Proceedings

Laboratory work provides only one of many skills needed by the experimental scientist.•

John Garratt
Department of Chemistry, University of York, Heslington, York, YO1 5DD
e-mail: cjg2@york.ac.uk

Introduction

I graduated as a Biochemist in 1959. Of course a lot has changed since then. One change that ought to have persuaded us to re-evaluate the way we teach science in general and chemistry in particular is the enormously increased need for a scientifically literate population – a population capable of understanding, discussing, and influencing those major issues of the day that are based on science. It seems to me to be absolutely right for a meeting about chemistry and education to think about the issues and questions that are important for our society. Of course many of these have only tenuous connections with science; issues to do with war, with terrorism, with refugees and such like. But lots of today's serious moral and economic issues are a consequence of advances in science; genetic engineering, cloning, and climate change come to mind immediately. General issues like these generate more specific questions; some relate to risk (is it safe to immunise children with MMR vaccine? to eat meat from cows with BSE? to continue to burn fossil fuels without constraint?); some relate to control measures (is a vaccination policy an effective method of controlling an outbreak of foot and mouth disease?); some relate to the availability of resources (what sort of resources is it reasonable to commit to a particular genetically deformed baby? to providing impotent men with Viagra? to keeping unhealthy pensioners like me alive?). Each of us will have our own examples of questions desperately needing an answer based on an understanding of the scientific process and of the nature of evidence. But in addition to having a scientific basis, any answer has to be applied in the context of a complex society with its own history, culture and conventions. It really is extremely important for everyone, including those not trained in science, to understand and debate these issues.

To meet this need for a scientifically literate population we need to review our responsibilities as teachers. When I graduated I thought that academics were supposed to educate an elite to **extend** our knowledge of the World. Now I know it is more than that. About ten years ago, at a meeting to debate 'should higher education address business needs?' I said that University graduates should "*know their subject, and be able to **ex**plain, **ex**ploit and **ex**tend their knowledge*". My choice of 'explain' was deliberate; it emphasises our role as evangelists of public understanding of science. Today I would amplify what I then called "*the triple X Experience*" by saying that our role as teachers is to educate scientifically literate evangelists. A truly scientifically literate evangelist will recognise that 'Laboratory work provides only one of many skills needed by the experimental scientist'.

Of course I agree that laboratory work is a defining feature of a natural science, though not of course exclusive to chemistry. That doesn't mean that I think we should therefore describe chemistry as a laboratory-based subject, since I don't believe this does justice to what we actually do. I suggest that a better description of science (including chemistry, of course) is:
'a discipline which is based on the logical and imaginative interpretation of purposeful observation'.

Making Purposeful Observations

I chose 'purposeful observations' carefully to distinguish them from what I call 'chance observations'. I call something a chance observation when it is commonplace enough to be made, but not noticed, by other scientists. Pasteur and Fleming famously made chance observations; they were such commonplace observations that they were made by dozens of others (including to my certain knowledge the uncle of my first lab technician). What distinguished Pasteur and Fleming was that they converted their chance observation into a purposeful one by imaginative and logical interpretation.

Of course there are other examples in history, but they get rarer because the kind of chance observations that can be turned into purposeful ones (even by the most creative thinkers) have mostly been made. New observations that contribute to our understanding of the world are hard to make. Nowadays purposeful observations (even if they are unexpected) are made under very special and unusual conditions. This gives us a clue about the way scientists work. I suggest that our work involves the following six steps:

i) decide what observations we would like to make,
ii) imagine the conditions in which such an observation might be made,
iii) plan how best to create these conditions,
iv) create the conditions to the best of our ability (usually in a laboratory),
v) observe carefully to see whether our imagination and our planning were effective,

• This is an edited version of the Galen Lecture delivered at the 4th Variety in Irish Chemistry Teaching on 27.3.2002.

vi) interpret the observations with a mix of logic and imagination.

We often use the phrase 'doing and experiment' to describe steps (iv) and (v), and may overlook the fact that these two steps are a part of a larger (seamless) process. Most chemists (but not all scientists) 'do their experiments' in a laboratory. That is why laboratory work is central to chemistry. But we need to put it into context by also emphasising the creative thought that goes into planning the conditions in which a purposeful observation might be made and into the imaginative interpretation of observations so that they expand our knowledge of the world. Doing laboratory work is necessary for the advancement of scientific knowledge and understanding, but it is not sufficient. Laboratory work is also difficult and expensive, so it is wasteful to do experiments before we have thought as carefully as possible about how to make the observations we want. In other words, real scientists put off their experiment until they have thought it through; minimising the need for laboratory work is, I suggest, a sensible principle for scientists. Of course, it turns out that laboratory work is so slow, and the need for purposeful observations is so great, that experimental scientists spend a great deal of time in the laboratory.

Here is my list of the things we think about before doing an experiment, which subsumes steps (i), (ii), (iii), and (vi).

- What question(s) are we trying to answer (what idea(s) are we testing?)
- What observations (data) would provide an answer to the question(s) (would be consistent with or refute the hypothesis)?
- How can we best create conditions for making the desired observation(s) (collect the data)?
- How will we process and evaluate the observations (data)? Note that this includes taking account of error and uncertainty in any observation (measurement) made.
- What will we do next – why did we bother?

I stress that **all** these are things we think about **before** doing the experiment (and therefore often things that we do outside the laboratory, with a consequence that our students may not associate them with laboratory work). The 'what next' point is an important one. The scientist is like a chess player; always thinking several moves ahead, even though the result of the next step is uncertain; in other words we predict but not rely on the result of the previous step. Without this element of the planning process, an experiment is not real science but becomes mere 'stamp collecting'. I don't think this is an insult to philatelists since they know they are not engaged in a pursuit designed to lead to the discovery of the secrets of the world through "*the systematic study of nature*" (a phrase borrowed from the Canadian novelist Robertson Davies).

I contend that in our teaching we over-emphasise laboratory work at the expense of planning and interpretation, and consequently we devise laboratory exercises that encourage a 'stamp collecting' approach to science. The laboratory exercises we give to our students actually discourage them from thinking scientifically about the process of science in which I include

- the nature of evidence and proof,
- the design of investigations,
- the limitations to knowledge imposed by the available procedures for obtaining it.

Let me illustrate what I mean with an example from a lab manual I picked up recently when I happened to visit a friend. This is a good and well thought out exercise; it's worth including in any undergraduate chemistry course. One thing that really impressed me was the clear and honest list of objectives heading the instructions in the lab manual. Here they are.

- To gain experience of monitoring reaction progress using spectrophotometry;
- To learn about pseudo-first-order kinetics;
- To compare manual and automated methods for data acquisition and analysis;
- To determine an activation energy.

Clearly someone has thought carefully about what the student is supposed to learn. All of them are important. The lab manual also contained a recipe with details of

- the concentrations of reagents to use,
- the temperatures at which to measure the rate of reaction,
- the method to use to measure this rate,
- the way to process the data to obtain the required activation energy.

Because this exercise involves following a well-tried recipe designed to give a result that is already known, it cannot really be described as 'doing an experiment'; the students have no need to think at all about the scientific process. I have always maintained that these recipe-following exercises are a necessary part of the process of learning about experimental work, but they are not sufficient.[1] The exercise I have just described is (like most other lab exercises) an excellent example of how to 'collect a new stamp'.

What such exercises do not do is provide any opportunity for the student to learn how to make a purposeful observation. It does not help the student to learn

- why anyone wants to know the value of an activation energy (when you have measured it, what are you going to do with it? what makes this measurement part of science rather than just stamp collecting?);
- how to formulate an hypothesis ;
- how to design an experiment to test that hypothesis.

Understanding how to test an hypothesis is a particularly important part of scientific literacy.

In our book *A Question of Chemistry*[2] we provide one of my favourite examples of the muddled attitudes to the criteria of proof. It concerns the safety of Rabbit Calicivirus as a way to control the rabbit plague in Australia. The extract below was taken from an article in the New Scientist.

"…to justify releasing the virus in the first place, the Australian government ***should have first obtained clear proof that it infects just one species****, the rabbit. Researchers claim to have done just that. They exposed 31 species of native and domestic animals to the virus. They measured the amount of antibodies and virus in the blood and organs of these animals, and looked for signs of sickness. Those tests showed that the virus did not replicate or cause disease in any test animal.* ***'Our testing of rabbit calicivirus is the most comprehensive study that we know of into the host range of an animal virus',*** *says Murray."*

In the book we ask students to consider whether the criteria specified for releasing the virus can be achieved, and whether Murray's statement is consistent with the supposed need for clear proof.

This passage illustrates that the nature of proof is often poorly presented by scientists, and so it is unsurprising that the public misunderstand the limits to scientific enquiry. Leading from this, I propose two key principles that scientific evangelists need to impress on the public.

- It is theoretically and philosophically impossible to prove the absence of something (an effect, a substance, etc).
- The ability to detect something positive depends on the precision of the method in use (the level of random error) as much as it does on its sensitivity.

I have plenty of anecdotal evidence to support the suggestion that neither point is intuitively obvious. I even have real evidence from a small study we carried out into what we called 'the language of error'.[3] We found that most of a sample of first year chemistry undergraduates believed that 'a qualitative method can be used to **prove** that a constituent is absent from a substance'. Few of them recognised that that the limits of detection of the analytical method merely set the upper limit of the amount of substance that can be detected. This, and other misconceptions we uncovered, hinder their ability to design convincing investigations. I will illustrate this point with some previously unpublished data.

Investigating Factors Affecting the Time of a Pendulum Swing

In this study, first year science students used a computer simulation called pendulumLAB (created by Jane Tomlinson). This allows the investigation of factors affecting the length of time a pendulum takes to swing. The user can choose the length of the pendulum, the mass of the bob, and the angle to which the bob is raised, and is asked to investigate the effect each of these has on the time of the swing. Before discussing the results from our volunteer students, I will suggest how the investigation might be carried out by someone who adopts the principle of minimising laboratory work.

Such a person might draw up a strategy along the following lines.

- First establish which of the variables has a detectable effect.
- Set up the pendulum and make replicate measurements with all variables constant to determine the precision of measurement.
- Change one of the variables (a lot) and make a similar set of replicate measurements.
- Now change another (a lot) and repeat.
- Now change the third variable (a lot), and repeat.
- If necessary investigate further with more measurements

Table 1. Data to show the effect of angle, mass and length on the time of a pendulum swing

Data to show the effect of		Angle	Mass	Length
Length	150	150	150	20
Mass	20	20	100	100
Angle	80	20	20	20
Readings	27.7	24.5	24.8	9.2
	27.7	24.8	24.9	9.0
	27.9	24.8	24.8	9.3
	27.1	24.5	24.5	9.3
	27.3	25.1	24.3	8.9
Mean	27.54	24.94	24.66	9.14
S.D.	0.29	0.24	0.22	0.16

- Big changes in each variable would be achieved with angles of 80 and 20, mass of 20 and 100 g and length of 150 and 20 cm (fairly arbitrary values chosen by common sense and for convenience).

Table 1 shows data obtained by using these principles. These illustrate that, on the basis of these 20 measurements, we can easily draw the conclusions that

- There is a clear effect of angle
- Any effect of mass is too small to be shown with this set of data
- There is a huge effect of length.

Of course it is now quite easy to add more data. One might do this to try to establish the shape of the relationship between angle and time or length and time, or to test whether an effect of mass could be detected by collecting more data at the same values of angle and length, or (preferably) with a larger range of values for the mass of the bob.

The data in Table 2 illustrates how difficult it can be to detect an effect if one makes the wrong choice of variables. Columns 4 and 5 show this with different angles, where the two angles chosen are close together, and columns 1-3 show the similar problem that arises if the length of the pendulum is so short that the time of the swing is difficult to measure with precision. This illustrates that careful thought makes the difference between a successful investigation from which unambiguous conclusions can be drawn and the collection of data that may be misleading.

Table 2. Data to show the effect of angle when the length is short and the angle variation is small

Length	20	20	20	100	100
Mass	100	100	100	100	100
Angle	40	20	80	20	40
Readings	9.3	9.2	10.1	20.2	20.4
	9.2	9.0	10.1	20.0	20.5
	9.2	9.3	9.9	20.3	20.7
	9.4	9.3	10.0	20.4	20.5
	8.9	8.9	9.9	20.1	20.6
Mean	9.20	9.14	10.0	20.2	20.5
S.D.	0.19	0.16	0.1	0.16	0.11

Investigations by Students

Turning to the data collected by students, an encouraging fact is that all of them carried out 'fair tests' in the sense

that they studied the effect of one variable at a time, whilst keeping the other two constant. However, only four of them had a clear policy of making replicate measurements. One of these took five replicate readings, and did so at each of the fifteen chosen of conditions; another took twelve sets of triplicate readings which made up about a third of the total; the other two took a single set of replicates (one of eight and one of ten). That left ten students who had essentially no information about the reproducibility of their data. I say 'essentially no information' because the data collected by nine of them did actually include one or two sets of duplicate or triplicate readings, but these look as though they were obtained by chance rather than by design since the separate values were obtained whilst systematically changing different variables, and it is doubtful whether the students noticed.

As part of the study, we had asked the students to predict the effect of each variable before starting their investigation, since this might help them to focus on what they were testing. Their predictions are summarised in Table 3. Note that any prediction that there will be no effect can **in principle** be refuted by demonstrating an effect. In contrast, any prediction that there will be an effect cannot **in principle** be refuted. The latter prediction can be confirmed by observing an effect. In contrast, failure to observe an effect may simply mean that the method in use is not sufficiently sensitive or precise to detect one. There is little to say about the effect of length. Only two students made the incorrect prediction that it would have no effect, and both changed their minds as a result of their investigation.

As we have seen, the effect of the angle is much smaller. Even so, one might have expected that the eight students who correctly predicted an effect would confirm their expectations; in fact only half did so. One (who measured the time of swing at 17 angles from 2-60 degrees, and at a length of 10 cm) concluded that the effect of angle is 'variable, increasing and decreasing the time taken'. Since this student made no replicate measurements, it is not possible to say whether he gave any consideration to the effect of error on the data. The other three students actually changed their minds and concluded that angle has **no** effect. None of them based this on what could be described as an exhaustive study; one took 9 readings, another took 10 and the third took 24 (including 10 replicates with all parameters constant). I believe that these students failed to detect an effect simply because they made a poor choice of conditions – all of them used a pendulum length of only 10 cm, which makes the effect hard to detect, and one of them made the task almost impossible by restricting the range of angles to 30-60 degrees. It seems that they changed their minds without good justification, and in a direction that is **in principle** dangerous, and they showed a fine disregard for the principle that the absence of an effect cannot be proved.

The other six students made a prediction that, as Table 1 shows, can be refuted quite easily. However, only three of them changed their minds. A fourth thought that there is probably an effect, but was not confident of the significance of this in the absence of statistical analysis. This conclusion was based on a two-minute study involving only 19 measurements (including a single replicate). Had the student made (say) 6 replicates at each of three different angles (18 readings instead of 19), the data would have convincingly demonstrated an effect without the need for statistical analysis. That leaves two students of these six who confirmed to their satisfaction that there is no effect of angle. Perhaps these two were so committed to their prediction that they failed to test it adequately; this is a practice that we may deplore, but which we are all too aware happens.

Student conclusions about the effect of the Mass of the bob lead us to similar conclusions. The one student, who predicted there would be no effect, confirmed this prediction on the basis of single measurements made at each of four masses ranging only from 10 g to 40 g. All thirteen students who made the incorrect prediction that Mass would have an effect, changed their mind. One actually concluded that the time of swing decreases as Mass increases – the opposite of the prediction. This was a remarkable conclusion to draw from 8 measurements each made at a

Table 3 Student predictions for the effect of variables on the time of swing for a pendulum

VARIABLE	LENGTH	ANGLE	MASS
NO. PREDICTING:			
NO EFFECT	2	6	1
EFFECT	12	8	13

different Mass and giving a range of values for the time of 14.1 – 14.8 s, since to test the reproducibility of the method, this student earlier made 10 replicate measurements, the results of which varied from 6.1 to 6.9 s. The remaining eleven all concluded that Mass has no effect, which is in accordance with the current state of physical knowledge. This would be a satisfying result, if the students had reached their conclusion after an exhaustive investigation. Unfortunately this was not the case, as is obvious from the simple observation that few of them had sufficient results to estimate the level of uncertainty (or error) in their data. Table 4 shows the number of results obtained by each of the students that give information about the effect of Mass. Only two of them could possibly be described as having carried out an exhaustive investigation. Far from being comforting that they drew a correct conclusion, this result can only mean that these students (like many of the others) have a very poor appreciation of the nature of evidence and, in particular, of the philosophical impossibility of proving the absence of an effect. Furthermore, the students generally seem to have little concept of the most efficient way of testing whether an effect can be observed, or of the problems created by experimental error.

Table 4 Data related to the effect of Mass on the time taken by one swing

Number of OBSERVATIONS	No. of students	RANGE of masses	No. of students	LENGTH used	No. of students
> 25	2	90	8	≥ 100	5
8 - 13	9	50 -70	2	20 - 50	4
3 or 4	3	30 - 40	4	10	5

I do not wish to give the impression that I think the students have demonstrated incompetence, since the outcome was very similar when we asked academic scientists to carry out the investigation.[4] One may excuse both groups on the grounds that they were put in a position that discouraged thought, and strongly encouraged them to do an experiment while they still had things to think about. But I also think that we do too little to provide opportunities for students to develop the kind of thinking which might help them both to design better investigations themselves, and also to recognise flaws in the design of other investigations. A scientifically literate evangelist needs this latter skill as much as any other.

Possible Remedies for Shortcomings in Laboratory Investigations

This conveniently brings us back to the point that most laboratory exercises we give to our students actually discourage them from thinking scientifically about investigations. The example mentioned earlier illustrates this. The exercise requires the students to determine the activation energy of a reaction. The following list summarises key aspects of experimental work missing from this exercise.

- The exercise gives no clue about the reasons why a chemist might need to measure the activation energy of a reaction; what makes the measurement a purposeful observation and not a piece of stamp collecting?
- The students have no opportunity to think about how to choose the conditions under which to make measurements in order to ensure that suitable data are collected.
- The only guidance given on the treatment of experimental error is an instruction to estimate the largest and smallest values that could fit the data, thus students are not encouraged to consider why literature values of activation energies (and other measured values) rarely offer a range of values.
- The determination of a value for the activation energy does not provide an opportunity to test an hypothesis.

How can we remedy these shortcomings? I have three suggestions

- Incorporate additional tasks into the lab manual;
- Introduce (or enhance) pre-lab and post-lab work;
- Integrate computer simulations with a lab exercise.

When we published our study of student understanding of the language of error,[3] we suggested that teaching in this area needs to be radically rethought and restructured. We proposed that lab manuals for work involving quantitative measurements should include instructions such as 'give evidence of the random error in your data', 'indicate precautions you took to avoid systematic error', 'comment on the comparability of data collected by different individuals (and whether differences are significant)', and so on. Of course these tasks would be difficult for students since most of them are unfamiliar with the relevant concepts. One way to overcome this problem would be to provide all students with a 'Glossary of terms used to deal with error and uncertainty in experimental data'. This would be more or less equivalent to a Data Book that many departments provide for students to use at all times including in examinations. Such a Glossary would provide much more than mere definitions of words and phrases, and in many cases would provide a substantial paragraph explaining a term and showing how chemists (scientists) have adopted a specific technical meaning for a word that may have a rather different emphasis in common usage; the use of 'accuracy' and 'precision' provides an obvious example. A well-written Glossary of this sort could be a valuable asset for most graduate chemists, and the careful design of questions incorporated into lab scripts would encourage them to become familiar with it and perhaps to continue to use it for many years after graduation.

This general concept of including small additional tasks into the lab manual for incorporation into the lab write up could, with advantage, be extended to non-quantitative lab work such as synthesis. For example, students could be asked to comment on the purity of their product (what impurities are most likely to be present, what is the maximum level at which they might be present?), or to comment on yield (was their yield of product satisfactory for a method intended to provide enough material for further testing, or for commercial exploitation on a kilogram scale, or on a multi-tonne scale). I believe that imaginative questions would benefit from the existence of a well-written Concepts book, which would have similar benefits to those perceived for the Glossary.

I believe that this sort of additional task would help to focus students' attention on factors which experimentalists need to take into account when interpreting their observations, and in this way could bring a formal exercise one step closer to the making of a purposeful observation. But there is a limit to what we can do in this way. For example, these tasks cannot bring students any closer to realising that one of the key things to think about before resorting to laboratory work is 'what will we do next?' One way to do this is by studying a relevant published paper in a pre-lab session. This was an idea I developed with Brian Mattinson at York.[1, 5] I think we have not previously seen this as a piece of pre-lab work deliberately linked to a specific piece of lab work to enhance understanding of purposeful observations. One of the papers we used happened to deal with the measurement of rate constants (not a million miles away from my exemplar lab exercise of determining an activation energy). The authors of this paper wished to determine the rate constant for the reaction between OH and NO in order to better understand possible effects on the stratosphere of an increase in supersonic air travel (a matter of concern at the time the paper was written). Imagine using this paper as a pre-lab exercise before students tackle the determination of an activation energy. It would surely be easy to convince them that the determination of rate constants and activation energies is not a piece of stamp collecting but is a purposeful observation. They could appreciate that the investigation, which forms the subject of the paper, is too complex to be carried out in an undergraduate laboratory, but that it is worth practising their laboratory skills on a simpler system. Thus this paper exercise, if carried out in conjunction with a lab exercise, should help students to appreciate the place of laboratory work in the broader canvas of experimental science.

The third approach that I advocate is the use of computer simulations. There are many ways of using these to complement and enhance laboratory work. Here I want to limit my discussion to two of these. One is illustrated by enzymeLAB, the first simulation I planned and used. Its purpose was to provide students with an opportunity to plan their own investigation of an enzyme.[6] It would be easy to design an analogous simulation dealing with the determination of an activation energy. This could simply involve using the Arrhenius equation and assigning a value for A and E_a to an imaginary reaction ($A + B \rightarrow P$). With such a simulation, students could decide for themselves on the conditions under which to measure k_{obs} in order to determine the activation energy. The exercise could be carried out before or after the laboratory exercise to illustrate the point that careful thought about the design of an investigation can usefully lead to the minimisation of time and effort spent in the laboratory. However, my experience shows the importance of both providing careful preparation, and also sensitive feedback, if students are to get maximum benefit from the use of simulations like this.

A limitation of this way of using simulations is that it can emphasise the stamp collecting approach to science, since it does nothing to show the importance of **purposeful observations**. To deal with this aspect we need a simulation that requires some form of hypothesis testing, such as pendulumLAB. What makes this particularly suitable is that one of the three variables to study has a very clear effect on the measurement, one has an effect that can be tricky to detect, and the third has no effect. Nevertheless, pendulumLAB is not suitable for inclusion in an undergraduate chemistry course, because the topic does not link with the knowledge base of the subject. One possibility would be to adopt the principle of a simulation that we called unknownLAB in which the subject of study is not given, but the user is asked to investigate the effects of three variables. The abstract quality has the questionable virtue that it does not obviously relate to a different discipline. I would make two changes to our original version of unknownLAB before recommending it for trial. First, I would increase the flexibility by having different versions of the basic relationship between dependent and independent variables, and these could be randomly assigned to different students. Secondly, I would provide background notes indicating that previous observations on the system provided an indication of the likely findings. I would do this partly because it is more realistic (Galileo would have known quite a lot about pendulums before attempting to carry out a definitive investigation), and partly because it would encourage users to think carefully about how to test an hypothesis.

The other real system, which I would consider simulating, is the solubility of alcohols in water (or may be the partition of alcohols between oil and water). I like this system because most students (and a surprising number of academics) are unfamiliar with it; they usually think that the solubility of straight chain alcohols (say C_3 to C_6) in water increases as the temperature is raised, whereas, of course, the reverse is true. This trivial observation can be converted into a purposeful one by using it to calculate ΔH and ΔS for the transfer of a CH_2 group from water to oil. Now that many chemists are aspiring molecular biologists, this is an important measurement, since it is the basis of an understanding of hydrophobic interactions. It also happens to fit my criteria for a useful simulation, in that ΔH is very small and therefore difficult to determine, whereas ΔS is large and negative. Because the system is unfamiliar to most students of chemistry, it seems a suitable one to ask them to investigate; they would do so with some expectations of what result they would find, and many of them would be surprised by their findings. This would provide useful opportunities for discussion of the process of science that go far beyond the narrower field of hydrophobic interactions.

Concluding remarks

In this paper I have argued that our practical courses typically over-emphasise laboratory work at the expense of the planning of investigations and the interpretation of data. A consequence is that our students are not taught the true meaning of scientific literacy and frequently have only a poor appreciation of some basic principles of the nature of evidence and proof that contribute to the scientific method (whatever this is). I provide some evidence from previously unpublished results, which supports my concerns and conclude by making some suggestions for teaching strategies that I believe could lead to improvements.

Readers convinced of my first point may be encouraged to introduce changes to their teaching strategies, some of which may even be based on the suggestions I have outlined. Bodner et al.[7] argue that when we introduce changes it is because we have "*perceived weaknesses in the current situation*", that we have "*formulated an hypothesis that a particular change will lead to a particular improvement*" and that we will "*wish to test or evaluate our hypothesis*". Alas, those who have introduced imaginative changes in their teaching know how difficult it is to evaluate their success; a likely outcome is that the students neither enjoy nor appear to learn from the experience. Neither finding should persuade us that the idea should be dropped, but both should be matters for some concern.

I do not believe that student 'enjoyment' of a learning experience is a good measure of its potential value, even though it often appears high up on course evaluation forms. Adverse feedback from students need not be taken at its face value, and we should heed the advice of Bodner et al.[7] that our evaluation should "*look behind the façade of answers to the question 'do the students like it?' toward deeper questions such as 'what do students learn that they were not learning before?'*". However, I do believe that negative student feedback provides evidence that the teacher's intentions have not been fulfilled; it may not be the idea that is at fault so much as the detail of its execution. It may simply be that the students are unaware of what we are trying to achieve – that we are so convinced of the need for change that we have forgotten that the students have both different starting points and different objectives. We need to try to get into their minds and change our presentation accordingly, remembering the tenet of Constructivism that new knowledge needs to be constructed in the mind of the learner, and built into the individual's existing framework of knowledge.

The question of whether the students learn has been addressed somewhat enigmatically by Bodner in his comment that "*we can teach – and teach well – without having the students learn*".[8] When challenged about this, he explained that he was pointing out that the criteria used by unbiased onlookers to assess teaching quality do not usually include that of student learning. This point was brought home to me when the use of scientific papers as a teaching aid, which I outlined above, was picked out by the TQA exercise as of particular merit even though I have no evidence at all that the students learned anything from the experience. My conclusion from the negative nature of the student feedback from the scientific paper exercises makes a good general conclusion to this paper. It is that the piecemeal introduction of innovations made by individual enthusiasts is always likely to produce disappointing results because the impact is too small in relation to the course as a whole. If we are serious about the need for increased scientific literacy amongst our students, then this must be reflected in a change in attitude of the whole department; it is no use any one member thinking that it is an issue that can be left to one or two enthusiasts.

Acknowledgments
I am grateful to the organisers of the 4th Variety in Irish Chemistry Teaching for inviting me to give this lecture, and to Galen Pharmaceuticals for their generous sponsorship of the lecture. Amongst the many colleagues who have helped me to develop the ideas I present here, I am particularly indebted to Jane Tomlinson, who not only created pendulumLAB but has also been a constant source of inspiration to me since we started to collaborate.

References

1. J. Garratt, *U. Chem*. Ed., 1997, **1**, 19.
2. J. Garratt, T. Overton, and T. Threlfall, *A Question of Chemistry*, Longman, 1999.
3. J. Tomlinson, P.J. Dyson, and J. Garratt, *U. Chem. Ed.*, 2001, **5**, 16.
4. J. Garratt, and J. Tomlinson, *U. Chem. Ed*., 2001, **5**, 74.
5. J. Garratt and T. Overton, *Ed. in Chem*., 1996, **33**, 137.
6. J. Garratt, D. Clow, and P. Groves, *Biochem. Ed.*, 1994, **22**, 15.
7. G.M. Bodner, D. MacIsaac, and S. White, *U. Chem. Ed*., 1999, **3**, 31.
8. G.M. Bodner, *J. Chem. Ed*., 1992, **69**, 186.