References

1. Meester M A M and Maskill R 1993 The practical side of chemistry *Educ Chem* 31 156-159
2. Bennett S W and O'Neal K 1998 Skills development and practical work in chemistry *U. ChemEd* 2 58-62
3. Clow D 1998 Teaching learning and computing *U. ChemEd 2* 21-29
4. Johnstone A H 1980 Chemical education research: facts findings and consequences *Chem Soc Rev* 9 365-390
5. Johnstone A H 1997 Chemistry Teaching: science or alchemy *J. ChemEd* 74 262-268
6. Johnston A H 1999 Chemed research: where from here? Proceedings *Variety in Chemistry Teaching* 14-16
7. Garratt J 1997 Virtual investigations: ways to accelerate experience *U. ChemEd* 1 19-27
8. Verdonk A 1993 The role of educational research in the quest for quality in the teaching and learning of chemistry Proceedings *Variety in Chemistry Teaching* (Ed) Aitken M (Royal Society of Chemistry, London)

Evaluation of teaching and learning: matching knowledge with confidence

COMMUNICATION

John Garratt,[a] Jane Tomlinson, [a] Simon Hardy,[b] and Doug Clow[c]

[a] *Department of Chemistry, University of York, York, YO10 5DD*
[b] *Department of Biology, University of York, York, YO10 5DD*
[c] *Institute of Educational Technology, The Open University, Milton Keynes, MK7 6AA*
e-mail cjg2@york.ac.uk

We have used a two-part questionnaire to obtain feedback from students immediately before, immediately after, and six weeks after carrying out a computer-based simulation. The simulation is intended to help students to develop investigative skills. The first part of the questionnaire tests knowledge by means of multi-choice questions. The second part asks students to assess their confidence in their understanding or in their ability to apply knowledge. The use of this evaluation strategy has allowed us to formulate hypotheses about ways to improve the student learning experience in future years. We conclude that this evaluation strategy can be a valuable and generally applicable way of identifying whether a particular learning experience helps students to develop an appropriate balance of knowledge, understanding and ability to apply knowledge.

Introduction

"Most British people, most educators and most students now believe that it is one of higher education's purposes to prepare students well for working life"[1].This sentiment is reflected in a number of recent reports which emphasise the need for degree courses generally[2] or specifically in chemistry[3,4] to adopt a more student-centred approach to teaching so that students develop a range of personal and professional skills appropriate to a scientific education.

Many individual teaching and learning strategies have been developed to bring about more active student participation in their education[e.g. 5–8]. One largely unsolved problem is evaluation of the effectiveness of such innovations. Their aim is rarely limited to that of helping students to achieve a higher mark in a conventional examination, and therefore it is not appropriate to evaluate them by attempting to measure a change in examination performance. Bodner et al have discussed the different reasons why this is inappropriate[9]. They argue that the main purpose of evaluation of any new teaching initiative is to discover what modifications to make which will maximise the positive effects and minimise the negative ones (since we should take for granted that any significant change will have some effect). We were faced with the problem of choosing an appropriate strategy for evaluating the success of introducing one of the eLABorate computer simulations[10] to a class of first-year biochemistry students. The particular simulation is enzymeLAB[11,12,13]. This simulation allows students to investigate the effect of substrate concentration, enzyme concentration and pH on the rate of an enzyme catalysed reaction. It is designed to build on (and hence consolidate) basic knowledge of enzyme kinetics, and to develop an understanding of how this knowledge is applied in the design of an investigation.

We were attracted by the strategy recommended by Draper for evaluating interventions in the classroom[14,15]. This involves two interesting features which we have not used previously. First he recommends the use of the same questionnaire as a pre-test, a post-test, and a delayed post-test (follow-up). Second he recommends that the questionnaire should be designed to make a dual evaluation by using *"a measure of the student's confidence in fulfilling the learning objectives, and a knowledge quiz."* We therefore decided to devise a two-part questionnaire. The first part was designed to be a knowledge quiz testing aspects of knowledge of enzymes which would be useful in planning a real investigation of an enzyme. The second part would seek information about their confidence in their understanding of concepts or in their ability to apply knowledge.

We report here the results of using our two-part questionnaire with a group of 23 first year biochemistry students. The number of students is too small for us to draw firm conclusions about the effectiveness of the simulation as a teaching aid. However, we believe that the results demonstrate that the evaluation strategy is a useful one, and could be used to advantage to evaluate a wide range of learning experiences.

Methods

The Simulation

The simulation deals with an enzyme which obeys Michaelis Menten Kinetics; that is, the rate of the enzyme catalysed reaction (v) is determined by the equation

$$v = \frac{V_{max}S}{K_m + S}$$

Where S is the concentration of substrate
K_M is a constant at constant pH (but may vary with pH)
V_{max} is a constant at constant pH and enzyme concentration (but may vary with pH and, at given pH, is proportional to enzyme concentration).

The key features are:

- the program selects the parameters (V_{max}, K_M, and their sensitivity to pH) so that each student is given a different enzyme;
- the user selects values for the pH and the concentration of enzyme and substrate at which the rate of the reaction is to be measured;
- the resulting value of v which is displayed has a realistic experimental error with a standard deviation of 5% of the correctly calculated value.

Aspects of the enzymeLAB program have been described elsewhere[8,11,12,13].

The Context

The first year cohort of Biochemistry students at York consists of about 25 students. This exercise forms part of a nine week module ‘Biochemical skills’. In this module, students perform a variety of tasks, usually working in teams. Examples include designing, carrying out, and interpreting biochemical investigations. The module is based on a workload of one day a week for eight weeks with an assumption of a small amount of additional private study.

The enzymeLAB exercise forms the practical component of one day of the course. The students are required to determine the value of K_M and V_{max} for their enzyme at pH 7, and to investigate whether the optimum pH of the enzyme is affected by the concentration of substrate. The precise wording of the task, as given in the student worksheet is as follows:

You have four tasks

(i) Show that the kinetics of your enzyme are consistent with Michaelis Menten Kinetics.
(ii) Find the value for V_{max} and K_M at pH 7.
(iii) Find the optimum pH of your enzyme.
(iv) Find out whether the optimum pH varies with the concentration of substrate.

First year students have little or no experience with this kind of problem which essentially involves recalling factual knowledge and applying it to a problem which cannot be solved by applying a fixed algorithm. The class therefore starts with a discussion designed both to bring key knowledge to the forefront of their minds and to help them to develop their own strategy for approaching the problem. The discussion lasts for 60 – 90 minutes. Students are then given access to the computers. They work on their own with one or two tutors available to answer questions and to provide support and guidance as necessary. About 45 – 60 minutes before the end of the session, the students are called together to discuss their findings.

Evaluation

For reasons outlined in the introduction, we aimed to evaluate the effectiveness of the exercise by a two part questionnaire. The same questionnaire was completed by the students before the start of the introductory session, after the end of the final discussion, and (without warning) in the final session of the course which is about six weeks later.

In 1998-99 we used six multiple choice questions, listed in Table 1 to test key aspects of the students’ knowledge of enzymes. Each question was provided with a correct answer, three distracters, and ‘don’t know’. In the interests of space, neither the correct answer nor the distracters are shown in Table1. The questionnaire was completed anonymously, and we emphasised that it was not a test, so that students should answer “don’t know” rather than guess at the correct answer.

The second part of the questionnaire, designed to establish the students’ confidence in their understanding and in their ability to apply knowledge, is shown in Table 2.

No time limit was set for the completion of the questionnaire, but students were encouraged not to spend time puzzling about their answers. All students completed both parts of the questionnaire in less than 10 minutes.

Results

The results obtained for 1998-99 from the multi-choice questions are summarised in Table 1. We have not distinguished between the different incorrect responses, though we recognise that further information about misconceptions might in principle be obtained from an analysis of the frequency of different incorrect responses. In this case, we decided that the number of students was too small to justify analysis at this level of detail. In general the completion of the exercise resulted in an increase in the number of correct responses, but for all questions except 5 it reverted somewhat towards the pre-exercise level after six weeks. The major difference between the pre-exercise and the follow-up responses is the smaller number of ‘don’t knows’ in the latter.

The responses to the second section of the questionnaire related to student confidence are shown in Table 2. As with the first part of the questionnaire, student confidence in their

Table 1: Student responses to questions in Section 1.

Number of students giving correct, don't know or incorrect responses in the pre-exercise test, the post-exercise test, and the follow-up test (six weeks later).

Note that all 23 students in the class completed the pre-exercise questionnaire; three were given permission to leave early and did not complete the post-exercise questionnaire; two students were absent from the final session when the follow-up questionnaire was completed.

		Number of student responses		
		Pre	**Post**	**Follow-up**
1. Which of these is the Michaelis-Menten equation, used in enzyme kinetics?	Right	13	20	15
	Don't know	2	0	0
	Wrong	8	0	6
2. Between what ranges of values would you expect the K_M of most enzymes to lie?	Right	6	16	14
	Don't know	12	0	1
	Wrong	5	4	6
3. Which one of the following *best* describes the conditions required for v to be near to V_{max}?	Right	13	16	11
	Don't know	5	0	1
	Wrong	5	4	9
4. Which one of the following is the best range of values of [S] to use to calculate K_M and V_{max}?	Right	2	6	4
	Don't know	10	2	1
	Wrong	11	12	16
5. Which one of the following *best* describes what effect changing the pH from the optimum would have on an enzyme's kinetic constants?	Right	7	11	16
	Don't know	7	3	2
	Wrong	9	6	3
6. Which one of the following is the most appropriate statement about the optimum pH of an enzyme?	Right	2	9	4
	Don't know	8	1	5
	Wrong	13	10	12
Overall Percentage	Right	31%	65%	51%
	Don't know	32%	5%	8%
	Wrong	37%	30%	41%

Table 2: Student responses to questions in Section 2.

Responses were awarded a score of 1 – 5 according to whether they responded 'no confidence', 'little confidence', 'some confidence', 'confident', 'very confident'. The number shown is the mean of these scores.

How confident are you that you ….	Average Score		
	Pre	**Post**	**Follow-up**
1. know what the Michaelis-Menten equation is, what it means, and how to use it?	3.17	4.30	3.29
2. understand what K_M and V_{max} are, what they mean, and the effect they have on the rates of enzyme-catalysed reactions?	3.44	4.10	3.43
3. have a good feel for the ways in which pH can affect K_M and V_{max}?	2.30	3.44	2.57
4. understand what it means for an enzyme to be saturated?	4.09	4.05	4.14
5. could plan a series of experiments to determine the K_M and V_{max} of an enzyme at a given pH?	3.35	3.75	3.33

understanding and in their ability to use knowledge shows an immediate increase as a result of completing the exercise, but appears to revert to the pre-exercise level after 6 weeks.

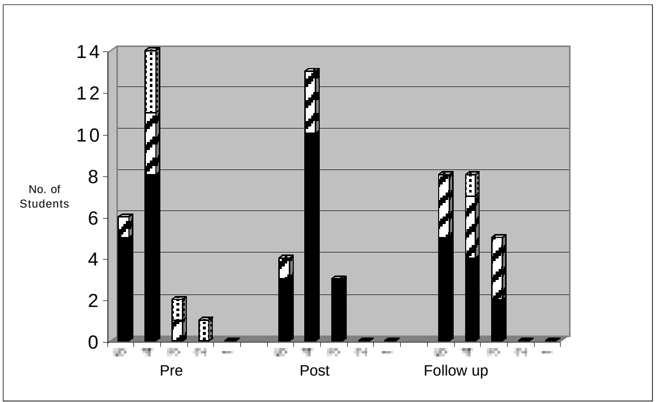
We examined the match between student knowledge (correct answers to questions in Section 1) and student confidence (Section 2). We illustrate our approach to this with questions 4 and 5 in Section 2. Question 4 asks "How confident are you that you understand what it means for an enzyme to be saturated?" In our view, for this confidence to be justified, the student must know that the observed rate of the enzyme-catalysed reaction (v) approaches the maximum rate (V_{max}) when the concentration of substrate is several times greater than the enzyme's K_M (which is the subject of Question 3 in Section 1 of the questionnaire).

Question 5 (Section 2) asks "How confident are you that you can plan a series of experiments to determine the K_M and

V_{max} of an enzyme at given pH?" This planning requires the selection of values of substrate concentration at which to measure the rate of the enzyme catalysed reaction; it is important to select values of substrate concentration which range on both sides of K_M and have a reasonable spread (say between 0.2 and $2K_M$). This knowledge is tested by question 4 in Section 1 of the questionnaire.

We have prepared histograms to show the data obtained from the questionnaires for these two confidence questions. These are shown as Figures 1 and 2. There are separate histograms for the data obtained pre-exercise, post-exercise and at the follow-up stage. The histograms show, for each level of confidence (very to none) the number of students responding correctly, incorrectly or don't know to the relevant question in Section 1. These histograms show that student knowledge (or lack of it) does not correlate well with their confidence in their understanding or their ability to apply their knowledge.

Figure 1 Student confidence in understanding of enzyme saturation.

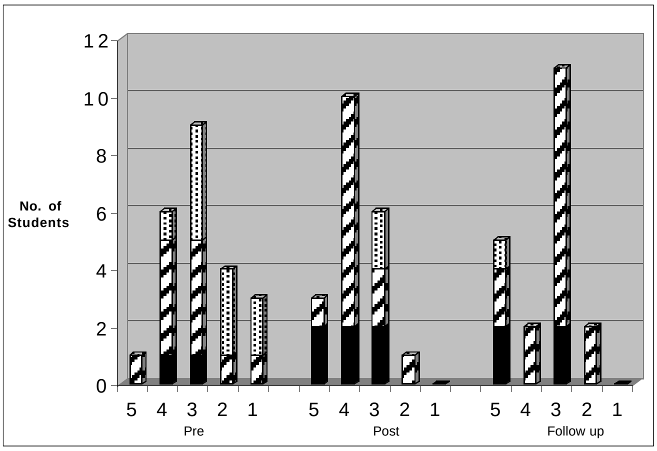


Histograms show numbers of students responding 'very confident' (5), 'confident' (4), 'some confidence' (3), 'little confidence' (2) and 'no confidence' (1) to Question 4 in Figure 2. These are correlated with answers to Question 3 in Figure 1:

■ correct; ▨ don't know; ▦ incorrect.

Data are shown for the Pre-exercise (PRE), Post-exercise (POST), and FOLLOW-UP responses.

Figure 2 Student confidence in ability to plan experiments to determine K_M and V_{MAX}



Histograms show equivalent data to Figure 1 for question 5 in Table 2 correlated with answers to Question 4 in Table 1.

Discussion

Our primary purpose in writing this paper is to present a critical retrospective analysis of the design of our questionnaire, in the belief that this will encourage others both to adopt the general approach, and also to avoid making some mistakes. This discussion is therefore in two parts. First, we make some interpretations of the data we obtained from the questionnaire. Secondly, we evaluate the questionnaire critically with a view to identifying specific ways of improving this particular questionnaire and also general features which we believe are important in other questionnaires of this kind.

Interpretation of Data

This analysis is designed to illustrate the sort of conclusions that can be drawn from the two-part questionnaires used pre-post-, and as a follow-up to the simulation exercise. Table 1 shows encouraging improvement in the number of correct answers to Section 1 of the questionnaire immediately after completing the exercise. Six weeks later it is encouraging that the proportion of correct answers is substantially higher than at the pre-exercise stage. We are not particularly surprised that there has been some reversion to the pre-exercise level six weeks later; we believe it would be unreasonable to expect a single exercise of this kind to reconfigure the long-term memory in such a way as to make this detailed factual information immediately available.

What is particularly notable is that the percentage of 'don't know' responses falls dramatically in the post-exercise questionnaire and stays low in the follow-up questionnaire. Questions 2 and 4 elicited the highest number of don't know responses in the pre-exercise questionnaire. This is not surprising since it is unlikely that answers to these questions would have been emphasised in first year lectures on enzyme kinetics. In the follow-up questionnaire, students rarely answered 'don't know' to either question. Unfortunately, for question 4, this resulted mainly in students giving incorrect answers! Question 6 is of interest in that it provided half of the 'don't know' answers in the follow-up questionnaire, even though we had hoped that the exercise would help the students to understand the effect of pH on enzymes. A possible conclusion is that, because different enzymes respond differently to pH change, students could not explore the possible range of effects by studying a single enzyme, and the class discussion did not bring this out sufficiently.

In contrast to the apparently improved knowledge-base demonstrated in table 1, the students' confidence appears to rise immediately after the exercise but reverts almost exactly to the pre-exercise level six weeks later. Bailey[16] has noted a fall in students' confidence in their skills after participating in exercises designed to improve these skills[17]. He attributes this to the students making a more realistic assessment of their skills when they understand better what is involved. It may be that a similar phenomenon is occurring here.

For reasons given in the next section, we limit our detailed discussion of the questions of confidence to questions 4 and 5, data from which are shown in Figures 1 and 2. Question 4 (Figure 1) is concerned with student *understanding*. In contrast, question 5 (Figure 2) deals with confidence in *ability to apply knowledge*. Figure 1 shows a high level of confidence in understanding of the meaning of enzyme saturation. This high level of confidence appears to be largely justified by the student knowledge of the conditions which would lead to saturation. The situation illustrated by Figure 2 is less satisfactory: here the confidence levels (though lower than those shown in Figure 1) are still reasonable, but this is not well-founded as is shown by their ignorance of one of the key facts which would allow them to design an investigation to determine K_M and V_{max}. If this truly reflects the student attitude it may show that they are not good at recognising the factors that need to be taken into account in designing an investigation. This may be evidence of a need for more opportunities to play a more active role in the design of experiments.

The Questionnaire

All the questions in Table 1 (if rewritten in free-response format) would have to be considered by anyone planning an investigation of a hitherto uncharacterised enzyme. Well-designed investigations are therefore likely to be planned by investigators who either can provide good answers to these questions or recognise their ignorance so that they can look up background theory before they start.

The questionnaires were completed anonymously. A disadvantage of this is that the potential to trace the development of an individual student's understanding is lost because we are not able to assign each questionnaire to a specific student. It is worth considering whether anonymity is sufficiently important to the student to require that it be maintained.

In evaluating responses to questions 1-3 of section 2 it became clear that the precise wording of the confidence questions is especially important: we cannot be sure how students would interpret questions which ask what is 'meant by' the Michaelis Menten equation, K_M or V_{max}, or what they would understand by 'having a good feel' for the effect of pH. Our rule for setting these confidence questions in future is that they should be of two types. One type would be worded "how confident are you that you *understand* (some well defined concept)", and a necessary condition for including it would be that we could write down a clear statement which we could accept as demonstrating understanding of the concept as described in the question. The second type would be worded "how confident are you that you can (carry out some task involving application of knowledge)", and a necessary condition for including it would be that we could write down a clear protocol for carrying out the task together with basic knowledge on which the protocol is based. Our suggested written statements on the confidence questions would serve two purposes. They would ensure that the questions were worded in a way which could be usefully interpreted by students. They would also emphasise the knowledge which should underpin the understanding and the application of knowledge. This would provide effective guidance on appropriate test questions (and relevant distractor answers) for Section 1.

Our view is that only questions 4 and 5 in Section 2 of our questionnaire meet these criteria, and that questions 3 and 4 in Section 1 are suitable complementary questions. It is perfectly possible to use more than one question in Section 1 to test knowledge on which confidence questions in Section 2 should be based. However this would lengthen the questionnaire. A key feature of the questionnaire is that the students should not recognise that the questions in the two sections of the questionnaire are intentionally correlated; this helps to give a clearer impression of any mismatch between their confidence and the knowledge on which this should properly be based.

This evaluation strategy has suggested to us ways of modifying the exercise to improve the student learning experience. In particular, our data indicates the importance of helping the students to recognise the logic behind the design of the investigation which they carry out. Similarly, the responses to Question 3 in Section 2 suggest that it would be worth emphasising the opportunity to gain more understanding of the effect of pH on enzyme activity. Thus we have at least partially achieved our primary objective of evaluating the exercise with a view to improving the student learning experience in future years. The combination of knowledge-based questions and questions of confidence has provided more useful information than we would have obtained by leaving out either part.

Our experience illustrates some ways in which carefully constructed questions related to self-assessed confidence and objective knowledge can together be an effective way of collecting feedback. We conclude that this evaluation strategy can be a powerful tool for testing the effectiveness of innovations in teaching, providing that the questionnaire is designed using the principles we have described. We recognise that the same strategy could be used in a more general context, without the need for repeated exposure to the same questionnaire, to evaluate whether students have a balance of knowledge, understanding and ability to apply knowledge.

Acknowledgements

The enzymeLAB software package is one of the simulations created by the eLABorate project which was funded by the University Funding Councils of England, Scotland, Wales and Northern Ireland under the Teaching and Learning Technology Programme. Further information on the eLABorate products is available at

http://www.york.ac.uk/depts/chem/staff/elaborate/

One of us (JLT) was supported by the ESRC under grant number RO22 25 0121.

References

1. *Towards a partnership* 1987 (The Council for Industry and Higher Education, London)

2. Dearing R 1997 *Higher education in the learning society* Report of the National Committee of Enquiry into Higher Education, HMSO London)
3. Mason G 1998 *Change and Diversity: the challenges facing chemistry higher education*. (Royal Society of Chemistry, London)
4. *General guidelines for the academic review of Bachelors Honours Degree Courses in Chemistry* 1998, (Quality Assurance Agency (Gloucester).
5. Garratt C J and Overton T 1996 Scientific papers as a teaching aid *Educ. Chem*. 33 127-138
6. Kee T P and Ryder J 1998 Developing critical and communication skills in undergraduates through chemistry *U. Chem. Ed.* 1 1-4
7. Belt S T and Phipps L E 1988 Using Case Studies to develop key skills in chemists *U. Chem. Ed* 2 16-20
8. Clow D J and Garratt C J 1998 Computer simulations: creating opportunities for science writing *U. Chem. Ed* 2 51-54
9. Bodner G MacIsaac D and White S 1999 Action research: overcoming the sports mentality approach to assessment/ evaluation *U. Chem. Ed* 3 31-36
10. Garratt C J Clow D Hodgson A Tomlinson J 1999 Computer simulation and chemical education: a review of Project eLABorate *Chem. Ed. Rev.* 14 51-81
11. Garratt C J and Groves P 1992 Computer simulations of experiments with enzymes *CHED Newsletter* Fall 1992 137
12. Clow D J M 1996 Computer simulations of laboratory experiments DPhil Thesis University of York
13. Garratt C J 1997 Virtual investigations: ways to accelerate experience *U. Chem. Ed* 1 19-27
14. Draper S W 1996 Integrative evaluation: An emerging study of classroom studies of CAL *Computers and Education* 26 17-32
15. Draper S An introduction to integrative evaluation *Proceedings of Variety in Chemistry Teaching 1996*: (eds. J Garratt and T Overton) (Royal Society of Chemistry, London)
16. Bailey PD 1999 Personal communication
17. Bailey PD and Shinton SE *Communicating Chemistry* (Royal Society of Chemistry, London)