Paper appears in: D. B. Zandvliet & J. Osborne (Eds) (2006), *Proceedings of the National Association for Research in Science Teaching (NARST) Annual Meeting*, San Francisco, United States, 3-6 April, electronic paper No. 210848

A MODEL OF FACTORS DETERMINING STUDENTS' ABILITY TO INTERPRET EXTERNAL REPRESENTATIONS IN BIOCHEMISTRY

Research into the role of external representations (ERs)[1] in science education has shown that ERs are not always powerful vehicles for learning and may often cause unexpected learning difficulties. The aim of this study was to develop a model of factors affecting students' ability to interpret external ERs used in biochemistry. The study was qualitative in design and employed the modelling framework of Justi and Gilbert (2002) to express the model. To validate the model, nine students were each interviewed thrice, using a clinical instrument, during their interpretation of three ERs, each representing antibody-antigen interaction. The data was analyzed by induction, where response patterns of interest emerged naturally rather than being predisposed. Empirical testing of the model allowed its seven component factors namely, the conceptual (**C**), reasoning (**R**), representation mode (**M**), reasoning-mode (**R-M**), reasoning-conceptual (**R-C**), conceptual-mode (**C-M**) and conceptual-reasoning-mode (**C-R-M**) to be verified and operationalized. The data suggests that in addition to framing researchers' thinking about students' interpretation of ERs in biochemistry, the model can potentially be used as a tool in any science education context to measure the *interaction* of the afore-mentioned factors and as a template from which to design strategies for improving ER interpretation.

Konrad J. Schönborn[a], Georg-August-Universität, Germany

Trevor R. Anderson[b], University of KwaZulu-Natal, South Africa

Introduction

External representations (ERs) such as static pictures, diagrams, graphs, photographs, micrographs, maps, flowcharts and computer-based dynamic visuals are usually assumed by science instructors to be excellent learning tools for constructing knowledge (e.g. Lowe, 2004). However, various research reports (e.g. Wu and Krajcik, 2006) have suggested that ERs do not *always* improve understanding and may in fact induce a range of conceptual and reasoning difficulties. This problem is largely due to naïve assumptions by instructors and textbook authors that what works for experts will be equally beneficial for novices (e.g. Schnotz and Lowe, 2003). Therefore, it is of paramount pedagogical importance that educators and authors of educational materials strive to better understand the factors that are

---

[1] Abbreviations used in this paper are: ER, external representation; Ab, antibody (immunoglobulin G) molecule; Ag, antigen molecule; I, interviewer; S, student; ELISA, enzyme-linked immunosorbent assay.

[a] Conducted this research within the Science Education Research Group (SERG), University of KwaZulu-Natal (Pietermaritzburg).

[b] Correspondence should be addressed to this author at the Science Education Research Group (SERG), School of Biochemistry, Genetics, Microbiology and Plant Pathology, University of KwaZulu-Natal (Pietermaritzburg), Private Bag X01, Scottsville 3209, South Africa, e-mail: Anderson@ukzn.ac.za, tel: +27 (33) 260-5464/29, fax: +27 (33) 260-5462.

responsible for the inadequate processing and interpretation of scientific ERs so that the difficulties can be addressed and remediated. As a response to this objective, we conducted a study in the context of biochemistry education, an area in which there has been a lack of research on students processing of ERs (Schönborn *et al*., in press).

In response to the shortage of research on students' processing of ERs in biochemistry, our findings in this context will serve to supplement recent findings concerned with students' interpretation of ERs in other domains (e.g. Lowe, 2004). Like other authors, we argue that an understanding of students' misinterpretation of scientific ERs requires ongoing research especially in domains where the visual content available to learners is constantly changing and being increased at a high rate. This argument is born out of the fact that the *way* ERs are processed is poorly understood by researchers because a huge diversity of ER forms is available to learners, each with their own instructional goals (e.g. Blackwell, 2001). In this regard, Ploetzner and Lowe (2004) have also alluded to the fact that when engaging in such studies, workers should make a concerted effort to consider the *cognitive* dimensions associated to learning with ERs so that researchers, educators and ER designers alike, are in a better position to measure what factors influence students' processes of ER interpretation. As a response to this general objective, the current study in the context of biochemistry education may contribute to a better understanding of the factors that affect students' interpretation of ERs in science as a whole.

This study used the modeling framework of Justi and Gilbert (2002) to develop and test a model of the factors that influence students' interpretation of ERs used in the teaching and learning of biochemistry. Concerning our approach to analyzing the data for expressing the model, we aligned ourselves with other ER researchers (e.g. Mayer, 2003) in arguing that a constructivist epistemology serves as a feasible platform from which to model the factors that affect students' ability to interpret ERs used in the teaching and learning of biochemistry. In this regard, our approach to obtaining and treating data in this study data could best be described as a "descriptive synthesis" rather than a process of data reduction (McMillan and Schumacher, 1993, p. 480) and is consistent with a postpositivistic approach to science education research.

In line with the arguments above, the present study addressed the following research questions. Firstly, what factors affect students' ability to interpret external representations (ERs) in biochemistry? Secondly, can the factors be incorporated into, and expressed as an appropriate model? Thirdly, how can empirical data be obtained to investigate the nature of the factors and to validate of the model? Fourthly, what practical applications will the model have and, will it be generalizable to all ERs in science?

Methods

Student group and external representations used in the study

The study was conducted from 2001 to 2002 with nine biochemistry students at the University of KwaZulu-Natal, South Africa, who had all completed a third-year level module on immunology. Each student was interviewed three times, an interview for each of three different ERs of antibody-antigen structure giving a total of 27 interviews. The three ERs used in the study are shown in Fig. 1 (E - G) and are multiple representations of antibody-antigen interaction that fall on a real to abstract continuum (e.g. Wheeler and Hill, 1990). The electron micrograph (Fig. 1 E) can be considered a “real” depiction of antibody and antigen interaction, the space-filling model (Fig. 1 F) a “semipictorial” (stylized) representation of antibody-antigen interaction and the graphical plot (Fig. 1 G) an “abstract” portrayal of antibody-antigen interaction. The electron micrograph (Fig. 1 E) shows trimer and pentamer complexes formed when Y-shaped IgG antibodies bind to the divalent hapten dinitrophenyl (DNP) (Valentine and Green, 1967). Fig. 1 F represents a three-dimensional, space-filling display of the binding of an antigen (lysozyme protein) to a Fab fragment of an IgG antibody molecule (Amit *et al*., 1986). Lastly, Fig. 1 G is a Cartesian graph of the quantitative results obtained from an enzyme-linked immunosorbent assay (ELISA) (Jackson, pers. comm.) of the binding interaction between antibody and antigen molecules. Each colored curve represents results obtained at different weeks of an immunization schedule. Absorbance at 405 nm is plotted against the negative logarithm of antibody concentration. This paper shall refer to each of the ERs in Fig. 1 as "ER E", "ER F" and "ER G", respectively. For each ER, both the ER and its caption were supplied to students during all interviews but only one ER was supplied at a time.

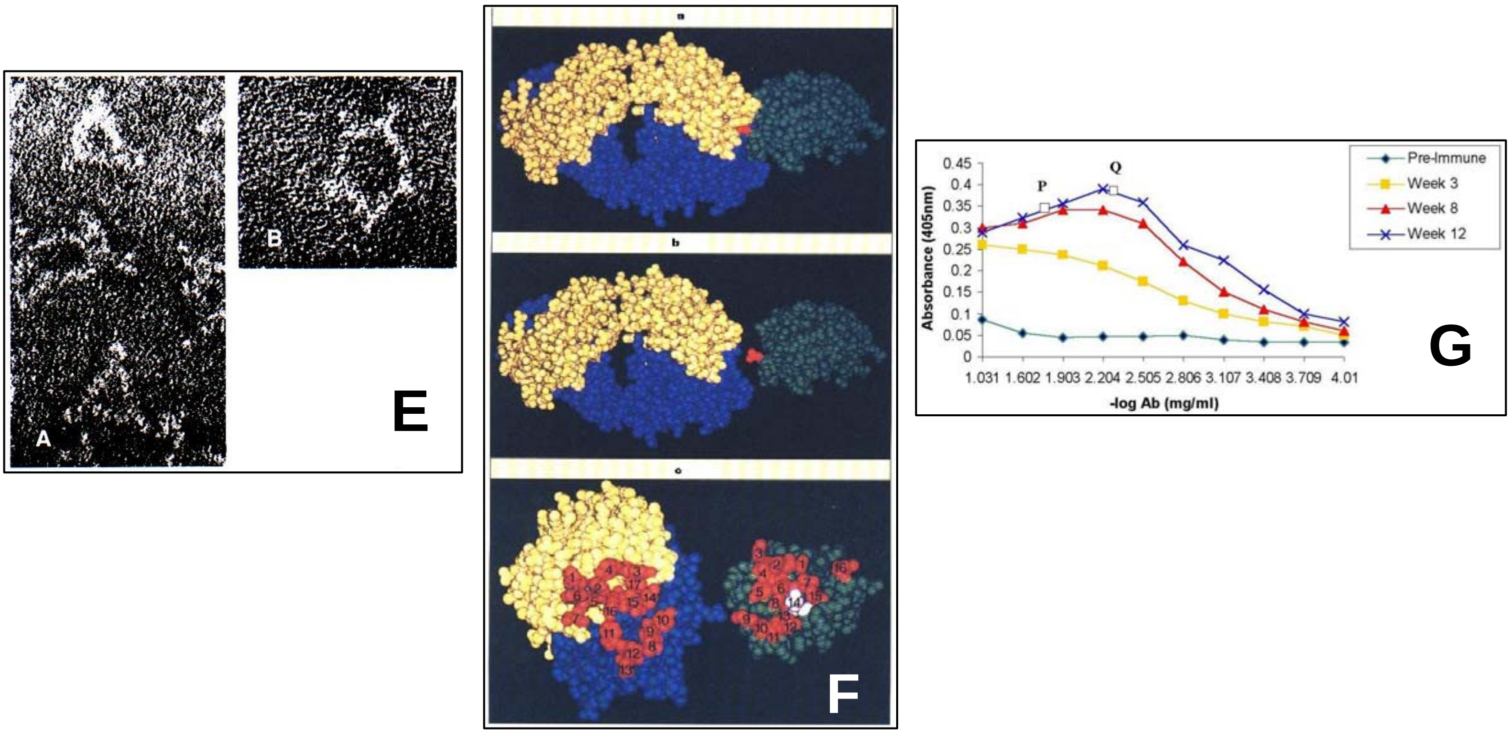

Figure 1. Three multiple ERs of antibody-antigen interaction, ***E****) Electron micrograph (x 1 000 000) of complexes formed on mixing divalent hapten with anti-hapten antibodies. The hapten links together the Y-shaped antibody molecules to form trimers (A), and pentamers (B) (Roitt, 1997),* ***F****) Space-filling model showing Fab antilysozyme and lysozyme molecules fitting snugly together. Antibody heavy chain, blue; light chain, yellow; lysozyme, green with its glutamine 121 in red. Fab and lysozyme models are also shown pulled apart in the second frame (Roitt, 1997),* ***G****) Antibody response curves obtained from an ELISA showing the relationship between absorbance (405nm) and antibody concentration (mg/ml). Three booster shots were administered and the antibodies collected at the weeks indicated in the text box (Jackson, pers. comm.).*

Development of a model of factors

The modeling process of Justi and Gilbert (2002, p. 371) was used to develop a model of factors affecting students' ability to interpret external representations (ERs) in biochemistry. Expression of the model involved a five-stage cyclical process. Firstly, the *purpose* of the model was decided upon based on previous research by the authors (Schönborn *et al*., 2002), the current authors' prior knowledge and experience of student difficulties with ERs, and a thorough analysis of the literature (see Schönborn, 2005). Secondly, a *mental model* was constructed and thirdly, the mental model was externalized as an *expression model*. Fourthly, conduction of various *thought experiments* as well as extensive discussion of the expression model between the authors helped decide on the validity of the model and whether to modify it. Stages 2 – 4 were repeated several times so as to optimize the expression model. Fifthly, empirical tests were designed and performed in order to decide whether to further modify, reject or accept the model as a *consensus model*.

Empirical testing of the model

The model was tested empirically using an adaptation of general qualitative clinical interview methods (e.g. White and Gunstone 1992) common in science education research (e.g. Orgill and Bodner, 2004). In this regard, an interview instrument termed the *three-phase single interview technique* (3P-SIT) was specially designed and piloted to gather data on students' conceptual understanding and reasoning ability (see Schönborn, 2005). For details on the nature of this interview method, an extensive description is represented at the present NARST conference (see paper number 210971). Empirical data generated from 3P-SIT was used to investigate the nature of the factors of the model, to formulate clear operational definitions for each component factor, to test the validity of the model, and to establish the *nature of interaction* between the factors of the model.

All 3P-SIT interviews were audiotaped and videotaped (e.g. Sumfleth and Telgenbüscher, 2001). The data collected during the interview sessions consisted of 27 video segments, 27 audio-transcripts, 134 student-generated diagrams (SGDs) and 27 researcher-generated field note items. Data was analyzed by means of a qualitative method in which categories of responses were "uncovered" (e.g. Lincoln and Guba, 1985, p. 203) rather than being predisposed. In this regard, data was analyzed by means of an inductive method in which categories of responses emerged from the data themselves (e.g. Grayson *et al*., 2001), and in which patterns were "made explicit from embedded information" (Lincoln and Guba, 1985, p. 203). In addition to exposing data for operationalizing the definitions of the component factors of the model, the interrelationships between the data across the 3P-SIT interview phases were investigated in an attempt to measure how successfully the ER was interpreted and, whether sound or unsound learning had occurred after exposure to the ER.

Results and Discussion

Development of the model and its constituent factors

The modeling process of Justi and Gilbert (2002) enabled us to successfully develop the model presented in Fig. 2 below.

The model, composed of seven factors (**C**, **R**, **M**, **R-C**, **R-M**, **C-M** and **C-R-M**), is represented in the form of Venn logic since our data suggests an interaction between the factors. Empirical testing of the model using 3P-SIT (see paper number 210971) allowed operational definitions for each factor to be established. For the purposes of clarity and for the reader to better appreciate the nature of the model, the operational definitions that were derived from the data are presented first, while the corresponding empirical evidence is presented afterwards.

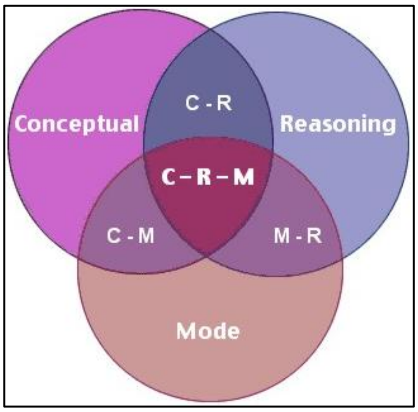

Figure 2. Venn diagram representing a model of seven factors that determine students' ability to interpret ERs. The model expresses three factors and four interactive factors affecting students' ability to interpret an ER

We defined the conceptual factor (**C**) of the model as the conceptual understanding and prior knowledge that a student holds *before* exposure to any ER. It embodies the collection of a student’s conceptual frameworks and mental models of relevance to the ER, including any alternative conceptions. Since “reasoning” is a cognitive *process*, one has to have something to reason “with” and therefore, reasoning processes cannot be defined in isolation. In terms of the model, we defined the reasoning factor (**R**) as representing those *cognitive processes* that a student employs when reasoning with the ER and with his/her own conceptual knowledge of relevance to the ER. More specifically, factor **R** represents a student's total reasoning ability, i.e. the skills needed to decode and perceive visual markings on an ER, to access and retrieve conceptual knowledge from long term into working memory (e.g. Baddeley, 1992) in order to perform ER-related reasoning; and, to assimilate information that is first perceived from an ER and then incorporated into already existing knowledge. Unlike a conceptual difficulty, which is context-dependent, a *reasoning difficulty* is *independent* of context (e.g. Grayson *et al*., 2001) and can be observed in multiple scientific content areas. The representation *mode* factor (**M**) of the model encapsulates the nature of the ER. This includes characteristics of the ER such as the graphical and diagrammatic features, the spatial arrangement of the ER elements, ER conventions, visual icons, visual cues, artistic devices, color, topography, level of abstraction, symbols, labels, linguistic captions and so on. Factor **M** can be considered distinct from both **C** and **R**, since it does not depend on any human constituent during the interpretation process and remains constant unless the ER is modified.

The **R-C** interactive factor, representing the relationship between the reasoning (**R**) and conceptual (**C**) factor, includes cognitive processes such as when a student selects, retrieves, actively adjusts or adds to their existing knowledge. **R-C** is indicative of a student's ability to reason with their conceptual knowledge of relevance to the ER because, in effect, they are using the collection of their concepts to 'think about something' or to 'solve a problem'. Congruently, within **R-C**, cognitive processes such as assimilation and accommodation can also be represented. The **R-M** interactive factor between the representation mode (**M**) factor and the reasoning (**R**) factor exemplifies a student's ability to decipher and reason with an ER and its graphical features. For instance, when reading an ER, a student will employ perceptual mechanisms such as recognition and organization of patterns, shapes and colors, visualization (McCormick *et al*., 1987), distinguishing relationships between ER features (e.g. Lowe, 1993) and mentally organizing the topographical information on the ER. The **C-M** interactive factor of the model was defined as representing the nature of the conceptual (propositional) knowledge represented by the ER, including the extent, complexity and soundness of such knowledge. It also includes both the conceptual knowledge represented by the graphical markings and symbolism used to construct the ER, and knowledge of the meaning of the symbolism and conventions employed in the ER to communicate the science.

The **C-R-M** interactive factor represents a student's ability to engage all factors of the model, by utilizing their reasoning skills (**R**) to reason with both their conceptual knowledge of relevance to the ER (**C** and **R-C**) and with the ER itself (**M** and **R-M**) so as to successfully interpret, visualize and learn from the propositional knowledge represented by the ER (**C-M**). For example, the process could take the following form. Upon reading the ER, the individual deciphers and decodes the visual information on the ER (**R-M**) and, in so doing, links their interpretation to, and filters their interpretation through, already existing current knowledge (**R-C**). The outcome of this process could result in the construction of a unique conception either consistent or inconsistent with accepted scientific knowledge (**C-M**). This scenario would depend on a combination of all three factors (**C-R-M**), during which all factors comprising the model would, at some time or other, be engaged.

<u>Empirical validation of the model</u>

The following empirical data validated the model and its component factors and informed the development of the above operational definitions for each factor of the model.

<u>Validation of the Conceptual (**C**) Factor</u>

Students' prior and current knowledge, i.e. the conceptual factor (**C**) is one component of the model (Fig. 2) that affects' the ability to interpret an ER. For example, consider the following student quotations obtained from two students before exposure to any ER of interest (Fig. 1). The first quote displays a sound conception of the nature of antibody structure and interaction with antigen while the second represents an alternative conception of the same phenomenon:

> S: The structure of an antibody…consists of four chains… two light chains and two heavy chains. On the N-terminal is where the antibody binds to the antigen…one antibody can bind to two antigens… there are two binding sites for binding two antigens.

> S: This is the antigen [inserts and labels Ag on Fig. 3] …ja [yes], the antigen. And the antibody would be like that [inserts top right Ab]… it [Ab] forms a complex when it binds on. That would be like one antibody to one antigen… the normal thing that happens is one antigen to one antibody…

In the first case, the student shows a sound conception of Ab structure and binding to Ag. However, the second student holds the alternative conception that an Ab only has *one* possible binding site for an Ag and that this site is the entire 'V' cleft of the Y-shaped Ab, instead of *two* variable binding domains. This alternative conception is supported by the following SGD (Fig. 3) which the student produced:

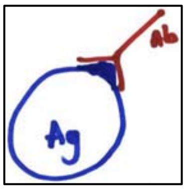

Figure 3. Student-generated diagram obtained during a verbal explanation of Ab-Ag binding

Both the examples above serve to support the **C** factor of the model. The **C** factor is clearly a key component of the model in that the nature of a student's prior knowledge will seriously affect their ability to interpret an ER representing such knowledge (e.g. Lowe, 1993; Winn, 1993).

Validation of the Reasoning (**R**) Factor

The **R** factor of the expressed model (Fig. 2), represents those *cognitive processes* whereby students reason with both the ER and their own conceptual knowledge in order to interpret an ER. This study identified at least five different reasoning mechanisms associated with students' interpretation of ERs. Firstly, some students employed *surface-level reasoning* (Chi *et al*., 1981) when processing the external markings on the ERs. These students interpreted the ER markings literally and at face value, without considering the deeper meaning of the markings (e.g. Ametller and Pintó, 2002; Cheng *et al*., 2001). Secondly, some students performed *inappropriate analogical reasoning* when interpreting the ERs (e.g. Sumfleth and Telgenbüscher, 2001). Thirdly, some students engaged in *inappropriate transfer* (Salomon and Perkins, 1989) when interpreting the ERs. Fourthly, some students found it difficult to *translate* between different ERs, which all represented the same concept or phenomenon (e.g. Gobert and Clement, 1999; Ainsworth *et al*., 1998), i.e. Ab-Ag binding in this case. Fifthly, we also discovered what we have termed the *superimposing* of one concept upon another. Here, some students tended to fuse two or more distinctively different concepts together into a single explanative model, often leading to alternative conceptions. The superimposing of concepts could be related to a recent finding by Grayson (2004), who found that some students struggle to *disentangle* distinctively different scientific concepts from one another.

Data corresponding to the **R** factor of the model is presented below to support the influence and importance of the above reasoning processes as components of the **R-M** and **R-C** factors of the model. In this regard, reasoning "as a process" acquires meaning only when it is observed in *action*. In other words, reasoning processes can only be observed if there is something to reason with, in this case with the ER (**R-M**) and with students' own conceptual knowledge (**R-C**). Hence, the **R-C** and **R-M** factors of the model can each be considered a subset of the overall reasoning factor (**R**) (see Fig. 2).

Validation of the Reasoning-Mode (**R-M**) Factor

As discussed earlier, the **R-M** factor represents the reasoning processes that a viewer employs when reading and deciphering the graphical markings in an ER. The following student quote provides empirical evidence for sound processing of the ER markings in Fig. 1 F and therefore supports the **R-M** factor of the model:

S: …[frame] 'c' shows what is involved in the binding… it shows the actual atoms involved in the binding… by highlighting the specific atoms and numbering them…

In contrast to the quote above where the topographical and colored features shown in 'frame c' of ER F were successfully processed, some students struggled to process the same graphical markings. This finding is supported by the following student quotation:

S: The antibody has receptors that go into this molecule [points to lysozyme on frame c] and then works on it [Ag] and breaks it [Ag] down… yeah, and that is how you get this glutamine [points to red spheres on frame c].

The nature of the **R-M** factor as a component of the model (Fig. 2) can be shown by presenting how both sound and unsound processing of the markings making up an ER can occur. Later, we shall demonstrate how such processing influences the overall manner in which an ER is interpreted.

<u>Validation of the Reasoning-Conceptual (**R-C**) Factor</u>

In addition to cognitive processes represented by the **R-M** factor of the model, interpreting an ER also requires a student to engage their conceptual understanding of the scientific phenomenon that is represented by the ER. A student's ability to *reason* with their conceptual knowledge of relevance to the ER is represented by the **R-C** factor of the model (Fig. 2).

The 3P-SIT technique (see paper number 210971) allowed us, once the nature and extent of an individual's conceptual knowledge (**C**) was ascertained, to measure *how* this knowledge was engaged by a student when interpreting an ER. For instance, consider the following quotation and corresponding SGD obtained from a student during the interpretation of ER G:

S: Yes, this is the primary one [green Ab]… it has got this C-terminal and C-terminal. So, when the secondary one [red] comes… it has an NH_3 [inserts on top right diagram] and will interact with this one [green Ab]…
I: How does what you were telling me earlier about lock-and-key… relate to what you're telling me about how the secondary [Ab] interacts with the primary [Ab]…
S: No, no. It is not like the lock-and-key. Because, the lock-and-key was due to the structure of the binding site and also due the antigen, how it [Ag] looks.

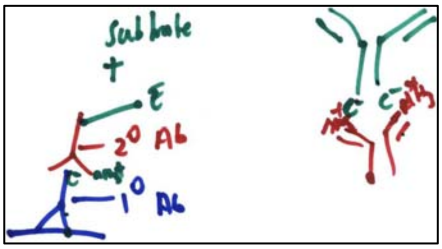

Figure 4. SGD showing how a student's conceptual knowledge influences the processing of an ER.

From analysing the above data, it is clear that the student was unable to transfer his earlier expressed idea of a lock-and-key binding situation to the context of an ELISA set-up in biochemistry. In this and other similar cases, it appeared that some students treated Ab-Ag interactions within the context of an ELISA as unique situations in some way different to otherwise identical Ab-Ag interactions. A source of this finding could reside in the way ELISA plate "set-ups" are conventionally represented pictorially. Often, they are represented with the Ab and Ag components arranged in a hierarchical, vertical and "sandwich-like" fashion.

In support of the above finding that validates the **R-C** factor of the model, other ER research has also shown that the interpretation of an ER depends to a large extent, on the knowledge that an individual "brings" to the ER (e.g. Roth, 2002; Cheng *et al*., 2001). In this regard, the above data suggests that students' interpretation of an ER is significantly affected by reasoning processes represented by the **R-C** factor, an essential component affecting students' ability to interpret scientific ERs.

Validation of the Representation Mode (**M**) Factor

As discussed previously, the **M** factor is concerned with the external information that corresponds to the nature of the ER in isolation and how well (or poorly) the graphical markings that constitute the ER represent what it is designed to represent. By validating the **M** factor, some of the external characteristics of an ER of interest (Fig. 1) that may cause student difficulties can be identified. In other words, data corresponding to the **M** factor consists of the nature, use and clarity of ER the features. This data can then be used to ascertain what external features of the ER may be initiating particular reasoning patterns.

Data pertaining to factor **M** can be obtained from experts including scientists and graphic artists as well as students' evaluation of an ER. For example, consider the following student's quote which provided an example of a datum from ER E that was used to model factor **M**:

> I: Is there anything that you don't understand or find confusing on this representation [ER E]?
> S: …The only thing is like…where the bonds form between the different antibodies.

The above student demonstrates how the graphical features representing the nature of the visual clarity of the trimer and pentamer Ab-Ag complexes influenced his/her reasoning. In this case, the student thought that the Y-shaped antibodies were somehow "joined" together. Due to the clarity of the visual information on the micrograph (ER E), it is impossible to actually "see" the hapten (antigen) molecules and, from a purely visual perspective, the antibodies do look like they are fused without hapten.

Data such as that presented above helps to locate and identify specific ER features, which may be responsible for inducing difficulties. The datum above has validated the importance of the **M** factor of the model as an integral factor contributing to students' ability to interpret ERs in science.

Validation of the Conceptual-Mode (**C-M**) Factor

The **C-M** factor encapsulates the nature of the conceptual (propositional) knowledge represented by the ER and its symbolism and markings. Like for factor **M**, it requires experts to judge or evaluate something in isolation from student interaction with an ER - in this case the propositional knowledge conveyed by the ER. Such propositional knowledge is obtained from textbook authors', from surrounding text that describes an ER, from the captions (figure legends) used by writers to describe an ER as well as from scientific findings that are presented in journals and other documents. In the current study, data corresponding to the **C-M** factor was obtained from the primary sources where the ERs were located and described.

The conceptual (propositional) knowledge represented by **C-M** is an indispensable factor that affects a student's interpretation of an ER. This is because the complexity, soundness and extent of knowledge that the ER represents will have a profound affect on how well the ER is interpreted.

Validation of the Conceptual-Reasoning-Mode (**C-R-M**) Factor

The previous sections of this paper have separately validated the three factors **C**, **R** and **M** and the three interactive factors **R-M**, **R-C** and **C-M** that influence a student's ability to interpret an ER. The current aim of this paper is to confirm the validity of the model as an integrated whole as implied by the overlapping nature of the factors (**C-R-M**) in the Venn logic used to represent it in Fig. 2. Therefore, to test the validity of the model as a whole unit (i.e. validate the **C-R-M** factor), data was needed to demonstrate that, at some time or other, students are required to engage *all* factors of the model in order to successfully interpret an ER. That is, the *indispensable* nature of each component of the expressed model (Fig. 2) needs to be confirmed. Our hypothesis is that interpretation of an ER requires the learner to use reasoning skills (**R**) to reason with both their conceptual knowledge (**C** and **R-C**) of relevance to the ER and with the markings on the ER itself (**R-M** and **M**) to make sense of the propositional knowledge represented by the ER (**C-M**).

The validation of factors **R-M** and **R-C** and therefore, validation of the **C-R-M** factor, was done by using a red colour to code engagement of the **R-M** factor and a blue colour to code engagement of the **R-C** factor of the model during students' interpretation of an ER. Coding of the transcript text in this manner enabled the authors to establish whether all factors were engaged during the interpretation process. In view of this rationale, the criteria for coding segments of the interview extracts either as corresponding to the **R-M** or **R-C** factors was based on an analysis of the *language discourse* contained in a student quote. For example, when expressing data corresponding to the **R-M** factor, the student used specific verbs such as "seeing" and "looking"; adjectives such as "distinct" and "blob-like" and, nouns such as "triangle" and "Y-shape" to reason (**R**) about the graphical markings on the ER (**M**). In contrast, when expressing data corresponding to the **R-C** factor, the student linked specific words or reasoning phrases (**R**) such as "since" and "that means" to reason with specific concepts (**C**) such as "amino acid sequence" or "lock-and-key" for instance. To illustrate this methodology, consider the coding of the following extracts obtained from a student's interpretation of ER E:

> I: Ok, so where would a lock-and-key interaction happen here [S mentioned this earlier]?
> S: …well on both sides of the hapten. Because, if you see here, it would happen on this side and on this side [indicates with bottom hapten on trimer in ER E]. So, there'd be like two lock-and-key interactions on both sides… because the antibodies want to bind to the hapten, they're going to have to stretch out more, to bind to it [hapten]… these antibodies have a hinge region… like a door has a hinge… it has flexibility to stretch out more because of that hinge.

It is evident from the datum presented above that, in order to successfully interpret the scientific knowledge (**C-M**) contained in ER E, the student has to engage all factors of the model. For instance, in order to successfully interpret the triangular-shaped markings represented by the trimer (**R-M** and **M**), the student has to engage sound conceptual knowledge (**R-C** and **C**) surrounding the lock-and-key interaction between Ab and Ag. Here, the student correctly suggests that, "there'd be like two lock-and-key interactions on both sides" of the Ab molecule. Further evidence for the above student's engagement of her conceptual knowledge (**R-C** and **C**) upon interpretation of ER E is provided by her analogical reasoning used to suggest that, "… these antibodies have a hinge region… like a door has a hinge… it has flexibility to stretch out more because of that hinge."

At some time or other, a student is required to engage and integrate all factors of the model in order to successfully interpret an ER. By coding the engagement of factors **R-M** and **R-C** within student quotes, the data demonstrates the indispensable nature of each factor of the model for sound interpretation of an ER.

Potential uses and applications of the model

Based on the empirical data above used to validate the operational definitions of the factors constituting the model, the potential practical applications of the model should be given attention. In this regard, the following six potential practical applications of the model were derived from this research:

- The model can be used to establish whether a student's overall *interpretation* of an ER is successful or not. This can be done, by comparing the student's "post" knowledge after exposure to an ER with the conceptual knowledge represented by the ER (**C-M**).
- The model can be used to establish whether any *learning* has occurred as a result of a student's engagement with an ER. Here, the student's "post" knowledge (**C**) obtained after interpretation of an ER is compared with data corresponding to their prior knowledge (**C**) obtained before exposure to any ER.
- The model can be used to determine which of the six factors *positively or negatively* influence a student's interpretation of a particular ER the most and, which the least.
- The expressed model could serve as a general *diagnostic* framework for guiding practitioners' and researchers' discussion and data analysis relating to the nature of a student's difficulty with an ER. That is, whether the student has a conceptual (**C**) or reasoning (**R-M** or **R-C**) difficulty or, whether the difficulty lies with the nature of the graphical features of the ER (**M** and **C-M**). The model hereby enables the prediction of the *potential source(s)* of difficulties with ER interpretation.
- Based on the nature of the data corresponding to each factor, the model could serve as a template for the development of *approaches* to teaching and learning including intervention strategies for improving student's interpretation of and learning from ERs.
- Based on the nature of the model and the operational definitions of its constituent factors, the model has a generic application to *all types* of ERs in science including not only static representations but also dynamic, animated and multimedia representations.

Conclusions

The modelling process of Justi and Gilbert (2002) provided a rigorous platform from which to express a model of factors determining students' ability to interpret ERs in biochemistry. Empirical data corresponding to each of the seven factors constituting the expressed model were gathered with a specially designed clinical interviewing method, termed 3P-SIT (also see paper number 210971). Data generated from these interviews were analysed by a qualitative and iterative method. In so doing, each factor constituting the model was validated as an indispensable component that plays a role in determining students' ability to interpret ERs in a biochemistry context. As a result of this process, the authors generated specific operational definitions for the meaning and nature of each factor of the model as well as potential practical applications of the model.

It is clear from the proposed applications of the model that many of its uses require specialized knowledge and research expertise before teachers and learners will be able to benefit directly from them. In this regard, we are currently preparing further manuscripts that alleviate these issues (e.g. Schönborn and Anderson, in press). Here, based on the theoretical foundations of the model, we have concentrated on deriving user-friendly approaches that science educators and designers of visual material could use to prevent or correct students' difficulties with ERs used in biochemistry. Based on our progress thus far, we suggest that such remediative approaches based on this model will find application in most science education contexts. This is because, to successfully interpret, or learn from *any* ER (**M**) in science, a student is required to posses the necessary scientific conceptual knowledge of relevance to that ER (**C**) and, is required to possess the reasoning skills (**R**) necessary to reason not only with their conceptual knowledge but to also reason with that ER. Through this fundamental principle lies the power of the model for improving student learning and for making a contribution to the already existing literature on ER processing and interpretation.

Acknowledgments

The authors would like to acknowledge financial grants from the National Research Foundation (NRF) (GUN Number 2053218) for supporting this work and the University of KwaZulu-Natal Research Fund (URF) for a postdoctoral fellowship.

References

Ainsworth, S.E., Bibby, P.A. and Wood, D.J. (1998) Analysing the costs and benefits of multi-representational learning environments. In: M.W. van Someren, P. Reimann, H.P.A. Boshuizen and T. de Jong (Eds.), *Learning with Multiple Representations*, Amsterdam: Elsevier Science, pp. 120-134.

Ametller, J. and Pintó, R. (2002) Students' reading of innovative images of energy at secondary school level. *International Journal of Science Education*, **24**(3), 285-312.

Amit, A.G., Mariuzza, R.A., Phillips, S.E.V. and Poljak, R.J. (1986) Three-dimensional structure of an antigen-antibody complex at 2.8 Å resolution. *Science*, **233**, 747-753.

Baddeley, A. (1992) Working memory. *Science*, **255**, 556-559.

Blackwell, A.F. (2001) Introduction: thinking with diagrams. *Artificial Intelligence Review*, **15**(1-2), 1-3.

Cheng, C.-H., Lowe, R.K. and Scaife, M. (2001) Cognitive Science Approaches to Understanding Diagrammatic Representations. *Artificial intelligence Review*, **15**(1-2), 79-94.

Chi, M.T.H., Feltovich, P.J. and Glaser, R. (1981) Categorization and representation of physics problems by experts and novices. *Cognitive Science*, **5**, 121-152.

Gobert, J.D. and Clement, J.J. (1999) Effects of student-generated diagrams versus student-generated summaries on conceptual understanding of causal and dynamic knowledge in plate tectonics. *Journal of Research in Science Teaching*, **36**(1), 39-53.

Grayson, D.J. (2004) A teaching strategy for helping students disentangle related physics concepts. *American Journal of Physics*, **72**(8), 1126-1133.

Grayson, D.J., Anderson, T.R. and Crossley, L.G. (2001) A four-level framework for identifying and classifying students conceptual and reasoning difficulties. *International Journal of Science Education*, **23**(6), 611-622.

Jackson, J.G. (2000) School of Molecular and Cellular Biosciences, University of KwaZulu-Natal, *Personal Communication.*

Justi, R.S. and Gilbert, J.K. (2002) Modelling, teachers' views on the nature of modelling, and implications for the education of modellers. *International Journal of Science Education*, **24**(4), 369-387.

Lincoln, Y.S. and Guba, E.G. (1985) *Naturalistic inquiry*, Newbury Park: Sage, pp. 39, 202-208, 268-273, 340-344.

Lowe, R. (2004) Interrogation of a dynamic visualization during learning. *Learning and Instruction*, **14**(3), 257-274.

Lowe, R.K. (1993) Diagrammatic information: techniques for exploring its mental representation and processing. *Information Design Journal*, **7**(1), 3-17.

Mayer, R.E. (2003) The promise of multimedia learning: using the same instructional design methods across different media. *Learning and Instruction*, **13**(2), 125-139.

McCormick, B.H., DeFanti, T.A. and Brown, M.D. (1987) Visualization in scientific computing - a synopsis. *IEEE Computer Application Graphics*, **7**, 61-70.

McMillan, J.H. and Schumacher, S. (1993) *Research in education: a conceptual introduction, 3rd ed*., New York: HarperCollins.

Orgill, M. and Bodner, G.M. (2004) Locks and keys: how analogies are used and perceived in biochemistry classes. *Proceedings of the National Association for Research in Science Teaching (NARST)*, April 1-3, Vancouver, Electronic paper number 203366.

Ploetzner, R. and Lowe, R. (2004) Dynamic visualisations and learning. *Learning and Instruction*, **14**, 235-240.

Roitt, I.M. (1997) *Roitt's Essential immunology*, *9th ed*., Oxford: Blackwell Science.

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

Salomon, G. and Perkins, D.N. (1989) Rocky roads to transfer: rethinking mechanisms of a neglected phenomenon. *Educational Psychologist*, **24**(2), 113-142.

Schönborn, K.J. and Anderson, T.R. (In press) The Importance of Visual Literacy in the Education of Biochemists. *Biochemistry and Molecular Biology Education.*

Schönborn, K.J (2005) Using Student Difficulties to Identify and Model Factors Influencing the Ability to Interpret External Representations of IgG-Antigen Binding, *Ph.D. Thesis*, South Africa: University of KwaZulu-Natal (Pietermaritzburg).

Schönborn, K.J., Anderson, T.R. and Grayson, D.J. (2002) Student Difficulties with the Interpretation of a Textbook Diagram of Immunoglobulin G (IgG). *Biochemistry and Molecular Biology Education*, **30**(2), 93-97.

Schnotz, W. and Lowe, R. (2003) Introduction: external and internal representations in multimedia learning. *Learning and Instruction*, **13**(2), 117-123.

Sumfleth, E. and Telgenbüscher, L. (2001) Improving the use of instructional illustrations in learning chemistry, In: H. Behrendt *et al*. (Eds.), *Research in Science Education-Past, Present, and Future*, Kiel: IPN, pp. 289-294.

Valentine, R. C. and Green, N. M. (1967) Electron microscopy of an antibody-hapten complex. *Journal of Molecular Biology*, **27**, 615-617.

Wheeler, A.E. and Hill, D. (1990) Diagram-ease: why students misinterpret diagrams. *The Science Teacher*, **57**(5), 59-63.

White, R. and Gunstone, R. (1992) *Probing understanding*, London: Falmer Press, pp. 65-81, 82-95.

Winn, W. (1993) An account of how readers search from information in diagrams. *Contemporary Educational Psychology*, **18**, 162-185.

Wu, H.-K. and Krajcik, J. S. (2006) Inscriptional Practices in Two Inquiry-Based Classrooms: A Case Study of Seventh Graders' Use of Data Tables and Graphs. Journal *of Research in Science Teaching*, **43**(1), 63-95.