development,* 43(1), 1042-1629.

Roth, W.-M. (2002). Reading graphs: contributions to an integrative concept of literacy. *Journal of Curriculum Studies,* 34(1), 1-24.

Roy, D., Ghitza, Y., Bartelma, J. & Kehoe, C. (2004). *Visual memory augmentation: Using eye gaze as an attention filter.* Proceedings of the Eighth International Symposium on Wearable Computers (ISWC'04).
(Available at: http://www.media.mit.edu/cogmac/publications/roy ghitza bartelma.pdf)

Schiessl, M., Duda, S., Thölke, A. & Fischer, R. (2003). *Eye tracking and its application in usability and media research.* MMI-Interactive'03.
(Available at: http://scholar.google.com/scholar?hl=en&lr=&q=cache:8sfPt4Tey3cj:www.eye-square.de.dokumente/EyeTracking-ResearchApplications.pdf)

Schönborn K.J. & Anderson T.R. (2006). The importance of visual literacy in the education of biochemists. *Biochemistry and Molecular Biology Education,* 34(2), 94-102.

Schönborn K.J., Anderson T.R. & Grayson D.J. (2002). Student difficulties with the interpretation of a textbook diagram of immunoglobulin G (IgG). *Biochemistry and Molecular Biology Education,* 30, 93-97.

Sears, D.W. & Thompson, S.E. (2007). *Pre-instructional assessment of the basic chemical and molecular literacy of biochemistry students.* Bioliteracy net.
(Available at: http://bioliteracy.net/Readings/paperssubmittedPDF/Sears.pdf)

Sears, D.W. (2008). *Moving toward a biochemistry concept inventory.* ASBMB Today September 2008, 19-21.

Shah, P., Freedman, E. G. & Vekiri, I. (2005). *The Comprehension of Quantitative Information in Graphical Displays.* In: The Cambridge Handbook of Visuospatial Thinking (Shah, P. & Miyake, A. eds), 426-476. New York, NY.: Cambridge University Press.

Shah, P. & Hoeffner, J. (2002). Review of graph comprehension research: Implications for instruction. *Educational Psychology Revue,* 14(1), 47-69.

Shah, P., Mayer, R.E. & Hegarty, M. (1999). Graphs as aids to knowledge construction. *Journal of Educational Psychology,* 91(4), 690-702.

Shaughnessy, J.M., Garfield, J. & Greer, B. (1996). *Data handling.* In: International handbook of mathematics education (Bishop, A.J., Clements, K., Keitel, C., Kilpatrick, J. & Laborde, C., Eds.),. (Part 1) pp. 205-237. Dordrecht, Netherlands: Kluwer Academic Publishers.

Siegrist, M. (1996). The use and misuse of three-dimensional graphs to represent lower-dimensional data. *Behaviour and Information Technology*, 15, 96-100.

Simon, S. (2008). *What's a good value for Cronbach's Alpha?* Accessed 2009 at:
http:/www.childrens-mercy.org/stats/weblog2004/CronbachAlpha.asp

Spelke, E.S. (2005). Sex difference in intrinsic aptitude for mathematics and science? A critical review. *American Psychologist*, 60, 950-958.

Spence, I. (1990). Visual psychophysics of simple graphical elements. *Journal of Experimental Psychology: Human Performance and Perception,* 16, 683-692.

Sylvester, J.J. (1878). Chemistry and Algebra. *Nature,* 17, 284. Accessed 2010 at:
http://www.archive.org/stream/nature15unkngoog#page/n312/mode/1up

Tairab, H.H. & Khalaf Al-Naqbi, A.K. (2004). How do secondary school science pupils interpret and construct scientific graphs? *Journal of Biological Education*, 38(3), 127-132.

Tobii (2004). *Product description. Tobii 50 series.*
(Available at: http:// www.srlabs.it/pdf/Tobii_50series.pdf)

Tobii (2006). *User manual.* (Available at: http:// www.tobii.se/linkpage.asp?sid=826)

Tracey, D.A. (1990). Toys, spatial ability, and science and mathematics: Are they related? *Sex Roles,* 17, 115-137.

Tufte, E.R. (1983). *The visual display of quantitative information.* Chesire CT: Graphics Press.

Tzanidou, E. (2003). *Improving usability of e-commerce sites by tracking eye movements.* Human Computer Interaction. INTERACT'03. Available:<URL:http//www.idemployee.id.tue.nl>

Van Tonder, A. & De Lange, R. (2002). *Biochemistry student's literacy and comprehension of scientific graphs.* In: International Visual Literacy Association Selected Readings (Griffin, R.E., Williams, V.S., eds.) 2002:205-210. IVLA, Loretto, PA, USA.

Vella, F. (1990). Difficulties in learning and teaching of biochemistry. *Biochemistry Education*, 18(1), 6-8.

Wenner, J.M. (2009). *Basic Graphing Skills.* Accessed April 2009 at:
http://serc.carleton.edu/quantskills/methods/quantlit/basicgraph.html

Winn, W. D. (1987). *Charts, graphs, and diagrams in educational materials.* (Chap. 5). In: The Psychology of Illustration, Vol I: Basic Research (Willows, D.M. & Houghton, H.A., Eds.). New York: Springer.

Wolf, W. (1970). *A study of eye movement in television viewing.* Final report. Columbus: Ohio State University, (ERIC Reproduction Service No. ED 046 254).

Wood, E.J. (1990). Biochemistry is a difficult subject for both student and teacher. *Biochemistry Education*, 18 (4), 170-172.

Zacks, J., Levy, E., Tversky, B. & Schiano, D.J. (1998). Reading bar graphs: Effects of extraneous depth cues and graphical context. *Journal of Experimental Psychology: Applied*, 4, 119-138.

BIOCHEMISTRY EDUCATION RESEARCH GROUP, UFS

Please answer the following questionnaire truthfully without guessing (this doesn't count towards your marks but will help us identify problem areas in the understanding of Biochemistry concepts and 'language'). Answer the questions or circle the letter corresponding to the answer of your choice and write a short motivation for your answer, if so required.

Student number: ._____.

TASK 1.

Answer the following questions using the graph below.
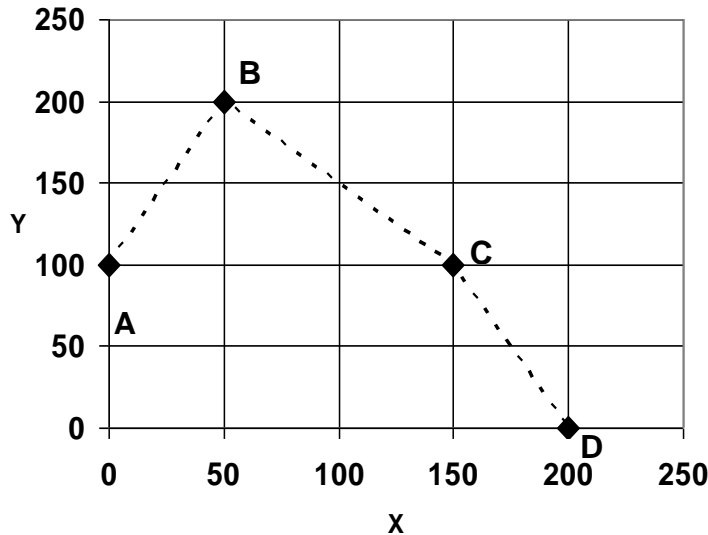
1.1 Which point(s) lie(s) on the y-axis?

Answer: **……...**

1.2 Which point would be labelled with the ordered pair notation of (50, 200)?

Answer: **……...**

1.3 Which point(s) has/have a y-coordinate of 100?

Answer: **……...**

TASK 2.

The following figure is a graph of reaction velocity versus pH for an enzyme-catalysed reaction illustrating the dependence of enzyme activity on pH. Answer the following questions:
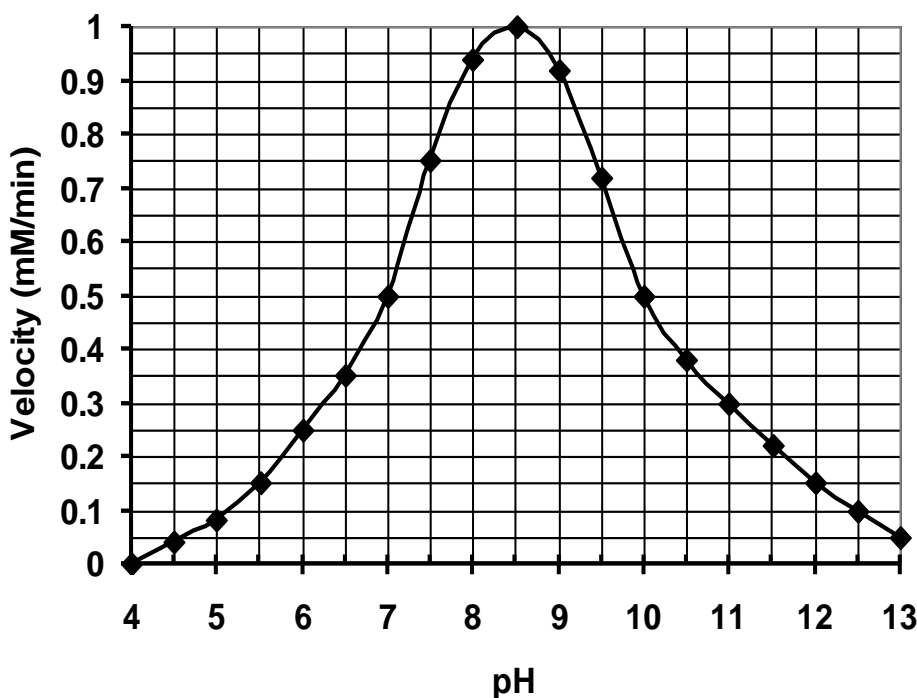
2.1 What is the maximum velocity observed?

Answer: **…………..**

2.2 At which pH does the maximum velocity occur?

Answer: **……….…..**

2.3 At which pH(s) is half the maximum velocity observed?

Answer: **…………...**

2.4 Which axis is the dependent variable?

Answer: **……………..**

2.5 Is the velocity at pH 6,0 a quarter or half of the velocity observed at pH 7,0?

Answer: **………..…...**

TASK 3.

The following figure is a graph of reaction velocity (units per millilitre or U/ml) versus substrate concentration ([S] in mM or mmol/l) for an enzyme-catalysed reaction. Answer the following questions:
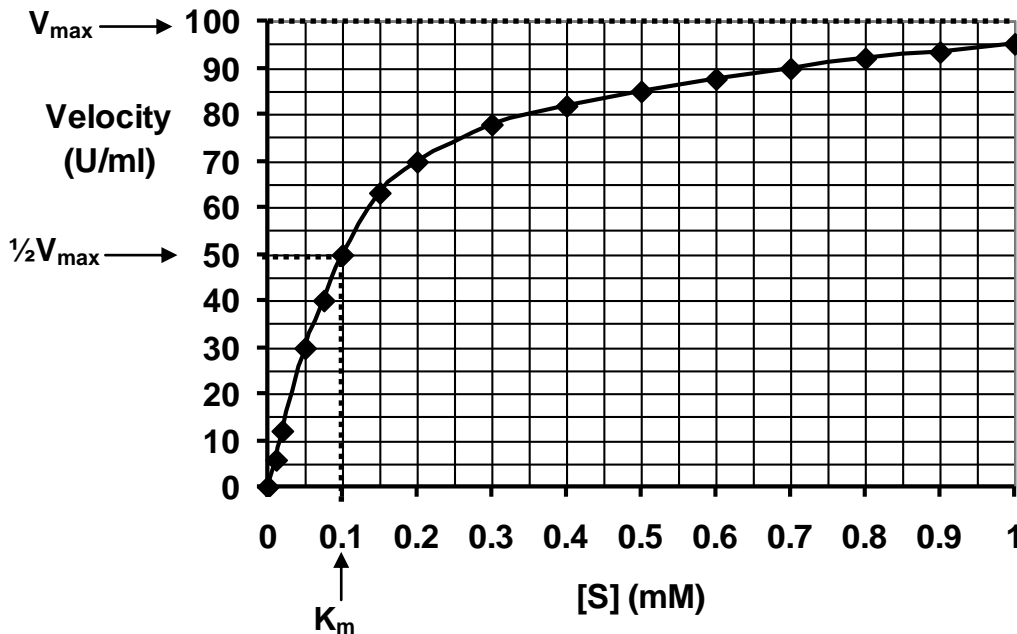
3.1 K_m will be expressed in terms of the following units (circle the correct answer):
- (a) units of velocity (e.g. units/ml)
- (b) v/[S]
- (c) same as the concentration units of S, the substrate (e.g. mmol/l or mol/l).
- (d) no units, it's relative.
- (e) I don't know.

3.2 If $[S] = 2K_m$ what will the velocity be?
Answer:…………

3.3 If $[S] = 3K_m$ what will the value of $\frac{1}{2}V_{max}$ be?
Answer:…………

TASK 4.

The following figure illustrates the effect of pH on the oxygen-binding affinity (Θ) of hemoglobin measured at different pressures of oxygen (pO_2), otherwise known as the Bohr Effect.
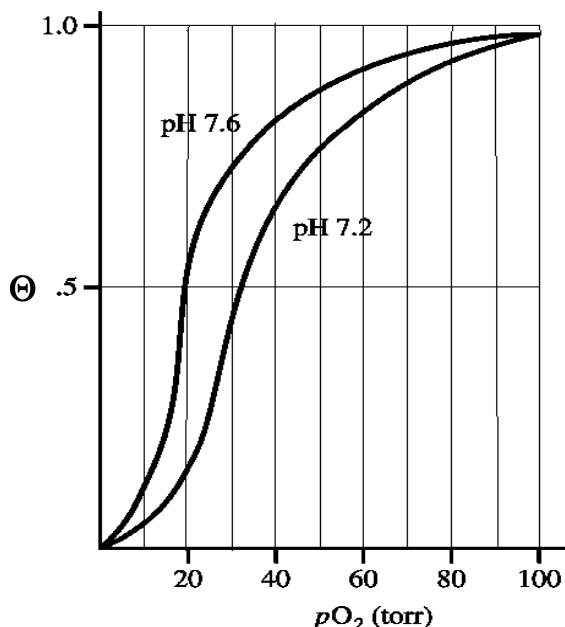
4.1 Name the independent variable:

………………………

4.2 Name the dependent variable:

………………………

4.3 It is possible to conclude that as the pH decreases the affinity of hemoglobin for oxygen:
(circle the correct answer)
- (a) increases
- (b) stays the same
- (c) decreases
- (d) I don't know

TASK 5.

The following figure is a graph (or progress curve) of absorbance readings (in units) versus time (in seconds) obtained for two different enzymes, A and B.
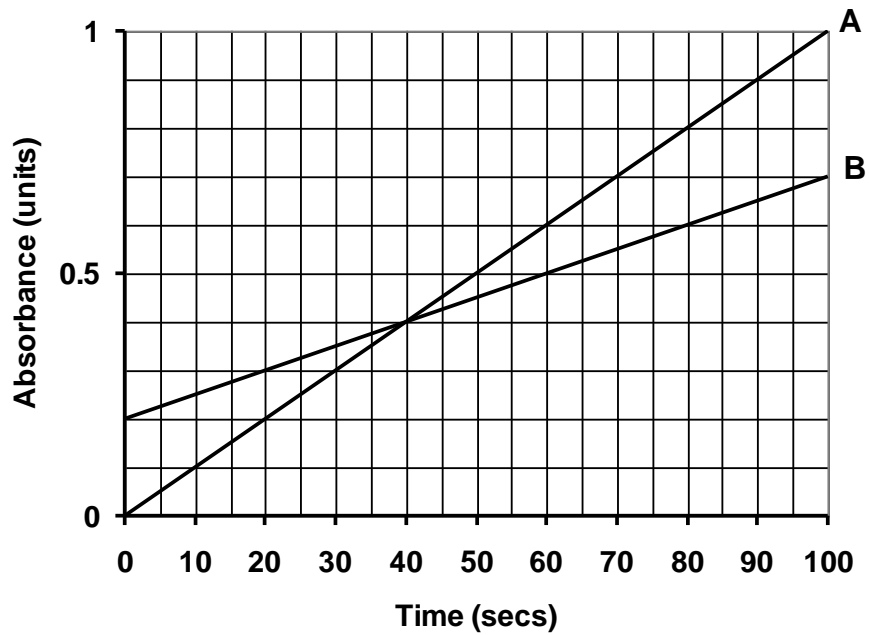
5.1 At the instant t = 20 sec, is the velocity (measured as change in absorbance per second or $\Delta Abs/sec$) of enzyme A greater than, less than, or equal to that of enzyme B?

Answer: ……………………..

Explain your answer: ………………………………………………………………………………

………………………………………………………………………………………………………

………………………………………………………………………………………………………

5.2 In the time frame 0 to 100 seconds, do enzymes A and B ever have the same velocity, and if so, after what time interval?

Answer: …………………………..

Explain your answer: ………………………………………………………………………………

………………………………………………………………………………………………………

………………………………………………………………………………………………………

TASK 6.

The following figure shows a graph of concentration (in mM) versus time (in seconds) obtained for an enzyme-catalysed reaction where substrate and product concentrations were measured.
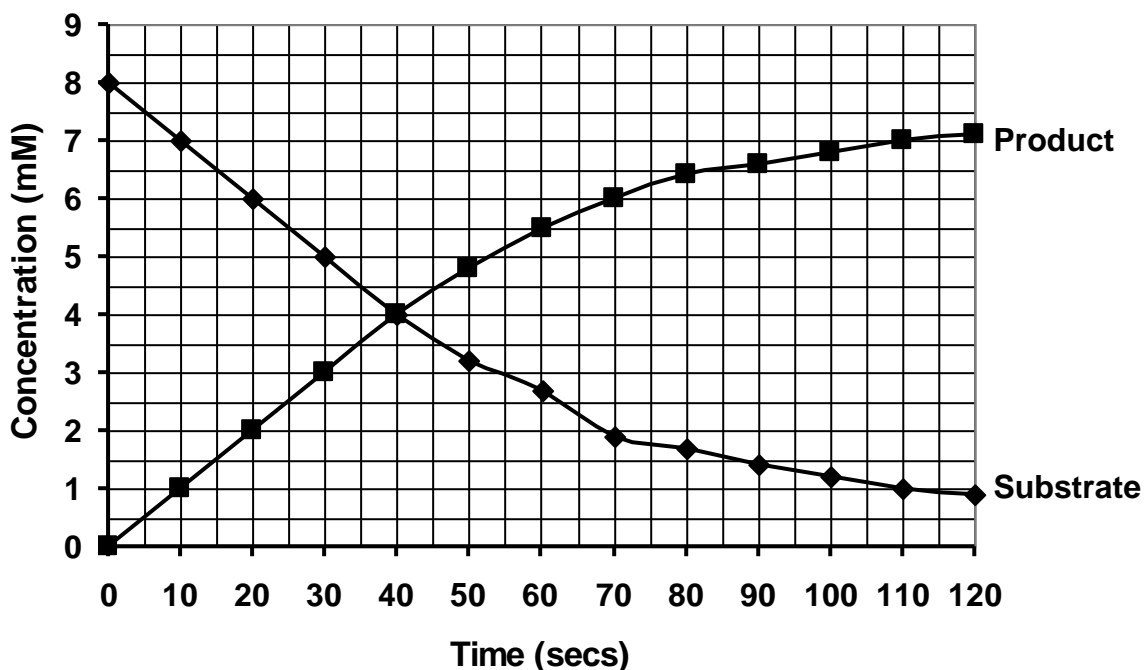
6.1 The maximum reaction velocity in mM/sec (appearance of product or consumption of substrate per second) occurs between: (circle a,b,c,d or e)

(a) 0 and 40 sec

(b) 40 and 80 sec

(c) 80 and 120 sec

(d) 0 and 120 sec

(e) I don't know

6.2 The graph shows: (choose one statement by circling a,b,c,d or e)

(a) A linear relationship between product concentration and time.

(b) An initial high release of substrate followed by a gradual tapering off with time

(c) As the substrate is consumed, the product of the reaction appears and increases in concentration up to a maximum with a gradually decreasing rate of production as substrate is consumed and eventually used up.

(d) An increasing rate of substrate consumption over time

(e) I don't know.

6.3 Will the substrate concentration ever reach zero?

Answer: ……………………..

Explain your answer: …………………………………………………………………………….
…………………………………………………………………………………………………….
…………………………………………………………………………………………………….

6.4 What will happen to the product concentration?

Answer: …………………………..………..

Explain your answer: …………………………………………………………………………
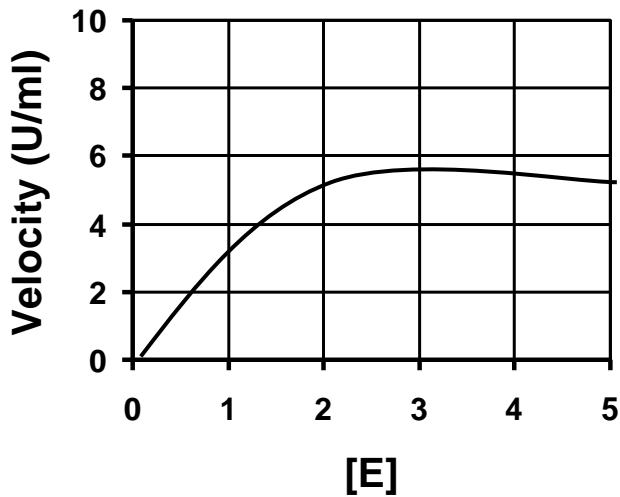…………………………………………………………………………………………………….

TASK 7.

Consider the following statement: *"The reaction velocity (in U/ml) of an enzyme-catalysed reaction increases by 100% (i.e. doubles) if the enzyme concentration, [E], increases by 100% (doubles)."* Which one of the following graphs would be an accurate representation of the preceding statement:? (Circle a, b, c or d)
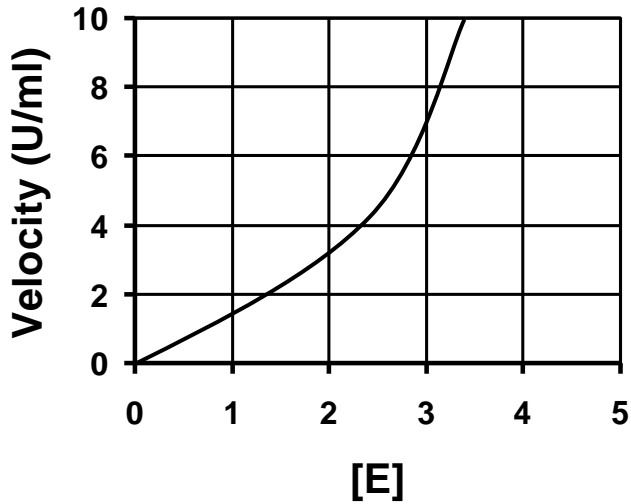
(a)   Graph 1
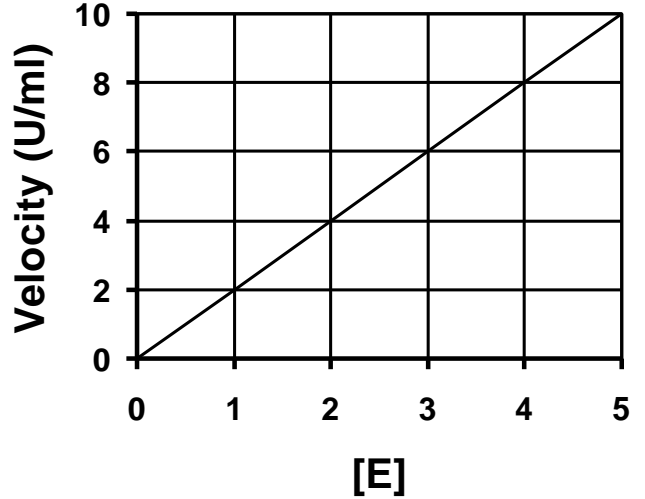
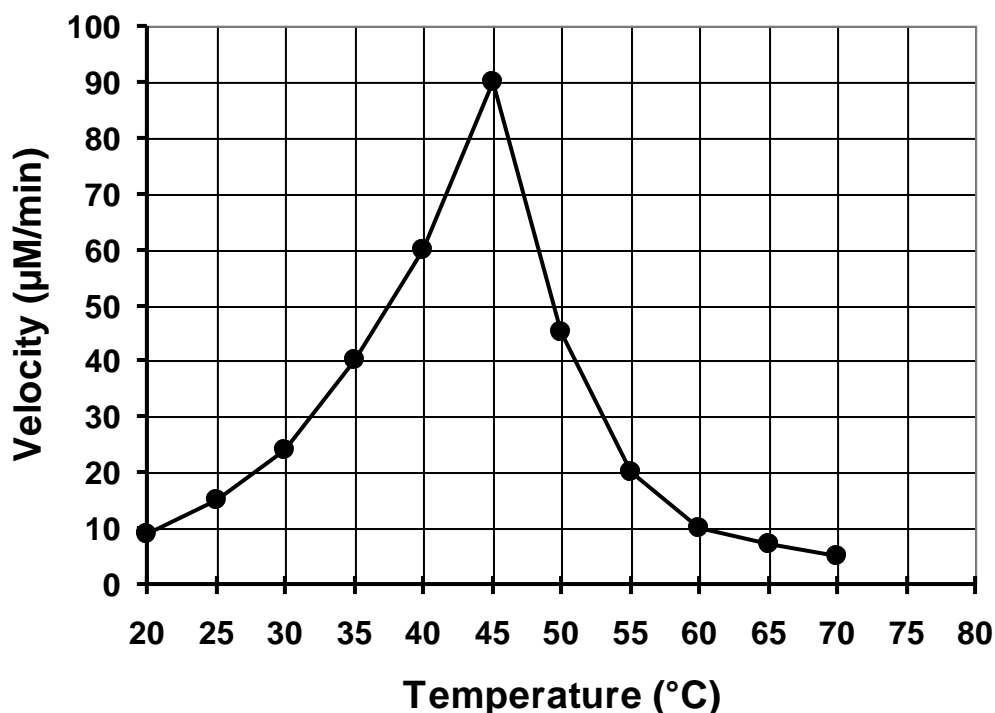(b)   Graph 2

(c)   Graph 3

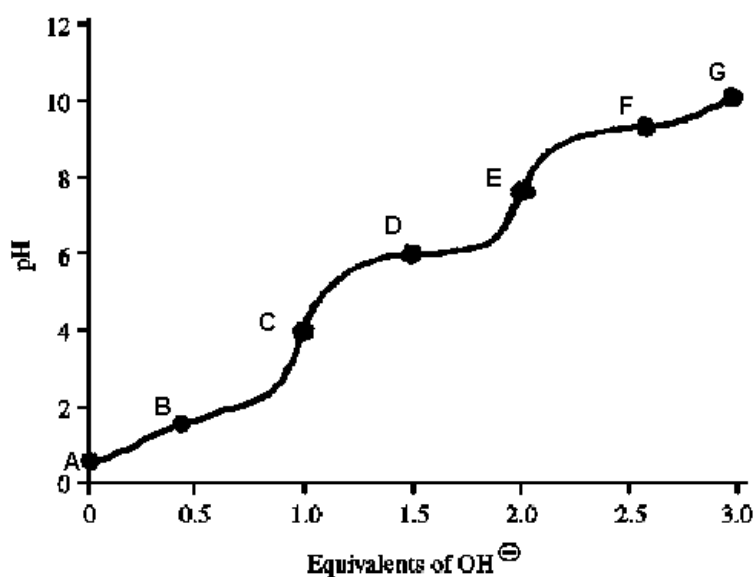**I DON'T KNOW**

(d)   Don't know

Give a reason for your answer:……………………………………………………………
……………………………………………………………………………………………………
……………………………………………………………………………………………………..
…………………………………………………………………………………………….…….…..

TASK 8.

Consider the following graph:

8.1 The maximum reaction velocity (the temperature optimum) is: (circle the correct answer)

(a) 45 µM/min at 50°C

(b) 90 µM/min at 40°C

(c) 10 µM/min at 20°C

(d) 90 µM/min at 45°C

(e) I don't know

8.2 We can say that : (choose one statement by circling a,b,c,d or e)

(a) The velocity of an enzyme-catalysed reaction increases continuously with an increase in temperature.

(b) The velocity of an enzyme-catalysed reaction decreases with an increase in temperature.

(c) The velocity of an enzyme-catalysed reaction increases initially to a maximum and then decreases as temperature is increased.

(d) I don't know.

8.3 At a temperature of 80°C the reaction velocity will most likely be: (Circle your choice of a,b,c,d or e):

(a) The same as at 70°C.

(b) Close to zero.

(c) -5 µM/min.

(d) 10 µM/min.

(e) I don't know.

TASK 9.

Consider the following titration curve for an amino acid and answer the questions.

9.1 Circle the correct statement:

(a) 'pH' is the independent variable and 'Equivalents of OH^-' is the dependent variable.

(b) Both 'pH' and 'Equivalents of OH^-' are dependent on each other.

(c) 'Equivalents of OH^-' is the independent variable and 'pH' is the dependent variable.

(d) I don't know.

Give a reason for your answer:………………………………………………………...……

………………………….………………………………………………………………………

..………………………………………………………………………………………………

9.2 At which of the lettered points on the graph is the change in pH the slowest?

Answer: …………………

Explain your answer: ……………………………………………………………………

……………………………………………………………………………………………………

9.3 At which of the lettered points on the graph is the change in pH the fastest?

Answer: …………………

Explain your answer: …………………………………………………………………….

……………………………………………………………………………………………………
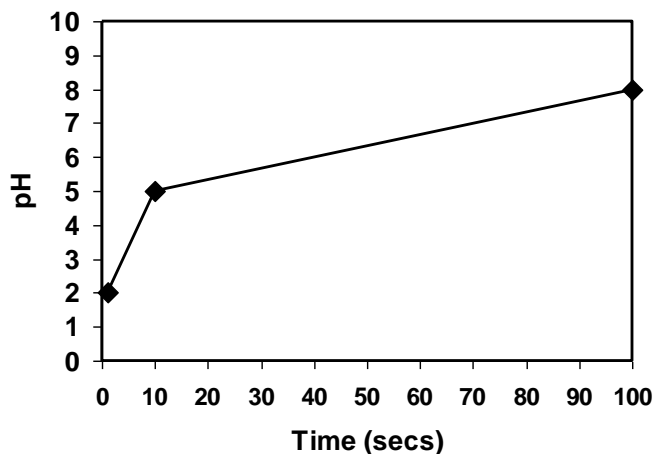
TASK 10.

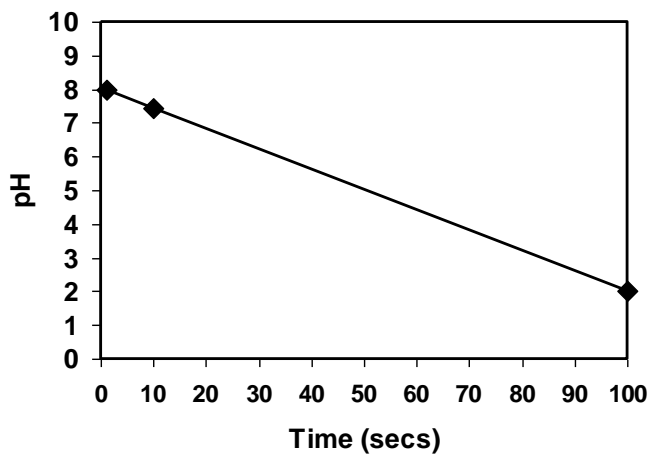The following figure is a graph of pH versus time for a reaction:
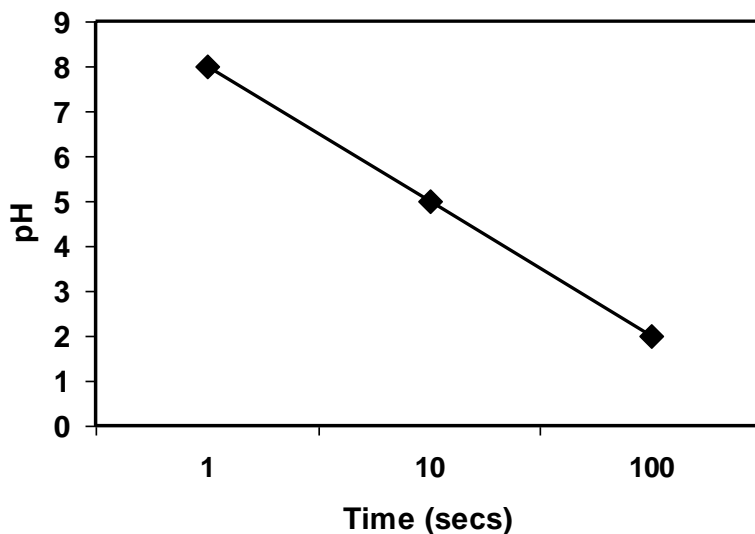
Which of the following three graphs would be an equivalent graph of the above graph displaying the same quantitative data? Choose your answer by circling a, b, c, or d
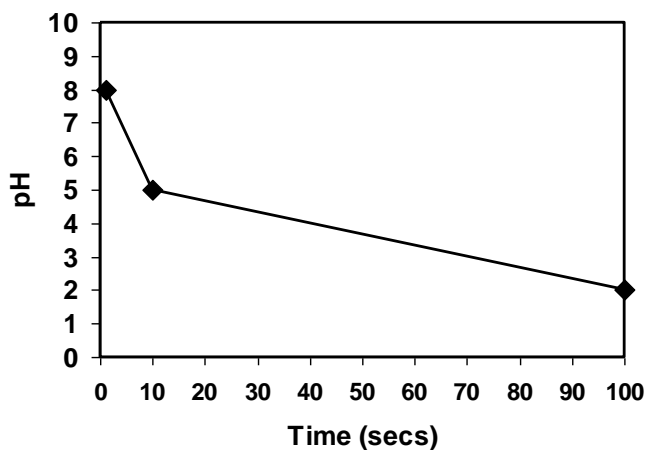
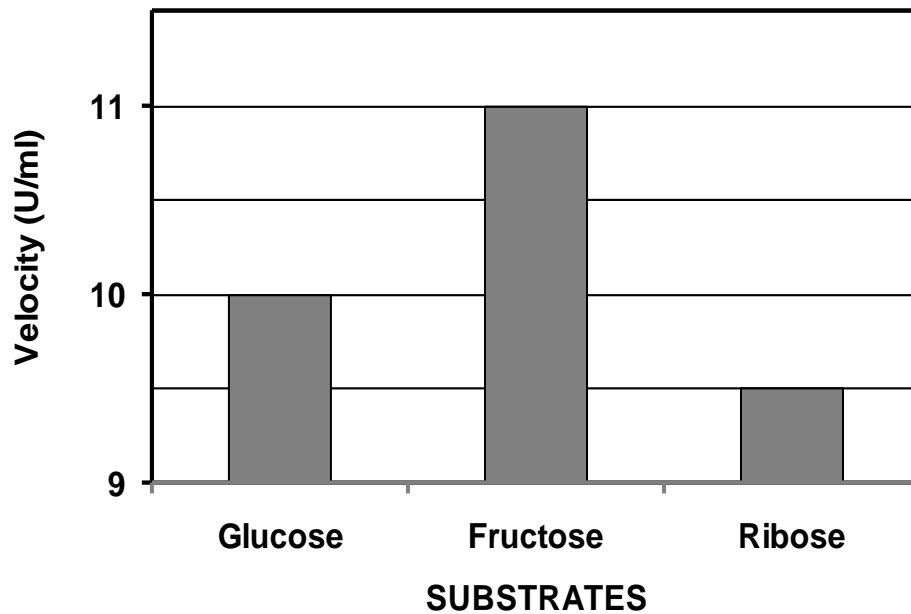
(a)   Graph 1

(b)   Graph 2

(c)   Graph 3

I DON'T KNOW

(d)   Don't know

TASK 11.

Consider the following figure, a graph showing the velocity of an enzyme-catalysed reaction obtained with three different substrates, and answer the questions.

Choose the correct answer by circling a,b,c, d or e:

11.1 The velocity obtained with glucose is:
  (a)   half that obtained with fructose;
  (b)   double that obtained with ribose
  (c)   half that obtained with fructose and double that obtained with ribose
  (d)   none of the above
  (e)   I don't know.

11.2 The velocity obtained with fructose is:
  (a)   double that obtained with glucose;
  (b)   four times that obtained with ribose
  (c)   10% higher than that obtained with glucose
  (d)   none of the above
  (e)   I don't know.

TASK 12.

The following figure is an elution profile obtained after the elution of a protein from an ion-exchange chromatography column. The protein was eluted with a sodium chloride gradient and was detected by measuring absorbance at 280 nm. Fraction volumes were 5 ml of eluant.
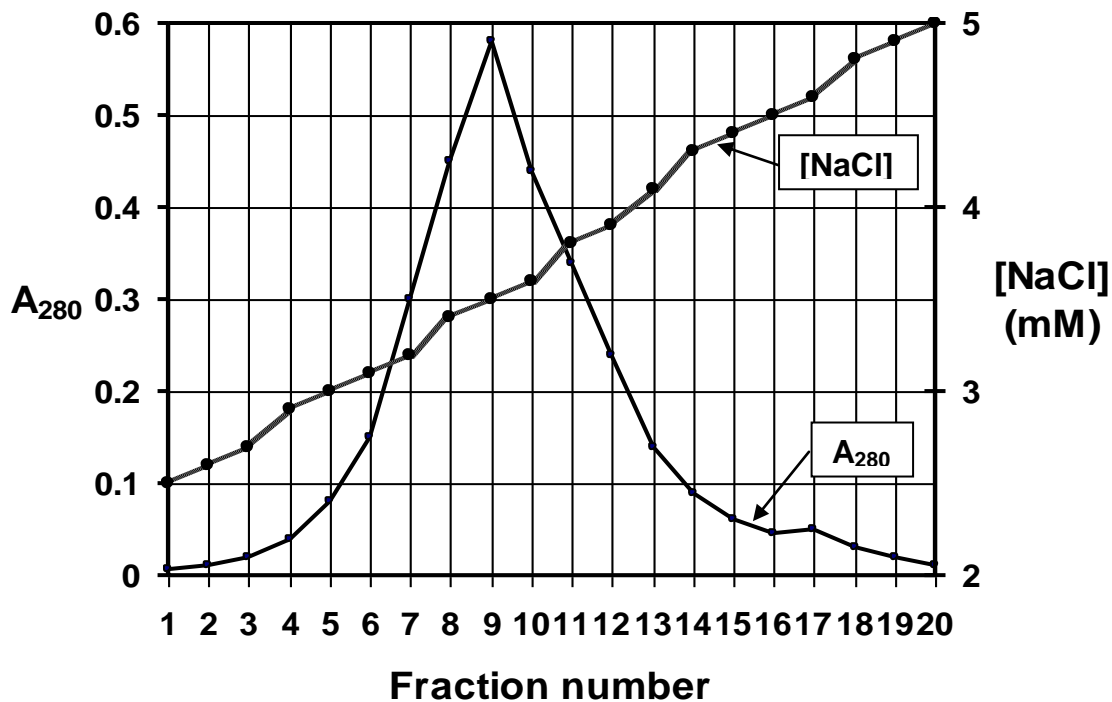
12.1 The sodium chloride gradient was: (circle a,b,c, or d)

(a) 0.1 mM to 0.6 mM

(b) 0.1 mM to 5 mM

(c) 2.5 mM to 5 mM

(d) I don't know

12.2 What fraction contained the most protein? Answer: ………………..

12.3 What was the sodium chloride concentration of the fraction with the highest protein concentration? Answer: ………….

12.4 This graph can be described as follows: (circle a,b,c, or d)

(a) As the sodium chloride concentration increases the protein is released from the column and elutes in a few fractions as an absorbance peak.

(b) As the protein appears in the fractions the sodium chloride concentration increases as a result of the presence of the protein.

(c) As the protein concentration increases, the sodium chloride concentration increases to a maximum and then decreases.

(d) I don't know

TASK 13.

The following figure is a graph of the inverse of reaction velocity (1/V) versus the inverse of substrate concentration (1/[S]) for two different enzymes, A and B. Substrate concentration was in mM. The intercepts on the x- and y-axis are used to calculate the maximum velocity (V_{max}) and Michaelis constant (K_m).
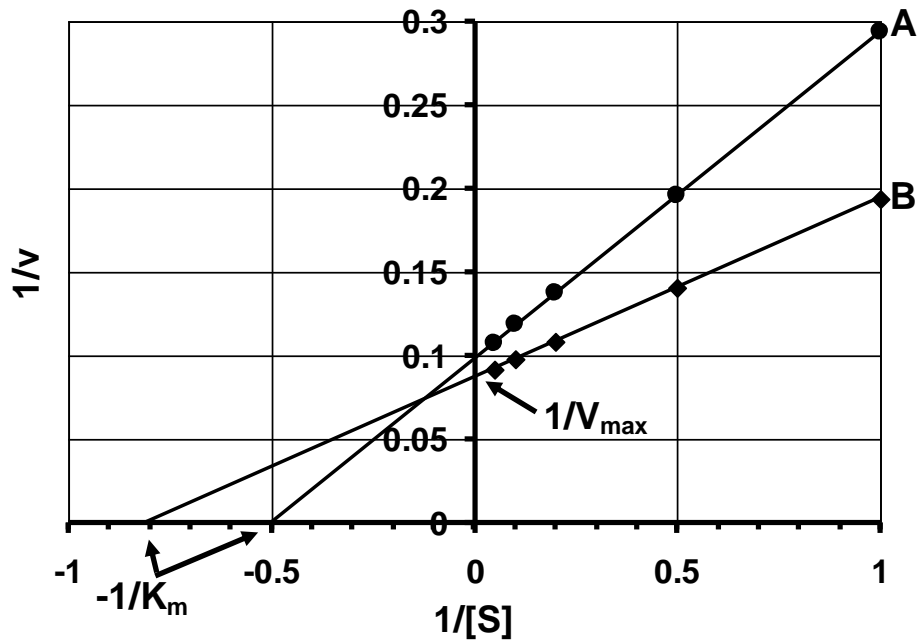
13.1   What is the value of V_{max} for enzyme A?

Answer: ……………………

13.2   Which enzyme has the highest velocity at a substrate concentration of 1 mM?

Answer: ……………………

13.3   Which enzyme has the lowest K_m value?

Answer: ……………………

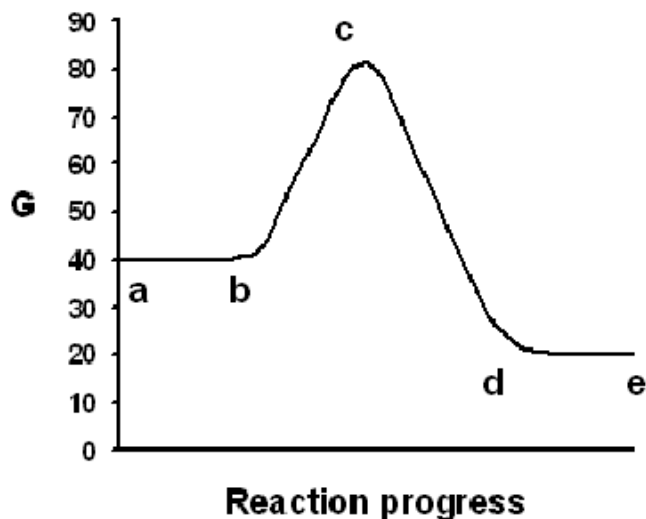**13.4**   Do enzymes A and B ever reach the same velocity and, if so, at which substrate concentration would this occur?
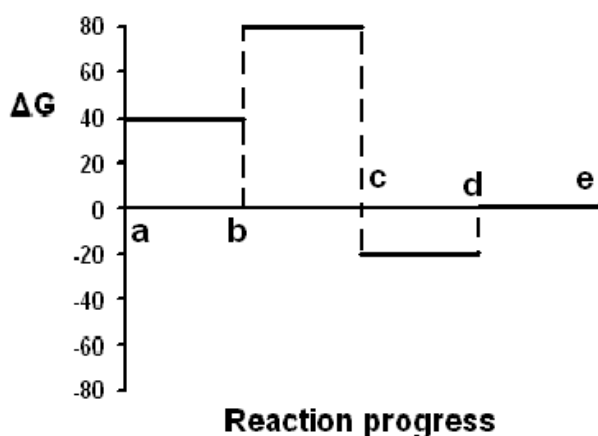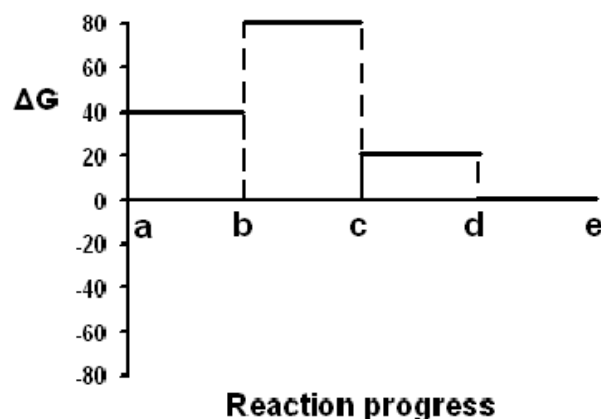
Answer: ……………………………………………….………………..

TASK 14.

The following graph shows free energy (G) versus reaction progress for a reaction proceeding from reactants to products.
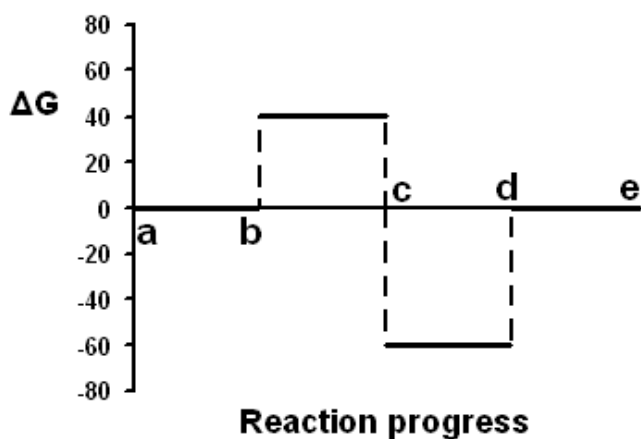

From the following three graphs, **choose the correct ΔG vs reaction progress** graph showing the change in free energy (or free energy difference, ΔG) as the reaction proceeds from a to e (circle the graph number).
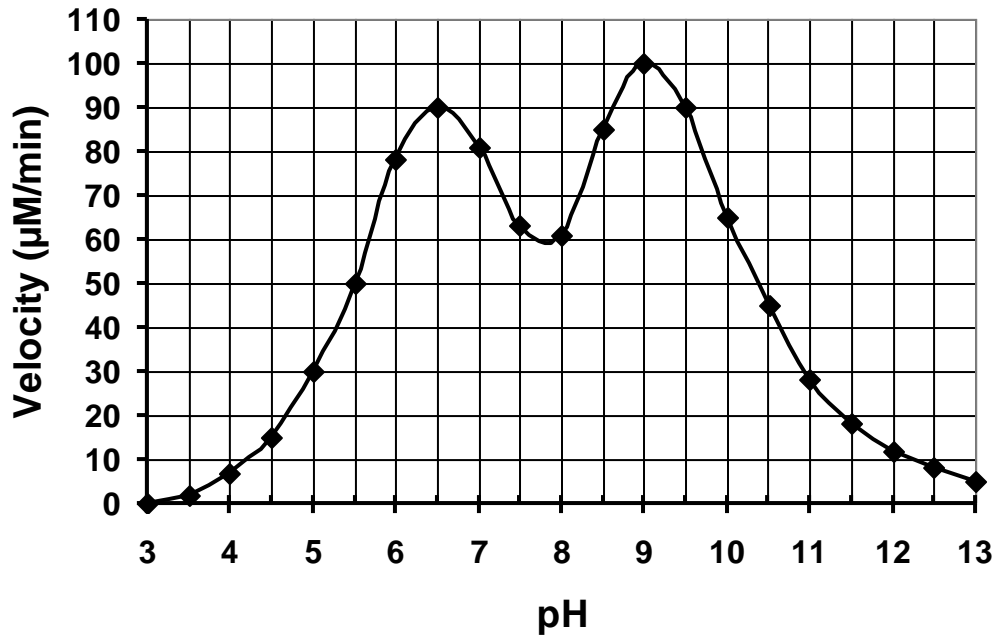

Graph 1

Graph 2

Graph 3

I DON'T KNOW

I don't know

TASK 15.

Consider the following graph:

15.1 The velocity observed for the second peak is: (Choose the correct answer by circling a,b,c,d or e):

(a) 25% higher than that observed for the first peak;

(b) four times higher than that observed for the first peak

(c) 10 µM/min higher than that observed for the first peak

(d) 100 µM higher than that observed for the first peak

(e) I don't know.

15.2 Describe the graph (i.e. what it is showing):

……………………………………………………………………………………………...……

…………………………………………………………………………………………………….

..…………………………………………………………………………………………………….

15.3 Describe a possible reason for the observed result:

…………………………………………………………………………………………...………..

..…………………………………………………………………………………………………….

..…………………………………………………………………………………………………….

TASK 16.

In a series of assays, a researcher measured the initial reaction velocity of an enzyme-catalysed reaction using a spectrophotometer equipped with a temperature-controlled cuvette holder.   The following figure, a graph of reaction velocity versus assay number, shows the results obtained.
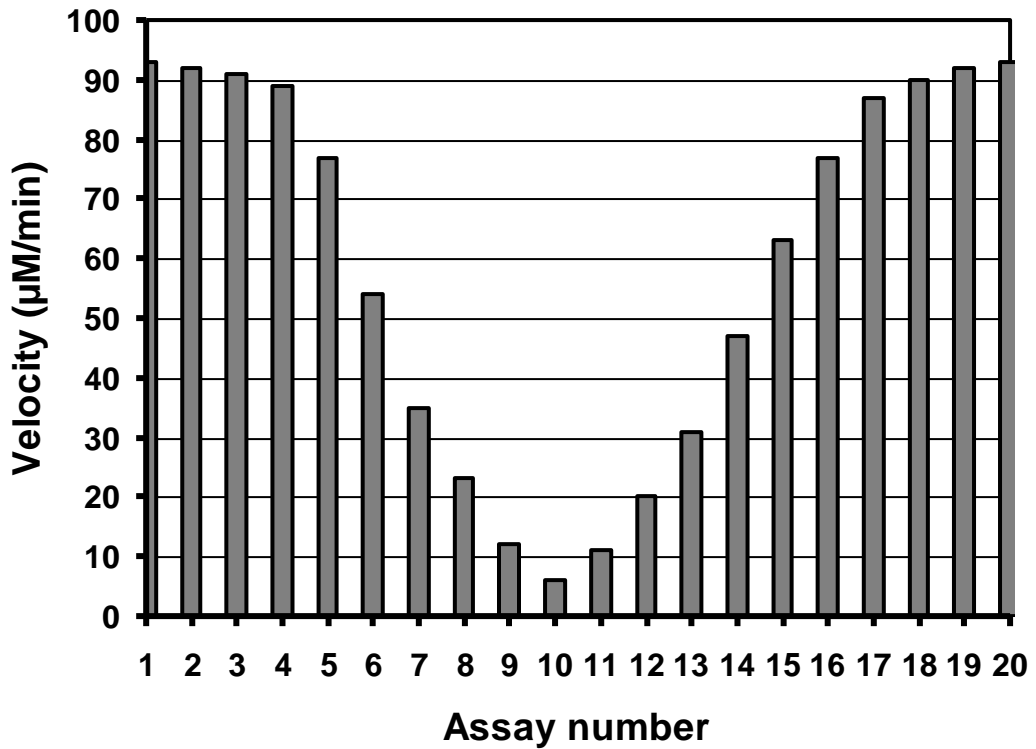
Briefly describe the possible sequence of steps the researcher could have carried out to obtain the observed results:

……………………………………………………………………………………………………..……..

………

…………..………………………………………………………………………………………………………

…………..………………………………………………………………………………………………………

Percent correct responses per question

STUDENT PERFORMANCE AT DIFFERENT LEVELS
(KIMURA)
% CORRECT PER LEVEL

APPENDIX C

133

STUDENT PERFORMANCE AT DIFFERENT LEVELS (CURCIO)

% CORRECT PER LEVEL

Legend: READ DATA, BETWEEN DATA, BEYOND DATA

STUDENT #

APPENDIX D