

The Chemistry Quiz, a Tool for Reinforcement Learning

PAPER

Stephen M Walker

Department of Chemistry, The Donnan Laboratories, University of Liverpool, Crown Street, Liverpool L69 7ZD

Email: sk01@liv.ac.uk

The Chemistry Quiz (The Quiz) comprises a series of programs used to improve students' competence with the everyday numeric manipulations required in chemistry and is designed to complement the existing range of software available. The Quiz allows the generation of random input data so that every student receives a different value even when attempting the same question. The Quiz is designed only for this type of question and makes no attempt to incorporate textual input, multiple choice, or other features found in other products. All computations use integer arithmetic to avoid rounding errors, with sensible use of significant figures. Quizzes have been produced covering volumetric calculations, spectroscopy, thermochemistry, particles and waves, and algebraic manipulation.

The Quiz has been used by staff and students for three complete academic years and, in this same period of time, it has been used to test students' ability to carry out volumetric calculations. The results suggest that an increase in reinforcement learning does improve the ability to perform these simple calculations as evidenced by higher pass rates and improved scores. In addition, there is a considerable saving in staff time compared with more traditional testing scenarios.

Introduction

Chemistry Departments have been struggling with the reluctance of students to carry out even simple mathematical exercises for some time now. Physical chemists have traditionally borne the brunt of this problem but the decline in mathematical ability has become so marked that it is now affecting all aspects of the subject. In seeking to find pragmatic remedies it was pointed out that 'learning reinforcement'¹ might be used to good effect.

Feedback and reinforcement are two of the most pivotal concepts in learning. Feedback involves providing learners with information about their responses whereas reinforcement affects the tendency to make a specific response again. Feedback can be positive, negative or neutral; reinforcement is either positive (increases the response) or negative (decreases the response). Feedback is almost always considered external while reinforcement can be external (extrinsic) or intrinsic (ie. generated by the individual). Information processing theories tend to emphasise the importance of feedback to learning since knowledge of results is necessary to correct mistakes and develop new plans. On the other hand, behavioural theories focus on the role of reinforcement in motivating the individual

to behave in certain ways. One of the critical variables in both cases is the length of time between the response and the feedback or reinforcement. In general, the more immediate the feedback or reinforcement, the more learning is facilitated. The nature of the feedback or reinforcement provided was the basis for many early instructional principles, especially in the context of programmed instruction¹. For example, the use of 'prompting' (ie. providing hints) was recommended in order to 'shape' (ie. selectively reinforce) the correct responses. Other principles concerned the choice of an appropriate 'step size' (ie. how much information to present at once) and how often feedback or reinforcement should be provided. These principles are often used in CAL drill and practice software.

Schools have employed this technique successfully in the past in extrinsic mode, but its use has declined in recent years and universities usually expect students to assume this responsibility for themselves (intrinsic). However, it seems that many students are insufficiently motivated to persevere for long enough for effective learning to result. The computer provides a medium with which to create an environment suitable for this type of learning **and** the motivation to practice effectively.

With careful design, a computer program which can deliver numerical problems for the student to practice their mathematical skills would also be able to set and mark a test designed to assess those skills. This was particularly relevant in Liverpool where staff in the first year laboratory, being concerned at students' inability to cope with the routine calculations associated with volumetric analysis, had initiated a test that all students were required to pass, even if it meant taking the test several times. This imposed a substantial workload on the staff involved, with the need for setting and marking multiple tests for classes of about 100. A number of packages have been produced with the aim of helping students with their mathematical skills². Despite the excellence of some of them, none of those available at the time (1995) provided sufficient numbers of exercises to ensure that each student would be given a unique collection of problems to solve. Because we regarded this as an essential characteristic we decided to create software for ourselves. (During the course of this project, other programs became available which do not suffer from this limitation.³)

Thus the objective was to create a series of programs, now described as the Chemistry Quiz (The Quiz) which could be used by students for learning reinforcement (to practice their mathematical skills) but which could also be used by tutors to administer an acceptable test (which, in the case of volumetric calculations, would replace an existing test).

Methodology

The requirements of the program were defined as follows,

- it must have a practice mode and an examination mode;
- it must be able to generate sufficient questions so that students can both practice and be tested without there being a significant risk that the test and the practice examples would be identical;
- it must be easy to use (by staff and students);
- it must be secure;
- it must provide automatic feedback or marking.

Five topics were selected to cover most of the material that students would be expected to encounter in their first year course. Within each topic, a number of different types of question were defined. The topics and the number of question types defined for each are listed in Table 1. A good description of the kind of question would be 'chemarithmetic'. The Quiz is designed only for this type of numeric question and makes no attempt to incorporate textual input, multiple choice or other features found in other products. All computations use integer arithmetic to avoid rounding errors, with sensible use of significant figures.

The twelve types of question defined for the topic 'volumetric analysis' are listed in the appendix. They illustrate both the style of question and also the method of generating a huge number of different specific questions from each original type. This depends essentially on two strategies:

- choosing for each question type an appropriate range of numerical values, and generating a random number within this range for each specific question;
- where appropriate, selecting at random one reagent from a set of plausible ones.

Modes

In order to satisfy the needs for learning reinforcement and for testing, The Quiz may be set up for student use in one of three different modes. The tutor controls which modes are available to the student through the part of the program "Quizmaker" (see section on Program Design).

1. Practice mode

This is the default mode. In this mode the student is presented with a total of 20 questions chosen at random from all the available question types for a particular quiz. The total of 20 questions has been chosen to minimise time and resources and to provide a reasonable length of study. In this mode a student may opt to exit at any time or, indeed, to repeat the quiz an

unlimited number of times. Students may save and review their work in order to gauge their progress and for this purpose the quiz considers the students' answers to be 'correct' if they are within 5% of the true answer. The student also has the ability to choose a particular question type at any time, which is useful to staff when setting tutorials and homework.

2. Directed practice mode

In this mode the tutor controls both the available question types and their relative frequency. This is useful if some of the question types are not covered in a particular syllabus, if the class requires extensive practice in a particular method, or if the quiz needs to be modified during a course as material is progressively put in front of students. Students are unaware of the differences between modes 1) and 2) - they are both perceived as practice mode. However, in the directed practice mode students cannot opt to choose a particular question type.

3. Examination mode

In this mode, the number of questions, their relative frequency and order are completely specified by the tutor. In complete contrast to modes 1) and 2), students cannot exit until they have seen all the questions stipulated (although not necessarily answered!) and when they finish, their answers are recorded in 'csv' format for compatibility with spreadsheets. Students have the facility to review their answers to all questions and to change the values should they so wish. In this mode students do not have the capability of saving the answers in their own file records.

Program design

The Quiz comprises four parts.

- *Quizmaker* is installed on the tutor's own computer and is never accessible to students. It is used to control which modes are available to students and to create tests using a file setup.dat.
- *Quizzes* creates individual questions from the basic question types.
- *Rich Text Writer* is used by the tutor to create Help files customised to the course.
- *Analyser* works the answers.

Quizmaker

The Quizmaker program consists of a number of pages. Quizmaker is not used if The Quiz is to be used only in the default practice mode.

1. Mode and student registry.

This page is used only to set up a test or exam. If directed practice is chosen then the program will immediately jump to the next page. If examination is chosen, this page allows the tutor to define the type of information required of a student taking the test - first name, family name and password. All of these are optional and the password field entered by the student is encrypted for security. This page also allows the tutor to set a time limit for the test.

Table 1: The five quizzes and the number of questions within a topic.

Topic	Question Types
Algebraic Manipulation	12
Thermochemistry	13
Particles and Waves	11
Elementary Molecular Spectroscopy	9
Volumetric Analysis	12

2. Quiz choice

The second page asks for the name of the quiz to be used. A list box is presented with the names of all the quizzes currently installed and a button is activated which allows access to the details of the particular quiz selected so that the lecturer can decide which type of questions should be included in the quiz.

The next box on this page asks if all of the questions should be included and the last box asks if the question order should be randomised. In directed practice mode the responses to these two questions are ignored since the program assumes, by definition, that only a sub-set of questions will be used and random order is obligatory.

3. Answers

The third page only appears in examination mode. It asks for the location of the students' answer file and its name. The second question asks for the number of questions to be included, up to a maximum of 20. If a time limit is not specified then there are no further questions. Alternatively, a final question box will appear inviting the time limit to be specified in minutes.

4. Questions

This page presents a grid for specifying the Question Types and their frequency. Initially the question types are assigned in ascending order to fill the total number of questions desired, with a wrap over if there are more questions than types. At this point the exact order and type of question may be specified. If random order has been specified then this grid is used to determine the frequency of occurrence of each type - the program will randomise the order for each student. In directed practice mode, randomisation is obligatory.

5. Final page

The last page summarises the choices and allows the user either to re-start or to save the data. The file called 'setup.dat' is then automatically created, saved and copied into the same directory as the particular quiz chosen.

Quizzes

When a quiz file is opened, the program looks for the file 'setup.dat' in its own directory. If it fails to find it then it will proceed in practice mode. If the set-up file is found and practice mode has been specified, the student will perceive the program as operating in practice mode. If the examination mode has been specified then the student has the choice of continuing to practice or opting to sit the examination.

Having opted to sit the examination the student must answer the quiz in the style defined by the lecturer. Any question can be skipped by clicking on 'Next Question'. Once all the required questions have been presented, the student has the option of reviewing the answers and changing them if desired. This can be repeated as often as required until the student is satisfied. At this point the student should click on "Exit Exercise" and the answers will be appended to the data file.

If a time limit has been specified then the quiz begins with a reminder to the student that this is in operation. The timing does not start until this message has been dismissed by the

student. The time remaining is displayed below the question number in minutes (to 1/10 minute). When 5 minutes are left this display changes colour to red and with 1 minute to go a message is placed on the screen. If the time runs out before the student saves the answers then the time remaining message changes to a flashing Time Expired notice and after 5 seconds the data are recorded and the program shuts down.

Rich text writer

Help with the calculations is written using the program Rich Text Writer by the tutor in the institution running the program. This is called 'Calculation Help' and gives advice on the calculation method for a particular question in terms appropriate to the lecture course. *Rich Text Writer* is designed to assist with this and generates the type of text needed by chemists. If a tutor decides that help is not required, then the program will simply report 'Sorry, no help available'.

Other help on how to run the software is built into the program.

Analyses

A separate marking program has been developed which completely automates the marking process and will, if required, generate a final '.csv' file for use in spreadsheets. This program is optional since the answer format is compatible with all spreadsheets and a custom template can be constructed using any preferred program. *Analyzer* allows the accuracy levels required and the mark awarded, to be set for each question. The major difference between computer marking and paper marking is that in the former it is an 'all or nothing' process with full marks for a correct answer and zero for one outside the prescribed accuracy limits. A human marker will probably give partial marks - for example, for writing a correctly balanced equation. Computer marking is thus likely to give apparently lower marks unless one realises that two different types of question are being asked and either adjust the marking scheme accordingly or accept that two different tests exist.

Examination implementation

Advice on the setting of computer-based tests is readily available⁴ The volumetric quiz was set up for examination with the following specifications,

- 1 Number of questions: 6
- 2 Question types (see Appendix): 1,3,5,7,10,11
- 3 Order of questions: strictly as set down in 2.
- 4 Time limit: 45 minutes.
- 5 No passwords required.

The appearance of the quiz is shown in Figure 1.

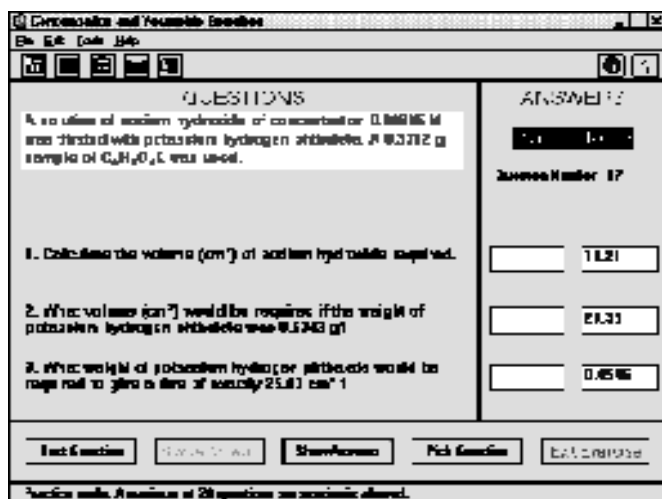
The students were given a 'window' of two weeks in which to take the test and acquire sufficient competency via reinforcement learning. They were able to take the test at any time within this period but the majority opted to do this on the final two days. There was no attempt to check for collaboration, impersonation or cheating and no form of traditional exam supervision was maintained. The examination window was scheduled for a period when all students had had time to learn the computer's operation and become confident in the mechanics.

Table 2: Examination results in volumetric analysis for 1995-98.

Method	Year and Test	Candidates	Performance	Mean Mark
Paper	1995 First test	81	27 Fail (1 withdrew)	52%
	1995 First resit	26	16 Fail	50%
	1995 Second resit	16	2 Fail (withdrew)	55%
Computer	1996 First test	101	11 Fail (1 withdrew)	72%
	1996 First resit	10	0 Fail	65%
	1997 First test	91	7 Fail (6 withdrew)	73%
	1997 First resit	1	0 Fail	
	1998 First test	89	7 Fail	73%
	1998 First resit	7	2 Fail (withdrew)	65%

Figure 1: The screen as seen by the student in practice mode with the computed answers displayed.

The students have access at all times to a calculator (with the ability to copy and paste into the answer box) and a table of appropriate relative molecular masses (so as not to introduce errors from this calculation)



Results

The Quiz program on volumetric analysis was introduced in all three modes in 1996. Prior to this all students sat a written examination of exactly the same form as that specified for the computer test. The written test was designed to be taken in a lecture slot of 45 minutes, whereas the computer test could be taken at any time within a 2 week period. Students failing the written or computer test are required to take resit tests until they either pass or withdraw from the course. The pass mark was set at 40%.

The results of the test and the resits for 1995 – 1998 are shown in Table 2. 1995 was the last year in which a written test was taken. The results are typical of previous years.

The key features of these results are that from 1996

- the mean mark shows a substantial increase;
- the number of failures shows a decrease;
- only one resit was necessary.

Discussion

The very clear implication of the results is that the major reason for the improvement is practice. In the allowed fortnight, students spent the first week practising and only attempted the test when they felt confident. The end result, whatever the means, is that the class can now carry out their volumetric calculations with confidence. This may simply reflect training in a specific type of calculation. Critical evaluation of the same students' performance in tutorials associated with the remedial mathematics course provides no evidence of improvement in their basic mathematical skills. It appears that improvement in one aspect of chemarithmic is not necessarily transferable to other aspects. However, even if this exercise has only improved one aspect of student numeracy, this must be of significant value. At best, it suggests that a similar strategy could lead to a more general improvement in mathematical skills.

The other four quizzes on spectroscopy, thermochemistry, particles and waves, and algebraic manipulation, have been used only in the practice modes (sometimes in conjunction with tutorials). They are popular with students, many of whom have purchased their own copies for use on their own computers. There is no information on how much they are actually used. Furthermore, in the absence of specific performance tests before and after the introduction of the quizzes, it is not possible to assess whether students show an improvement in ability to deal with these areas of chemarithmic. It would be interesting to know how much the impending test increases the students' motivation to take the opportunity for the 'learning reinforcement' offered by the programs.

As far as staff are concerned, the computer-based volumetric analysis test has been a significant advantage. The time taken to prepare, invigilate, and mark an examination is now 10 minutes compared to approximately 20 hours for the series of paper-based tests. Of course, it ignores the time required to write the programs and staff time devoted to running the computer network. However, the development time for the programs can (in theory) be spread over their total use; if more institutions adopt the programs, the development time becomes more effective.

Students are generally supportive, although it is difficult to extract anything sensible when asking their opinion of examinations. They like to practice, they like the two-week window and they like to be trusted. The only genuine negative note concerns their inability to indicate their thought processes and to add textual comment to their answers (a bonus as far as staff are concerned!). It may surprise some that they made no attempt to work together (there was, they said, insufficient time for chat), nor to personate, nor did they make any attempt to cheat in this completely unsupervised examination. Since 1997 the computer classroom has been fitted with a remote surveillance camera for security reasons and spot checks confirmed that serious individual work was carried out.

There is no evidence for the 'under-marking' traditionally expected when using a computer based test since the marking scheme for the paper test concentrated on numeric accuracy alone.

Appendix 1: The Volumetric Quiz

The volumetric quiz contains the following 12 question types,

• Type 1. Concentration-1

A random weight (from 0.2 to 2 g) of a typical volumetric reagent, chosen at random from a list of 10, is dissolved in water (from 50 cm³ to 1000 cm³) with no details of how the solution is prepared (e.g. acid required etc.).

One question is asked: find the molarity of the reagent.

• Type 2. Concentration-2

Similar to Type 1. Two questions are asked,

Question 1. Find the molarity of the reagent.

Question 2. Calculate the concentration of one of the elements in the reagent.

• Type 3. Concentration-3

Similar to Type 1. Three questions are asked,

Question 1. Find the molarity of the species.

Question 2. Calculate the volume of solution required producing a molarity of exactly 0.10000, 0.01000, 0.00100 etc. as appropriate.

Question 3. As question 2, but with a randomly chosen molarity.

• Type 4. Acid/Base-1

Randomly chooses either the standardisation of sodium hydroxide by potassium hydrogen phthalate or of hydrochloric acid by potassium hydrogen carbonate. Unknown molarities are in the range 0.08 to 0.13 with the unknown solution being titrated against random weights of the standard substance.

One question is asked. Find the molarity of the unknown reagent.

• Type 5. Acid/Base-2

Similar to Type 4. Three questions are asked,

Question 1. Calculate the molarity of the unknown solution.

Question 2. Repeats the calculation for a different weight of standard.

Question 3. Asks for the weight of standard required to produce a titre of exactly 25 cm³.

Acknowledgement

I would like to thank Dr H. Aspinall of this department for devising the questions used in the volumetric test and agreeing to submit her class to the experiment.

References

- 1 Deterline W A, 1962 An Introduction to Programmed Instruction, (Prentice-Hall, New York)
- 2 Active Learning Issue 1 *Computer Assisted Assessment*, CTISS, Oxford (1994)
- 3 Mackenzie D, Tripartite Interactive Assessment Delivery System (information can be obtained from Earth Sciences, University of Derby).
- 4 Head S, Chapman K and Heath S, Protocol for the implementation of summative computer-assisted assessment (CAA) examinations, University of Aberdeen (1997)

• Type 6. Permanganate-1

The standardisation of unknown permanganate solutions by sodium oxalate. Molarities are in the range 0.018 to 0.024. One question is asked. Find the molarity of the permanganate.

• Type 7. Permanganate-2

As Type 6. Three questions are asked,

Question 1. Find the molarity of the permanganate.

Question 2. Repeats the calculation for a different weight of oxalate.

Question 3. Asks for the weight of oxalate required to produce a titre of exactly 25 cm³.

• Type 8. Dichromate

The standardization of potassium dichromate with either iron(II) sulfate or iron(II) ammonium sulfate. Molarities are in the range 0.014 to 0.019.

One question is asked. Find the molarity of the dichromate.

• Type 9. Thiosulfate

The standardisation of sodium thiosulfate with potassium iodate using starch indicator. Molarities are in the range 0.08 to 0.13.

One question is asked. Find the molarity of the thiosulfate.

• Type 10. Chromium Ore

A back titration example. The chromium is oxidised to dichromate and reacted with an excess of iron(II) sulfate (25 cm³ of ca. 0.4 M). The remaining iron is titrated with potassium permanganate.

One question is asked. Find the weight percentage of chromium (in the range 8 to 13%).

• Type 11. Calcium Ore

The calcium (25 to 40%) is precipitated as oxalate, dissolved in acid and titrated with permanganate.

One question is asked. Find the weight percentage of calcium.

• Type 12. Argentimetric

Titration of a halide (randomly chloride, bromide, or iodide) against silver nitrate using dichlorofluorescein. Molarities are in the range 0.08 to 0.13.

One question is asked. Calculate the percentage of halogen in the unknown.