Improving Students' Data Analysis Skills in the Laboratory

PAPER

Paul C Yates

Department of Chemistry, University of Keele, Keele, Staffordshire ST5 5BG
e-mail: cha18@keele.ac.uk

Basic courses in mathematics for chemists often do not allow students to practice analysing data within a true laboratory environment. This is a vital skill which students often find quite intimidating. This problem has been addressed by using the first session of the practical component of a physical chemistry module to perform a data analysis exercise with the assistance of staff and postgraduate demonstrators. Student reaction suggests that such an exercise is needed, and that more encouragement needs to be given in the use of technology to support laboratory data analysis.

Introduction

The emphasis within the physical chemistry laboratory is very much on producing quantitative results and assessing their reliability. It is well known that many students of chemistry currently have great difficulties with the mathematical aspects of the subject. In a previous study[1] I concluded that students in the physical chemistry laboratory experience difficulties as a result of their weak mathematical ability. Some of the principal difficulties experienced by students are:

- general lack of confidence in using numbers
- failure to appreciate orders of magnitude
- lack of ability in basic algebra
- failure to appreciate the importance of including units.

These problems arise despite the provision of basic mathematics courses including such topics as estimating the maximum probable error and elementary statistics[2].

These difficulties are exacerbated by the students' inexperience with technique which means that their experimental error is greater than it would be for an experienced scientist. Thus the whole problem of data quality and uncertainty analysis becomes an important issue, and this is an area which requires a degree of experience not often possessed by the average first year chemistry undergraduate. The introduction of pre-laboratory work may help to overcome these problems[3].

Much pre-laboratory work has concentrated on providing a knowledge-base which will help students to understand underlying theory behind a laboratory experiment[4,5]. A different approach was needed to overcome the problems which students experience with data analysis. It occurred to me that a single pre-laboratory experience could prepare students for a complete laboratory course. This approach has been tried before with some success[6]; in the study described here all students analysed the same set of data while their activities were carefully monitored by staff and demonstrators, with ample opportunity for discussions.

The main difference between such a session and the analysis of data generated from a real experiment is that the instructor retains more control over the quality of the data and thus identifies mistakes more quickly. Furthermore all problems associated with experimental technique are eliminated, thus allowing students to concentrate completely on data analysis. The tutor does not need to spend a large amount of time trying to identify an error in a single student's piece of work and so is free to have meaningful discussions with a large number of students. This enhances the learning of each student, and should reduce the time spent in looking for mistakes later in the course when students do move on to analyse their own data.

The introduction of modularisation required the revision of the second semester course on "Energy and Dynamics". This provided an opportunity to introduce an exercise which would allow students to go through a specimen (typical) calculation with support from a tutor in the expectation that this would develop the confidence and skill to enable them to do other calculations unassisted.

Data analysis exercise

The data analysis exercise was designed to form a part of the first session of a second semester course consisting of twelve 3-hour laboratory sessions. Typical student numbers in this laboratory are 35-40, with three postgraduate demonstrators and one member of academic staff in attendance. The other activities included use of the Chemistry Courseware Consortium software on Using Tables and Graphs[7] and the straight line[8], and viewing of the video showing the experimental arrangement[9]. Thus 60 – 90 min was available for the data analysis exercise.

Data for analysis was chosen using the following principles:

- It should be provided by an experiment which was conceptually simple enough for the students to be able to envisage how the data had been collected without actually doing the experiment;
- A demonstration (or good video presentation) of the procedure should be available;
- The experiment should not be included later in the course;
- In order to fit the available time the data manipulation should not involve calculus, but only the simple algebraic manipulations of addition, subtraction, multiplication, division, indices and logarithms.
- The data processing should involve the use of a straight line graph, since this is one of the most common features of data analysis in physical chemistry.

Using these principles, the experiment selected was one in which the equilibrium constant for the dissociation of dinitrogen tetroxide is measured at various temperatures in order to determine the standard enthalpy change for this reaction.

Prior to modularisation, this experiment had been used in the second year physical chemistry laboratory. It was abandoned because it is technically difficult to carry out so that most students obtained poor data. Since constraints of time limited the collection of data to three temperatures, students had difficulty analysing straight line graphs. Although the technique is difficult, the theory required for this experiment is covered in the first year physical chemistry module. Furthermore the reaction itself involves a clearly visible colour change which is shown on a video in the Basic Laboratory experiments series[9].

The actual data used was modified from that provided by Alberty and Silbey[10]. Their data is essentially error-free, and reasonable error was introduced in order to encourage students to explore ideas of uncertainty analysis.

Description of the experiment

The laboratory experiment involves filling a calibrated flask with nitrogen dioxide to reach atmospheric pressure, and weighing the amount of nitrogen dioxide (which is, of course, a mixture of NO_2 and N_2O_4). The measurement is made at three different temperatures. The variation in mass reflects the degree of dissociation of the nitrogen dioxide. The calculation of the equilibrium constant for the reaction and for the molecular mass of N_2O_4 (M_1) uses the value for standard atmospheric pressure (p^o) and experimental values as follows:

- the volume of the flask (V) which is calibrated by weighing it empty and full of water, and using the density of water to determine the volume;
- the mass of nitrogen dioxide (w) from the weight of the flask when evacuated and filled with nitrogen dioxide;
- atmospheric pressure (p);
- the temperature at which the flask is filled with nitrogen dioxide.

The calculation itself involves the following steps:

- calculation of the apparent molecular mass of the gas in the flask from equation 1;

$$M = \frac{w}{V}\frac{RT}{p} \qquad (1)$$

- calculation of the degree of dissociation (α) from equation 2

$$\alpha = \frac{M_1 - M}{M} \qquad (2)$$

The equilibrium constant is then obtained from equation 3.

$$K = \left(\frac{4\alpha^2}{1-\alpha^2}\right)\left(\frac{p}{p^o}\right) \qquad (3)$$

The calculation of $\Delta H^{\ominus}$ for the reaction involves determining the slope of a graph of lnK against 1/T.

Running the data analysis exercise

This exercise has now been run for two years. The data analysis exercise itself has not been changed, but in the second year, the arrangements for running it were modified: the class was split into groups each under the supervision of a specific demonstrator and the data analysis was undertaken in rotation, with the other activities described above. In both years peer group discussion was encouraged as long as it was relevant to the work.

Students were given a verbal outline of the exercise and a set of instructions for analysing the data. They were also asked to estimate an uncertainty on their final value of ($\Delta H^{\ominus}$. Although methods of uncertainty analysis are covered in an earlier Mathematics for Chemists module, they had not previously been expected to apply these methods in a real situation. The format allows extensive discussions with the students, so that it is possible to emphasise that the level of certainty cannot be known, but only estimated, and that several methods are available for determining this[11]. The value of discussing the uncertainty estimation for each experiment with a member of staff or a demonstrator is stressed. The whole aspect of uncertainty analysis is something which students have traditionally found difficult in physical chemistry, and this approach allows some indication of the need for further discussions at the appropriate time. No formal treatment of this subject is made in the first year, but it is covered in more depth subsequently. The aim here is to give students an appreciation of the importance of error analysis and to consider how it may be applied in each of the experiments met.

No guidance is given as to how students should produce the required graph, and this is another area where informal discussion can be beneficial. With only three sets of data, statistical analysis of the line of best fit is inappropriate and students have invariably drawn the graph by eye. This easily leads to a discussion of how the line was chosen, and how this reflects the error or uncertainty in the result. The choice of appropriate scale is another feature which leads to useful discussion.

Some students completed the exercise relatively quickly while others struggled to finish in the time available. The format used allows flexibility by allowing the instructor to help, for example by supplying selected missing values in order to allow a student to move to the next step. In this way, each student can be shown the whole process, and given the opportunity to generate the required graph and discuss it.

Feedback given to students

All students are required to complete the exercise to the approval of the demonstrator. This judgement is based on an evaluation of how well the student has addressed the data analysis, and whether they are judged competent at data manipulation. It is thus possible for an incomplete exercise to be judged satisfactory, which goes some way towards addressing the different speeds with which students work.

Once the data analysis exercise has been completed to the satisfaction of the staff member, the student is provided with a sample report for the experiment which includes a full analysis of the data.

The sample report is approximately 500 words plus the relevant graph and includes sections headed *Introduction*, *Experimental Method*, *Results*, *Calculations*, and *Conclusion*. It is intended to parallel the form of the report required for experiments which are actually performed in the laboratory during the rest of the module.

Evaluation of effectiveness

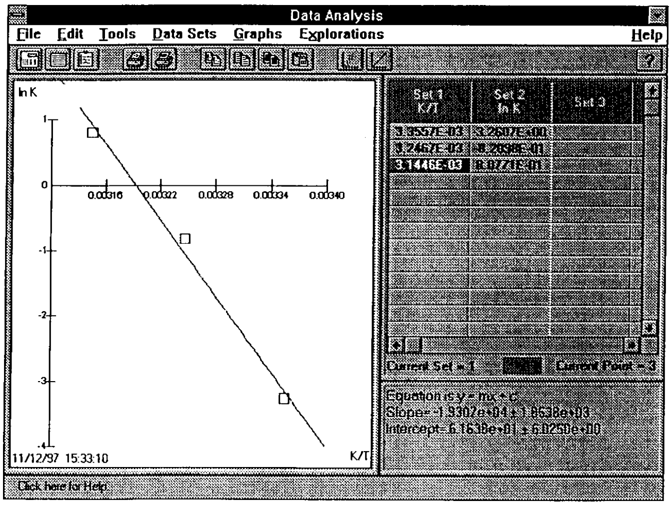
In the first year of operation the data analysis exercise ran with one large group while subsequently the class was divided as described earlier. There are merits to each approach; in the former less time is lost due to moving around, but in the latter additional supporting activities could be introduced, and students were under the supervision of an allotted demonstrator.

Observation of the students during the exercise showed that many found it difficult even though no new mathematical techniques were being introduced, and full instructions for performing the calculations were provided. The format of the exercise encouraged the students to interact constructively with the demonstrators, and the quality of the questions they raised, and their response to guidance gave a strong indication that this was an effective learning experience.

This view was supported by postgraduate demonstrators during the reflective critical review at the end-of-module review.

Student opinion was canvassed by informal discussions with the students throughout the module. These confirmed that they viewed the exercise as a useful preparation for the analysis of their own data and allowed them to tackle this task with confidence. Compared with previous years the submission of experimental reports with correctly completed calculations was considerably improved.

Figure 1: Result of using the data analysis tool in this experiment



Although I conclude that this exercise was of considerable benefit, I can see that improvements could be included in future years. One point we noted was that very few students even considered the possibility of plotting the graph using a computer. This suggests that this technique could usefully be incorporated into the exercise. There are many packages available for performing straight line plots and least squares fitting, including the Data Analysis Tool in the "Using Tables and Graphs" part of 'The Chemistry Tutor'[12] (see Figure 1).

The value of data analysis exercise might also be enhanced if combined with an appropriate post-lab after the experimental work as described by Johnstone et al [3]. At present some of this time is used for student oral presentations, and there is an opportunity to incorporate such work into the final timetabled session.

Availability

Copies of the data analysis exercise and the sample report described are available from the author on request.

References

1. Yates PC 1998 Making the best use of the first year physical chemistry laboratory *Research in Assessment* in press
2. Yates PC 1997 *Chemical Calculations* (Blackie Academic and Professional, London)
3. Johnstone AH, Sleet RJ and Vianna JF 1994 An information processing model of learning: its application to an undergraduate laboratory course in chemistry *Studies in Higher Ed.* **19** 77-87.
4. Kolodny NH and Bayly R 1983 Enhancing the laboratory experience in introductory chemistry with Apple-based pre-lab quizzes *J. Chem. Ed.* **60** 896-897
5. Starkey R and Kieper D 1983 A computer-based pre-laboratory quiz for organic-chemistry *J. Chem. Ed.* **60** 897-898
6. Craig DW 1994 A concept modeling and data-analysis pre-laboratory workshop for introductory chemistry *Abstracts of Papers of the American Chemical Society* **208** 194
7. Using tables and graphs 1997 *The Chemistry Tutor* (Chemistry Courseware Consortium, Liverpool)
8. Line: Y = mX + c 1997 *The Chemistry Tutor* (Chemistry Courseware Consortium, Liverpool)
9. Basic laboratory chemistry experiments, programme 17, Educational Media Film and Video Ltd
10. Alberty RA and Silbey RJ 1997 *Physical Chemistry Second Edition* (Wiley, New York)
11. Gordon R, Pickering M and Bisson D 1984 Uncertainty analysis by the 'worst case' method *J.Chem. Ed.* **61** 780-781
12. Runtime files and tools 1997 *The Chemistry Tutor* (Chemistry Courseware Consortium, Liverpool)