Evidence and Interpretation: Teachers' Reflections on Reading Writing in an Introductory Science Course

Ezra Shahn and Robert K. Costello
Hunter College of The City University of New York

Abstract

The use of writing as a means of assisting students to learn and of assessing their understanding in an introductory science course intended primarily as a terminal course for non-science majors is considered in the context of a discussion of cognitive development. We suggest that, particularly where students are asked to justify their understanding by referring to concrete evidence, writing samples are a sensitive indicator of cognitive position. We demonstrate this with examples of four different types of writing used in our course: short answer exam questions, exam essays, take-home essays which may be revised, and informal journal writing. The information gained from writing assignments can be useful as feedback to an instructor regarding (a) an individual student's assumptions about what can be known in science and what form this knowledge takes, (b) what individuals and the class as a whole are prepared to understand, and (c) in what ways particular subject material is likely to be misunderstood. We conclude that these different probes can reveal different aspects of development, and that the use of any of them requires attentive reading by the instructor.

Originally published in *Academic.Writing: Supporting Interdisciplinary Perspectives on Communication Across the Curriculum*: Shahn, Ezra, and Costello, Robert K. (2000). Evidence and Interpretation: Teachers' Reflections on Reading Writing in an Introductory Science Course. *academic.writing*. http://aw.colostate.edu/articles/shahn_costello2000.htm

Publication Date: March 26, 2000

 DOI: 10.37514/LLD-J.2000.4.2.05

Introduction

While it is generally accepted in many circles that writing can be useful both to enhance and to assess learning (Kleinsasser *et al*., 1994), introducing writing as an integral part of college science courses remains an elusive goal. This is largely due to the fact that knowledge of science is traditionally thought to reside in such skills as identification of facts (memory) or quantitative problem solving (algorithmic thinking). Thus biology lab "practicals" may require the naming of organs identified by pins bearing numbers, multiple choice tests in several disciplines may involve selecting the correct names of processes and relationships hidden among distractors, and solving word problems can require students to use the appropriate knowledge to balance a chemical equation or find the range of a projectile. Where in these activities is there a place for writing? In this paper we briefly describe an introductory lab science course designed to incorporate writing, discuss the nature of several different writing exercises that we have used, and examine some examples of student writing as a means of demonstrating what may be expected from non-science majors. In fact, in most instances the "prompts" for the writing assignments have been constructed so as to emphasize specific cognitive activities. Thus, not only are we frequently looking to see how the students use evidence to justify their answers, but the writing samples themselves are the evidence that we are using as the basis for our interpretation of the students' cognitive positions.

Despite the activities included in traditional science courses, many teachers who have taken such courses acknowledge that they only *really* learned a subject when they had to teach it. If our goal is for students to learn science, then we must rethink our course requirements to include activities that will engage our students in the same sort of processes that we go through as we prepare our new courses. This does not mean that we have to make our students teachers in fact; we can, however, get them to approach information in a manner that somehow mimics what we do. Outlining is of course part of this process, but for what purpose? and in what context? Most science texts are highly structured, and simply rewriting the chapter and section headings is not what we have in mind. Rather, when we prepare a course, we think of what we will say about each major point. This being the case, it is reasonable that we find ways for our students to do likewise. Because students are less skilled and knowledgeable than we, they should not be required to say it (i.e., organize and present their thoughts orally), but they should commit their connected thoughts to paper. In this way writing can be brought into the science class as the appropriate way to encourage learning.

Foundations of Science

The course we discuss in this paper is *Foundations of Science* (Shahn, 1990). This is a one year course with three hours of lecture and three hours of lab each week. The course is introductory, and can be taken by freshmen. In fact, since it is primarily taken by non-science majors in partial fulfillment of a distribution requirement, it has students at all levels, but the instruction remains introductory. Lab sections are small – 15 to 20 students – and discussion is encouraged in them, covering lecture and reading material as well as lab activities. All sections meet for the same lecture, typically about 75 to 100 students. With regard to content, the course is multidisciplinary in the sense that it covers material drawn from the more traditional areas of astronomy and physics, chemistry, and biology and geology. We have organized the course around three themes, each of which is covered in about 10 weeks. Topically, the three themes deal with celestial and earthly motion, the nature of matter, and the history of the earth and life on earth. Alternatively, these themes can be characterized as dealing with the emergence of the heliocentric model of our planetary system, the fundamentally particulate nature of matter, and the theory of evolution. Each of these stories is treated historically; rather than state contemporary beliefs, we devote our time to following the development of the major concepts that lie at the foundations of science today.

We have chosen this approach for reasons that are discussed in detail elsewhere (Shahn, 1990). These include the idea that this historical approach demonstrates the fact that today's scientific concepts have resulted from a process of continual modification. We believe that, compared to a simple declarative statement, this repeated demonstration is more sound as a way of countering the often implicit belief that scientific knowledge is a form of "truth" that is "discovered" in a form that lasts forever. In addition, we believe that for many students the story-lines that we develop provide a structure that can support the scientific information that on its own may be too forbidding. (Unfortunately, interviews with our students have shown that a number of them view this historical framework as just that much more material that has to be memorized. As will be indicated below, this immediately tells us something about the cognitive positions of those students.) Finally, the use of our narrative structure enables us to show the frequent instances where science is a part of a cultural whole and both depends on the contemporaneous intellectual environment for its development, even as it contributes to this environment.

The content of our course includes material that can be covered in the traditional way. That is, we can ask students to recognize names, reactions and objects, to solve problems dealing with motion and reactions, and to say (i.e. write) something about sequences of discovery or

patterns of events. Because we know that many of our students are weak in math we have tended to avoid an exclusive emphasis on numerical manipulation. While problem solving of this sort is important, if this were too heavily stressed we believe we would be dooming too large a part of the class to poor grades before the course even began. Moreover, given that most of the class has little intention to continue in science, it is not clear that success in algebraic and arithmetic problem solving would have significant future benefit. We have also tried to avoid the necessity of memorizing names and relationships. Many of our students think that such rote learning is equivalent to knowledge (they may have learned in high school that memory is the road to academic success), but we consider understanding that can be demonstrated by giving individual and personal responses to questions to be more important than memorization.

Writing and Cognitive Development

Our approach to assessment, which we believe enhances learning for understanding, is to pursue two different but related uses of writing involving short answers and more extended essays. We have also experimented with informal journal writing which seems to tap yet other avenues of learning. In all cases, we are looking for students to demonstrate through writing mastery of both factual knowledge and understanding. As an audience for their essays, we ask students to choose other students, say classmates who have missed part of the course work; we are not looking for mini-encyclopedia entries or sections of texts. In reading our students' work we can easily see whether the facts are correct; but while necessary, we see this as being only part of the way towards providing a fully satisfactory ("A" grade) response. Beyond this, we look for the way in which evidence is used to justify answers, and the way in which this evidence is initially selected; subsequently described, summarized, or identified; and finally evaluated in the process.

Implicitly, we believe that the successful outcome of the study of science is science literacy (Shahn, 1988), and this entails a growth in cognitive ability. Three models which are relevant to appreciating this statement have been provided by Piaget (1972), Perry (1970), and Kitchener and King (1990a,b). (The following summary provides a background against which our student writing samples can be judged. It is not intended to represent the complexity of the discussions in developmental psychology that have grown out of criticisms and extensions of these works.)

Piaget (1972) identified several stages in the cognitive development of children that to a large extent can be described in essentially mathematical or quantitative terms. The "highest" of these is called formal operational thought, and includes a number of cognitive strategies: the isolation and control of variables, combinatorial, correlational, probabilistic

and proportional reasoning. Related to these is the ability to recognize a contradiction between a prediction and an observation. Formal operational thought follows a "preoperational" stage, in which reality is closely connected to the individual (up to about age 7-8), and an "operational" stage, in which the significance of such reversible operations as addition and subtraction are mastered (by age 11-12). Although Piaget thought that individuals became formal operational by late adolescence, it has in fact been documented that many if not most students entering college do not function at this level (Herron, 1975). It has also been shown that acquisition of this level of thinking can be enhanced by instruction (Lawson, 1985). The spread of abilities among our entering class is further justification for not stressing quantitative problem solving as one of our major goals. But also for this reason we structure our labs with enough time to work through the numerical aspects of data acquisition and reduction.

Apart from mathematically related abstract thinking, concern with the use of language has been part of the history of cognitive development theories from the beginning. In his earlier work ("Judgment and Reasoning in the Child") Piaget (1959) considers such aspects as grammar and logic (Chapter 1), formal thought and relational judgments (Chapter II), and the notion of ideas of relativity (Chapter III) in terms that are not so quantitative as appear later. This association between language and formal thinking has been further investigated by Lawson and Shepard (1970), who were interested in the relationship between written language maturity and formal reasoning. They used a quantifiable concept of the "T-unit" (Hunt, 1965) (involving the number and length of independent and dependent clauses in a sentence) as a measure of language maturity, and standard Piagetian tasks to assess formal reasoning. They concluded that there was a significant correlation between the two for males, but not for females. While language maturity in this study was quantifiable, its relationship to "writing" in a more extended context, and "thinking," remained vague. At best, a correlation was shown to exist, but not a way of using writing samples as an indication of cognitive level.

There are a number of related approaches to describing and analyzing cognitive development which go beyond Piaget. These are called "post-formal operational" or "post-Piagetian." The latter designation may just as well refer to the fact that they were developed after Piaget. As will be obvious, their description does not so heavily depend on quantitative concepts, and many people who might be described as almost innumerate may still place highly on one of these scales. In a sense, Piaget describes cognition in terms of how children *work with* the world, the alternate approaches somehow deal more with how children (and adults) see the world, or *conceive of* knowledge about the world. Most of these post-

Piagetian models grow out of the Perry scheme, originally developed by William Perry (1970).

Moore (1991) has summarized the Perry scheme and discussed it in conjunction with a number of assessment techniques, and the results of some longitudinal studies. Following Moore's approach, the scheme posits 9 positions that have been grouped into four major categories: *I* Dualism (1-2), *II* Multiplicity (3-4), *III* Contextual Relativism (5-9), and *IV* Commitment within Relativism (7-9). The earlier positions (1-5) deal primarily with cognitive growth involving knowledge and knowing, the latter with ethical concerns involving issues of identity and commitment. For our purposes, we are only concerned with cognitive growth.

In *I* (Dualism), the individual's view of knowledge is truth, or fact. This knowledge is possessed by and obtained from specific personal experience and from authorities. Thus this view of knowledge is tightly tied to an approach to education. While position 2 acknowledges other opinions or beliefs, but only as being wrong, people at position 1 cannot even get that far.

With positions 3 and 4 in Category *II* (Multiplicity), the situation changes; this occurs as people encounter discrepancies among authorities. Since an authority cannot lightly be dismissed, these disagreements are seen as reflections of uncertainty, but initially against a backdrop that asserts that certainty will emerge. In the process of confronting and accommodating many sets of multiple answers, peoples' responses change from "We don't know yet" (but we will or we can – position 3) to "We'll never know for sure" (position 4). It follows from this that if we *can't* know, one person's answer or knowledge is as good as another's. Multiplicity is thus tightly bound to what some perceive as relativism.

The movement to position 5 is noted by Moore to be the most significant in the Perry scheme "because it represents a fundamental shift in one's perspective – from a vision of the world as essentially dualistic, with a growing number of exceptions to the rule in certain specific situations, to the exact opposite vision of a world as essentially relativistic and context-bound with a few right/wrong exceptions. This transition [in the view of knowledge] transforms the student's attitudes about learning and his/her role as a learner ...; the self is finally understood to be a legitimate source of knowledge along with the authority ... ." Compared to position 4, position 5 provides significant options because "the person has come to understand the significance of defining rules to determine the adequacy of arguments in specific frameworks; the person has become more comfortable with developing his/her own expertise; the person has explicitly acknowledged him/herself as a judger and a chooser."

Perry positions were originally determined as the result of interviews. An alternative approach in which writing samples were used is

described by Hays and Brandt (1992). They look at the way in which arguments for specific points are structured, as a debate might be, to convince or persuade audiences that are described to be sympathetic or hostile. By looking at the number of instances in which evidence is used in short essays, and the ways in which this is related to the thrust of the argument, they are able to assign cognitive levels to the writers. This approach is clearly cast in terms of situations in which there is a "pro" or a "con;" it is not clear that more traditionally academic subject matter can be treated in a comparable fashion.

Kitchener and King (1990a,b) have developed the Reflective Judgment Model which has its roots in the work of Perry (1970) and John Dewey. It relates cognitive development to a set of assumptions about what can be known and corresponding changes in how beliefs are justified when people are faced with uncertainty. These assumptions develop through seven stages. In (1), which they note is probably found only in young children, "knowing is characterized by a concrete, single-category belief system" based on a person's concrete experience. In (2), truth or knowledge is assumed to be attainable, but possibly still not at hand. For this reason, some people may hold "wrong" beliefs. Kitchener and King say that this stage "is most typical of young adolescents, although some college students continue to hold these assumptions."

By stage (3) the inaccessibility (if only temporary) of truth is acknowledged. "Beliefs," they say, "can only be justified on the basis of what feels right at the moment." They note that "[s]tudents in their last two years of high school or first year of college typically score at about Stage Three." In stage (4) "the uncertainty of knowing is initially acknowledged and usually attributed to limitations of the knower." This does not refer to a mental failing; it means that some things are just not susceptible to knowing. In stage (5) knowledge is contextualized. They say that the reasoning characteristic of this stage is most typical of graduate students. Stages (6) and (7) are characterized by an increasing appreciation of the relationship of knowledge to interpretation and context, and are rarely found among undergraduates.

Both Moore (1991) and Kitchener and King (1990a, King and Kitchener, 1994) discuss means of assigning appropriate positions or stages to individuals. Perry's original work grew out of interviews, and interviewing remains one of the preferred ways of probing a person's view of knowledge. But because it is extremely time-consuming, attempts have been made to develop standardized essay prompts and paper-and-pencil instruments for this purpose.

The results of both approaches indicate that students seem to improve gradually more as a function of schooling than as a function of age alone. That is, older students entering college for the first time will typi-

cally be positioned with their class, rather than with people of the same age who have completed several more years of school. Statistically, Moore (1991) notes that college freshmen have a Perry position of about 2.75, while seniors are at about 3.0. Where students have been followed for a semester, roughly half of them show an increase of 1/3 position or more. Kitchener and King (1990b) observe a change in average stage from 3.6 for freshmen to 3.99 for seniors. This sits in the middle of a pattern that shows continual growth from 2.79 for high school freshmen to 5.04 for advanced graduate students.

It is clear that the Reflective Judgment model is in many ways quite similar to Perry's scheme. In their description of it, however, Kitchener and King choose to emphasize how individuals deal with evidence, rather than on the learning environment in which knowledge is acquired. This makes it a particularly appropriate way to look at how science students approach the content of their courses. Thus, beginning students often believe that science deals exclusively with facts, and see a science course as one in which these facts are transmitted from teacher to student. It takes time for these students to realize that aside from measurements (which, in fact, may not even be exactly reproducible), science is a process of determining relationships among facts, and that this process requires interpretation of facts. That is, inferences must be drawn, and when appropriate, tested.

As for a relationship between actual "scores" or positions and academic performance, it has been frequently noted that when students approach a "foreign" subject (such as science), they are likely to regress, and function at levels below those which they show on other tests.

Writing in *Foundations of Science*

From these points of view, we can now describe how we use writing in science courses. We want students to express their understanding of science in terms of facts, application, and appreciation of the process by which significant generalizations have come about. This latter is part of the understanding that over time, even the most solid concepts of science have been and are likely to be subject to modification. That is, in developing an appreciation of the validity of scientific knowledge, students should also acquire a feeling for the limitations of science. Directed writing provides a means for ensuring that students devote the time and reflection necessary to develop this appreciation.

In Foundations of Science we use writing in two different ways: on exams, and in essays. Midterm and final examinations consist of a choice of 25 of 30 or so questions which can be answered in one or two sentences. The exam questions are selected from a larger number which have all been distributed at the beginning of the term. In all, we have prepared about four questions per lecture which comprise this set. By design, the

answers to these questions summarize the content of each lecture. We hope that students will direct their attention to preparing answers to this small number of questions. Since these are all available, students may check their answers with their peers; if a group agrees on the substance of an answer it is highly likely to be correct. For these reasons, our expectations of the answers to these questions are fairly high.

The second use of writing is in essays. There are four short to medium length assignments (1000 - 2000 words) per term. Three of these are returned to the students with extensive comment/criticism which they can use as the basis for a revised version. The revised paper is then used for grade determination in the course. In addition, beyond the three revised and one unrevised papers that are prepared at home, the final includes an essay question that is written under traditional exam conditions. The exact wording is not distributed beforehand, and there is no opportunity for revision.

Within this general format, we try to make the specific essay assignments increasingly sophisticated. Thus the first essay asks students to define, describe and give examples, but not to explain. The second essay asks for summaries of the use of models in explanations. The third essay asks that lab work be related to concepts that have been covered in reading and class contexts. And so on for subsequent assignments. Thus, we start with facts dealing with what the students actually observed, and descriptions of phenomena in terms that do not require explanation. We then proceed to use more complex relationships, and eventually to require that students evaluate their evidence to justify their conclusions. In the most common example, lab data must be examined from the point of view of reproducibility to establish its validity. By the end of the year, when we ask students to discuss the way in which different geological theories have been supported in the past, or for the type of evidence that supports the theory of evolution, we are expecting a much higher level of performance. We do not accept simple statements that a fact supports or is consistent with a conclusion; we want students to evaluate the evidence and lay out the reasoning that makes it relevant. For students with no prior writing experience it is unreasonable to expect cognitive growth to occur at this rate. But college courses typically do expect students to write some sort of explanatory text, and we feel that by working up to this stepwise, if not slowly, we may be able to help students realize that there really are differences in cognitive positions.

Clearly, the bulk of the student's grade is writing dependent (the only part that isn't is a 10 point contribution from lab work) over which the student has a considerable degree of control. By making writing important, we believe that students are shown that they should take it seriously.

Writing Samples

The following samples will be discussed in the context of their reflection of the students' cognitive level or position. All samples are drawn from student work in *Foundations of Science* at Hunter College. The majority of the students in the course – as in the College – are women, so we have elected to use female pronouns inclusively when we refer to an individual student's work. Our comments will concentrate on two points in the continuum of cognitive development discussed above, the student's relationship to knowledge (is it "truth"?) and the use of evidence. Neither we nor any other group we know of has reported significant cognitive development as the result of instruction in the course of one semester. Moore (1991) is quoted above as observing that roughly half of the students given pre- and post-course tests to determine cognitive position had increased theirs by more than one third, but both he and Kitchener and King (1990b) say that the change during all of college is only about 1/2 of a position. It is not clear how these two findings should be reconciled. In a personal communication, Kitchener notes that the 1/2 figure is based on averages across several non-equivalent samples; she also notes that longitudinal studies show individual changes of up to two stages. In fact, we are not really concerned with determining a student's absolute or relative cognitive position. In what follows we discuss student writing from a "naturalistic" perspective, because we believe that an instructor's awareness of how an essay can be read for purposes other than "writing ability" or scientific content can help in gauging the mode of presentation of material for a class, structuring assignments and providing constructive criticism for the student.

Short Answer Questions

Our experience with these questions has been mixed. Because they are all distributed beforehand, we are likely to see the results of pre-considered answers that have been memorized, or that are recollected. Also, since students are encouraged to study together, this recollection may reflect group effort, and not an individual's abilities. For these reasons, this is not the best way to see how individual students approach knowledge or knowing. In fact, we often see that students get high grades on some questions and low grades on others, indicating that in their preparation was uneven. Thus an average grade of 75 would not mean that the student got roughly 75 percent on each question, but, more closely, got 75 percent of the questions correct. In part, the fact that students do not get all correct answers indicates the way in which students misunderstand questions. That is, their answers are valid representations of what they believe to be correct, even with time to reflect on them. Some students, for instance, will avoid or do poorly on those questions that clearly demand

more abstract reasoning or understanding. Thus the differences that show through on these questions represent the variety of cognitive positions within the class.

The question from which answers will be discussed is one of the few on the final which specifically asked the student to deal with evidence; the others dealt with the more factual aspects of the course. **"What experimental (observational) evidence convinced Count Rumford that heat was a form of energy?"**

The type of answer we were looking for was supplied by one student: "As he rotated a cannon he found that heat was being produced and when he stopped the cannon from rotating heat was ceased and so was the work." A variation on this was, "He observed that when a force was exerted on a cannon, making it spin against a cutting tool, heat was formed. As long as work was done on the cannon, heat was produced. When the work stopped, the heat production ceased." These answers might be edited for print, but otherwise they are "textbook" examples of a long story in short form. They deal with evidence in a context in which a traditional inference is made.

The next answer, on the other hand, is not just "wrong;" although some of the imagery is correct, it confuses description of evidence with explanation. That is, it assumes the concepts that the observations are supposed to justify. "Heat was transformed against a cannon which was rotated against a cutting tool. Heat was exerted against the object to make it move because heat that was being generated was a form of work. When work ceased, the object did not move." The picture of a cannon being rotated against a cutting tool is fine, but heat is being used in too many ways: it is being exerted and being generated. These two usages are connected by "because," which also adds to the confusion; if anything, the second (correct) clause should precede the "because," not follow it. If this student had to name forms of energy, it is possible that she would have included heat, but this is not clear; while she equates heat with work, this may be simply a restatement of the question. Thus, while she may have the rudimentary concept of conservation of energy, she is not at all comfortable in discussing evidence, and may still see all "true" statements as of the same sort. That is, the description of an observation and the reasoned conclusion based on this observation may have the same truth value.

And what should one make of the following answer? "He accelerated two pieces of gum together and shot a ball out of a cannon." This student also has the cannon in the picture, but little else. In a multiple choice context, where the key words in different answers were cannon, phlogiston, caloric, and calx, it is likely that she would have made the right choice, but that would not really have indicated any sort of understand-

ing. In her written answer, however, one can actually see some other aspects of recall, given the context of the course. We discussed inelastic collisions, where mechanical energy was specifically not conserved, and considered the example of two projectiles fired at each other with chewing gum on them so that when they hit they would not rebound. We asked what would become of the kinetic energy, and suggested that it would be dissipated as heat.

A different piece of fancy is seen in this answer: "He observed the trembling and temperature change (very hot) of a cannon after it was fired – concluded that the energy from the collision (inside the cannon) was stored in the form of heat." The ring of truth here is the temperature change of a cannon, but for Rumford this was *not* after it was fired (he didn't fire the cannons he worked with), and not *because* of any collisions (the kinetic-molecular theory of heat was introduced much later, even though this "experiment" provided foundational evidence), and not because any energy was *stored* as heat (as usually described, heat is not stored in this procedure). There is certainly something to work with in this answer, but little to grade.

Finally, these students may know more than they are able to write about. The answer, "He observed a drilling of the hole on a cannon." provides the setting for the evidence, but that's all. In general, time was not a factor with students taking this test, so one cannot justify the idea that she was rushed. More likely, she really was not comfortable with the idea of evidence, and she was trying to construct a minimal understanding that included our story. A similar explanation may account for the answer, "When two objects were rubbed together quickly they grew hot." This is not wrong, but it omits the details of cannon boring, and the fact that Rumford noted that large amounts of cold water could be boiled away. Indeed, the realization that two blocks of ice could be melted by rubbing them together, and hence that work against friction caused heat, is often attributed to Davy. This student may really have a fairly good grip on the subject, but it is not expressed in context, that is, it does not address Rumford's experiences and his line of reasoning.

This last pair of answers gets to the nub of the problem of using short answers as a gauge of any sort of student mental activity – too much may be left to the reader's imagination. Often we are left with less than the hoped-for distinction between a right answer and a wrong answer; we seem to want to know *why* an answer is wrong so that we can distinguish between a wrong answer and a very wrong answer. And here we realize that the *why* above has at least two meanings. In the first place it is "in what way?" Is it a matter of fact that is misstated, or is it an interpretation, or an inference, that is wrong? In the second place it may be "for what reason?" That is, does the student go astray because he or she mis-

remembers; or reasons incorrectly; or is incomplete; or the correct answer is counter-intuitive, and the student is tied to recollections of prior personal experience; or ...?

Nevertheless, because they in some way force the student to go through the motions of formal thinking and justification, short answer questions of this type are still probably more valuable, and for many students less threatening, than the multiple choice questions that would take their place. While students may still have the option of viewing their grade as a sign that they are either "right" or "wrong," they also have the opportunity of seeing how their answers can be improved, and where they went astray in interpreting the statement of the question. Thus the reasoning processes that we hope are the central point of the students' learning experience are also the focus of the students' exams and grades. But in addition to the utility of this sort of question as a means of assisting in grade determination, the answers also provide the instructor who has the time with a clue about the way the student is thinking. The fact that several questions have been asked above, does not mean that we have reached an instructional dead end. Rather, these questions may be used as the framework of a conversation with the student. In this way both the instructor and the student may find out what the difficulty is, and the teaching process will have passed through a door that the use of a different assessment strategy would not have opened.

Exam Essays

The exam essays provide still other insights into what students know, and how they approach knowledge. Because there is more time and space devoted to the answers, they have more opportunity to show what they know. However, because we are not using questions that they have previously seen, we are not getting canned answers, but rather responses that are developed on the spot, and under some sort of pressure. For this reason they are also not edited, and show a variety of human errors.

The question discussed below asked students to **"Discuss the formulas for water proposed by Dalton and by Avogadro, and the evidence used by each."** The subject matter of this question was critical in the development of chemistry at the beginning of the 19 century; Dalton proposed that water was composed of only two atoms (one each of hydrogen and oxygen), whereas Avogadro suggested the still accepted three atom formula of H_20. Not all students had to answer this question – there was a choice – and of those who did, there were a number who ignored part of the answer, often that dealing with evidence. Another frequent confusion was the use of Avogadro's EVEN hypothesis (Equal Volumes of gases at the same temperature and pressure have Equal Numbers of

particles). We spent time on this in class, and its inappropriate inclusion seems to reflect partial knowledge, or simply an association.

The essence of a complete answer is shown in the following. It not only discusses the formulas in the context of who did what, but alludes the reasoning processes that are ascribed to Dalton and Avogadro. In this sense it ties together the process of science as we see it today, and by comparison indicates how it has changed. "Dalton and Avogadro used different techniques in determining chemical formulas for compounds. Dalton based his formulas on work done by Lavoisier whereas Avogadro based his formulas on work done by Gay-Lussac." [The work of Lavoisier referred to here consists of three parts: the identification of elements as simplest types of matter, the generally accepted notion of conservation of mass, and the beginning of the realization that chemical reactions take place between fixed proportions of the masses, or weights, of substances. Dalton used this as the basis of his atomic theory. In this context Gay Lussac is noted for making critical measurements of the ratios of the volumes of gases that combined with each other in chemical reactions. He concluded that not only were the masses proportional, but in the case of gases, so were the volumes. Avogadro inferred from this that equal volumes of gases must contain equal numbers of particles, even though these particles could not be seen or counted.]

"Dalton believed that elements combined in simple ratios. So his chemical formula for water would be H + O –> HO. This would mean that one [sic] volume of Hydrogen would combine with one volume of oxygen to [produce] one volume of water.

"Avogadro was familiar with Gay-Lussac's law of combining volumes. This law found that [two volumes] of Hydrogen would combine with one volume of Oxygen to make two volumes of water. Avogadro then found the chemical formula for water to be H_2O. He found that in their natural state, Oxygen is O_2 and Hydrogen is H_2.

"Dalton and others could not accept this but eventually they had to. By using the formulas H_2 and O_2 in other experiments it was found that elements combined in fixed ratio relations but not always in a 1:1 ratio."

This answer refers to the requested evidence in two ways. In the first place it uses the accepted shorthand of mentioning the names of the people who did the work – Lavoisier and Gay-Lussac. Then it goes further to give the main thrust of their observations and summary conclusions, especially the latter. Finally, it summarizes the reasoning used by Avogadro, and contrasts it to the beliefs of Dalton. In the context of this course, this answer is excellent. It shows the student's grasp of the both the scientific (i.e. phenomenological) and historical facts, and how they are related.

The following answer gets off in the right direction, but the writer doesn't seem to know when to stop. It's similar to the young child who

continues counting cars on a train after the train passes – the concept is not quite fixed. Evidence is mentioned, mostly in appropriate places, but it does not seem to be really digested. This student is using a rote approach (even though this particular question had never been asked), and will most likely build understanding on top of it. "Dalton believed in the atomic theory but believed that atoms combine in simple 1 to 1 ratios which tied in with Lavoisier's law of conservation of mass. According to Dalton water would be created by 1 atom of Hydrogen combining with 1 atom of oxygen to get 1 [atom] of water, which is true [sic], but Avogadro found when using electrolysis on water (which separates the hydrogen from oxygen) there was twice as much hydrogen as oxygen being separated out. Dalton felt that this could be explained by saying that perhaps hydrogen atoms are bigger than oxygen atoms and take up more space. Avogadro didn't think this was the case. He believed that elements could contain perhaps more than one atom in combination. The smaller atoms move faster than the heavier atoms bouncing off the wall of an enclosure more frequently than the heavier slower atoms." The last sentence has gotten off into the details of the kinetic-molecular theory, which was beyond the scope of the question.

This next student misinterpreted the use of the word "formula" in the context of the question. She wrote: "In this time, Dalton had his way of writing the chemical formulas which were circles in different amounts describing the compound. Avogadro had written his formulas in symbols according to their Latin meaning. This made it easier for the formulas and compounds to be understood. They were more like symbols than circles. It was less of a hassle to remember, and draw on to the paper." In answering in this way she completely ignored the last part of the question asking what evidence was used. Is this an instance of simple mis-reading, or of the fact that the concept of evidence was so foreign that it was not even seen?

By comparison, this answer shows that the student knew what evidence was in principle; it is part of the structure of her essay even though there are many mistakes of fact. Despite this, the justification for the "feelings" and "beliefs," that is the reasoning, is lacking. "Dalton felt that everything combined in a one to one ratio so he expected water to look like HO, while Avogadro believed that water combined as H_2O. Not all formulas will combine in a one to one ratio as they all have different weights.

"Dalton felt that 1 vol of hydrogen plus one gram of oxygen is equal to one volume of water. While Avogadro believed that 1 volume of hydrogen plus one volume of oxygen would give 2 volumes. ..."

The following sample ends with the personal parenthetical note to the teacher: "I know this is not well written and vague and sounds stupid, I can't explain how difficult it is..." In fact, it hits a number of the key points

very well. "Dalton and Avogadro disagreed about chemical compounds. Dalton believed that fixed numbers of atoms of one element combined in the same number of fixed atoms of another element to produce molecules. He believed in the simplest 1 to 1 ratio and, so, concluded water for example, to be one atom of hydrogen and one atom of oxygen or HO. He worked with data provided by Lavoisier and believed that atoms were indestructible and indivisible.

"Gay-Lussac determined relative atomic wts and Avogadro's hypothesis was influenced by the atomic wt of particles. He determined that molecules could combine according to volume but the atomic structure of a molecule could be that of more than one atom of a same atom combined with a different number of atoms in another element although the simplest ratio was used. He thus observed that a water molecule could be two atoms of hydrogen with one atom of oxygen.

"This was controversial at the time, because scientists believed that "like" atoms would repel each other. The bonding principle of atoms of the same element was difficult to accept.

"Dalton believed in the simplest ratio of atoms combining only on a one to one ratio so could not accept this theme. Avogadro proposed that the volume could still be proportional with the same number of particles, but they could be rearranged in proportions [with H_2O] as the simplest form..."

The organization of the next essay, as well as its general literate tone indicate a high degree of understanding, but while it refers to the sources of evidence, it still manages to avoid detailed discussion. This is possibly due to the student's feeling that she really didn't understand the details. However, the structure that is presented here is certainly one that could be filled out. In reading this it is also interesting to note the care with which a number of distinctions are made. At this stage in the development of the atomic theory, there is a real conceptual difficulty in distinguishing among atom, molecule, and particle, and this student's response indicates an awareness of these problems. Her discussion of Avogadro's concept of the water molecule seems to touch all bases.

"... Dalton's point of departure was his belief that atoms were the simplest form of any element and it logically followed that they would combine in the simplest possible of ratios. Inherent in this concept was the belief that one particle of something would only have one atom. He took details of his theory from the work of Lavoisier, who showed that the sum of reactants in an experiment were equal to the sum of their individual weights – so Dalton interpreted the combinations of particles to be in the simplest forms possible. He hypothesized that water could only include one particle of hydrogen and one of oxygen and that they each had one atom each. He used the work of Lavoisier on conservation of mass, experi-

ments in combustion, and the reactions of elements with each other to back up his hypothesis. ..."[Here, and in the next sentence, the knowing reader will be able to follow the argument that is being presented, but the reader lacking specific knowledge will not be able to infer the specific evidence being referred to by names. This is where a detailed discussion would have helped.]

"Avogadro analyzed the work of Gay-Lussac concerning the combinations of substances and proposed that particles could contain more than one atom and combine with different ratios than Dalton had thought, and that a water molecule could only logically be constructed by the combination H_2O. His own notion [was] that each element had its own unique number of atoms per particle with which to combine, and that many of the ratios of common elements ... were always made with fixed proportions ..."

In the discussion of short answers we observed that the constraint of space might have made it difficult to interpret that student's intentions. In the several examples of exam essays just given we see that an incomplete answer might often be accounted for in a number of ways. In some cases it may seem clear that the student probably couldn't have done any better, but in others it may be that the limitation was a combination of reticence and inexperience. Even though our course had been concerned with the use of evidence, and this concern was exhibited in both the reading assignments and the way in which the corresponding material was discussed in lectures, and we had asked students to consider this aspect of the work in essays during the term, we may not have adequately indicated what our expectations were. Our recommendation to people planning to introduce a writing component in their courses would be that they be scrupulously clear to their students regarding the general nature of their expectations, and specifically how these expectations should be met.

Essays

Our experiences with more formal student essays have revealed several features that have been obvious to any reader trying to get past the question of whether a particular answer is "right" or "complete." One of these deals with definitions; many students confuse a particular example with a more general consideration. For example, "An angiosperm is a rose," rather than "An angiosperm is a member of the group of flowering plants." Beyond this is the issue of description; here the common problem is a confusion between an explanatory account and an observation. Even when this is the accepted explanation, it might not enable the reader to recognize the object or event independently. An instance of this sort of error is describing an eclipse as the passing of the moon between the earth and the sun (a solar eclipse), or into the shadow of the earth (a lunar

eclipse). While these explanations are correct, they do not convey the experience of either kind of eclipse. Last in this chain of accounting for events is the process of explanation; here students are often unable to string logical arguments together for more than one or two steps. This apparent inability to write clearly may well be correlated with a difficulty in reading with understanding as well. If so, this would be indicative of a need for changing instructional strategies so as to emphasize these particular rhetorical devices and provide students with the opportunity to master them, particularly early in an introductory course.

Other features of student essays reveal the students' preconceptions of the nature of scientific knowledge. Most simply, some students believe that science gives "true" answers, and that declarative statements are the hallmark of science. This is often seen in conjunction with the use of explanations when descriptions are requested. Strangely related to this is the fact that while many students refer all knowledge to themselves and their own experiences, they do not include in this experience an emphasis on clear observations; rather, they seem prepared to settle for an impression. More demanding yet is the ability to summarize and generalize; many students simply repeat what they find in their source and make no attempt to reduce it or relate it to other things they may know. Finally, there is the issue of being able to draw a conclusion that ties together the substance of the paper. The poor results offer platitudes reflecting the value of the progress of science; the good ones show a considerable degree of reflection.

In the context of reports on lab experiments these problems can all be seen to a greater or lesser extent. Most significant for all of them is the difficulty in dealing with evidence – in recognizing it, describing it, relating it to other contexts, interpreting it, and evaluating it. Indeed, these are the hallmarks of the measures of reflective judgment discussed by Kitchener and King (1990a,b).

The examples given below are identified with individual students by capital letter, and by the number of the essay. The same letter indicates the same student.

Our first essay deals with the roots of science in explorations of different kinds of phenomena. The assignment follows two hours of lecture and discussion in which we suggest that science has grown out of attempts to account for human experiences with and "outside world." It asks students to distinguish among (i.e., define), give some examples of, and describe periodic, episodic and craft-based phenomena. We are looking for awareness of personal experience with this outside world, and the ability to discuss it objectively. That is, we want students to demonstrate an awareness of those aspects of experience which are shared with other people, and to be able to describe clearly observations which may be

considered "evidence." We realize that this may not be the students' natural way of looking at things, but we believe that this is a way to introduce them to the way of the scientist. At this stage we are not looking for interpretation, evaluation or reasoning based on this description. Nevertheless, students' preconceived views of what science is – and what knowledge is – are sometimes seen in these answers.

(A1) One student's paper shows her inability to separate herself from the world. She writes: "Another example of [periodicity] would be my role as a student. It's the same routine over and over, I get up at six o'clock in the morning, eat breakfast, take a shower, get dressed and leave for school at eight o'clock in the morning. This is everyday. Now my schedule of classes isn't the same due to the fact that it's different classes everyday, different break hours, and everyday it's a different time when I go home. A third example would be the direct and retrograde motion of the planets; their movement is always either in the direct or the retrograde motion." While we were looking for personal experiences, we wanted them to be related to the outside world, and specifically to phenomena which were a stimulus for inquiry in early science. This is in part touched on with the afterthoughts dealing with the motion of the planets. One cannot tell from this example why this student did not focus more quickly. Did she not read the assignment carefully or consider it before she started writing? If so, then the remedy to be suggested is simple to give, and simple to adopt: think before you write. But if her approach was more dictated by a worldview which does not involve much consideration of phenomena in a natural setting then no such simple prescription is available. The procedure might be to find some way to engage the student in a one-to-one conversation as a prelude to giving any advice. If this is the approach to be adopted, it is clear that the commitment has to be acknowledged in the planning stages of a course, and reasonable time has to be set aside to pursue it.

(B1) By contrast, the following sample shows a more appropriate use of personal experience. "Not willing to admit defeat easily, man quests for knowledge in order to control nature. To accomplish this we see the need for an understanding of the world we live in. Perhaps this is what inspired us to literally get above it. One way to do this is aboard an airship such as a Zeppelin, a blimp. It took a questioning mind, a few disasters, more experience, knowledge and skill to create and perfect this desired phenomenon. First man had to discover the potential qualities of hot air, and then, those of gases that exist in nature, specifically helium and hydrogen. Next he had to apply his findings and develop the means to capture the gas and navigate it up to send us soaring. I saw one the other night lit against the black sky and moving so slowly between two skyscrapers. It was brilliant yet ominous. I thought, "What a strange phenom-

enon. I mean what a fascinating, craft- based phenomenon!"" While this is more modern technology than an example of craft-based phenomena that lie at the foundations of science, it is still an example of how one can generalize personal knowledge and relate it to a larger context. We must realize that few of our students have first-hand experience with metallurgy, glass making or ceramics.

(C1) Another student shows the common problem of looking for truth in short statements, and of confusing description with definition and explanation. "Hurricane seasons are periodic because there is a season for hurricanes in some part of the world. Typhoons are periodic. Places where typhoons are periodic are China and Japan. The area is mostly surrounded by water and that causes typhoons whenever the seasons come around." Some of these statements are wrong, and others may just be debatable, but that is not the issue. Even if these were totally correct statements, they would not address the question of defining or describing typhoons as examples of periodic phenomena. In fact, of course, they are not; the seasons in which they occur are periodic, but the typhoons themselves should be considered episodic in the language of the question.

(D1) By distinction, an example of a good description is the following. "To the ancient observer, one pattern that changed with some predictable regularity was the rising and setting of the moon, which assumed different shapes in a regular cycle. They observed the moon at its smallest crescent shape, when the convex side was facing west, and saw that the moon in this shape would set right after the sun. [They noted] that when the moon had filled out to become a half circle, it would take more time to set, a full six hours after the sun." Her transition to a discussion of episodic phenomena shows a recognition of the distinction between precise knowledge and indeterminacy. "People then began to distinguish between phenomena which occurred on a regular basis, and events which recurred from time to time, but were difficult to predict. These are called episodic phenomena. With episodic phenomena, one can only assume that an event may recur, but this is unverifiable, and cannot be accurately predicted. For example, many scientists agree that California is due for a major earthquake along the San Andreas fault within the next twenty years, but they cannot determine the exact date, because earthquakes do not occur with precisely predictable regularity."

(E1) In addition to the dependence on declarative sentences, the confusion between explanation and description is shown in the following. "A lunar month is the average time between one new moon and another. The standard time between a new moon and another is slightly more than 29 days. The moon moves forward in its own orbit, while the earth has been rotating, so the earth must move farther than a complete rotation before "catching up" with the moon. Thus more than 24 hours pass be-

tween moon risings. The period of the moon's revolution is used as the basis for the calendar month." This passage has the form of standard scientific prose, but aside from avoiding the process of description, its concluding statement is wrong, that is, not in accord with convention. The length of the calendar month may at one time have been connected to the lunar "moonth," but it is now a matter of consensus, and is independent of the moon. The middle sentence, dealing with the relative speeds of the earth and the moon is also open to a considerable degree of questioning. Specifically, consideration of a detailed model of the currently accepted motions of the earth around the sun and the moon around the earth would probably show that the moon has to catch up with the earth, not vice versa. But note where this criticism has taken us: instead of describing a phenomenon, which was called for in the assignment, we are discussing a model, and in particular the language needed to describe the model. In the long run this may well be a critical part of the process of science, but that is beside the point of asking students to master the technique of description of their own experiences with commonly encountered phenomena.

(F1) Typical problems arose with the examples of an eclipse and the tides. In this selection, both appear in the same paragraph. "A solar eclipse involves either partial or total darkening of the sun when the moon comes between it and the earth. A lunar eclipse occurs when the earth's shadow is cast on the moon leaving it partially or totally darkened. Tidal action is caused by a combination of the gravitational attraction between the sun and the gravitational attraction of the moon. This combination causes an accumulation of water in both oceans and seas at two opposite points on the surface of the earth. As the earth rotates it has a series of two high tides and two low tides each day." It is clear that this student "knows" the material. In some contexts this response would earn high grades, and quite possibly she has been rewarded for this sort of answer. In this paper, however, she is confusing explanation with description in a way that shows little or no reflective thought; neither is presented in any depth. Thus it is likely that a person familiar with eclipses only from this description would be able to identify one should it occur, but it is highly unlikely that the ebb and flow that is experienced at the seashore would be associated with the account of the tides that is presented.

(G1) The conclusion of an essay may provide considerable insight into where the student is coming from. In this first essay we are asking for relatively little in the way of "higher level thought processes," and thus the conclusion we are looking for is rather modest. Nevertheless, we have this sample: "Our curiosity may not ever be satisfied because we are delving further and further into space as time passes, but at least we have a basic idea of how our world and heavenly bodies surrounding it be-

have." The paragraphs that preceded this were in fact quite good, but such an end seems more like the hero riding off into a sunset than a reflection on what has been discussed. Having said that, however, it is not immediately clear how an instructor should respond. The question of why the student chose to conclude in this fashion remains unanswered. Was she responding to previously learned ideas of the proper form of essays? Was she searching for some form of closure in terms of her own needs? Was she expressing her beliefs regarding general human motivation? And if the latter, were these long-held and carefully considered, or were they born out of the assignment? Clearly, answers to these or other similar questions are needed; how one encourages a student to grow, and the types of growth to be expected will depend on the ground in which this growth is rooted.

In our second essay we were looking for an ability to summarize old arguments, and equally important, for an understanding of what a "model" is and for ways in which evidence can be used as a reason for accepting a model or for changing it. Specifically we were asking students to consider early models of "the universe," to identify these models, name a person associated with each, describe the problems (i.e. the phenomena) they dealt with, and indicate how (outline the reasoning by which) the model accounted for the phenomenon in question. In this essay we are approaching the difference between science — and its concern with an outside world – and other disciplines.

(H2) This student had shown in her first essay that she was unusually able to revise her first draft; she went far beyond the all too typical minimalist approach of simply changing the "offending" word or grammar that had been marked on the first reading. On the first draft of the second essay her descriptions were commended. For example: "Heracleides, a contemporary of Aristotle, proposed a model in order to simplify Aristotle's model of the Universe. Heracleides proposed that if the earth was rotating on an axis, this would produce the same visual appearance of the celestial objects moving although they would in fact be still. [H]e observed that this would explain why objects appeared to move in smaller circles in particular areas of the sky and larger circles in other areas. He reasoned that when the stars are located closer to the axis of the earth in motion, that they would appear to move in smaller circles, and those objects further away would form larger circular motions. The visual appearance produced in the sky of the planets apparent motion is one that people are able to observe on earth. This new theory of Heracleides eliminated the extra spheres of Aristotle's model since he no longer needed to account for the motion of the celestial sphere." In spite of this relatively sophisticated passage, this essay concludes: "From the examples outlined throughout this essay, it is evident that the models proposed by various scientists

have allowed men to become aware of the many possibilities that can account for phenomena. Conclusively, models are the basis for scientific explanation which will always be beneficial towards man's constant exploration of phenomena." This peroration puts science on a pedestal, and does not recognize the importance of the use of models that had been so well described earlier.

(I2) A better conclusion, which did not follow as competently written descriptions of particular models, is the following: "Models are a way to explain how a particular phenomenon may occur. However, just because a model can provide a possible explanation for a phenomenon, it isn't necessarily an accurate explanation. Philosophers have and are presenting models that disprove the theories behind some models as well as reinforcing the theories behind other models. Hipparchus developed a model that enforced Aristotle's theory that the earth was the center of the Universe. Later philosophers presented models that contradicted this school of thought. Through reasoning and observations philosophers are constantly using and changing models to develop a better understanding of the universe." This sample is interesting in that internally it uses evidence of a historical nature to substantiate the conclusion that it is presenting.

(J2) Another example of how the goal of this essay can be realized is presented in this sample, which is included here as a demonstration of how much room there is for individual expression in the context of essays. We are not looking for uniform answers. "The Greeks, through many hundreds of years, had developed an approach to scientific problems which would eventually lead to an understanding of how the universe functions. They had learned the value of models, particularly mathematical schemes, in realizing the relationships that existed among the celestial objects. They had learned to test speculative theories against observations – to use empirical knowledge to validate or disprove theories. This was a large step in the direction of a modern scientific approach." Of course there is room for discussion here, but not for pejorative criticism. But it is the type of discussion that one expects to have with advanced students, not those taking a freshman course. Which again points to the utility of using essays in introductory courses. Given the time and the commitment, it is possible to approach each student at her level, and not be bound by "right" and "wrong" answers. In this case the fact that the Greeks used their observations of the otherwise unapproachable heavens as a test of models could be contrasted to their avoidance of theoretical and experimental approaches to other aspects of their world

(B2) This woman's writing has been presented before, and will appear again. It is recognizable for both its extremely personal style and for its competent way of dealing with content as well as for the cognitive

position it illustrates. "Heracleides discovered, much to my content, that Aristotle's model could be simplified if he reversed the order of things and rotated the earth, instead of the celestial sphere, eliminating all the extra spheres that were needed to cancel out motion. [Consistent] with what he observed, if the earth were spinning smoothly and slowly things on it would appear to be stationary but the heavens would appear to be moving. Like Plato, Heracleides succeeded in shifting our perspective tremendously.

"In conclusion, I am inclined to repeat myself and say how difficult it is to put myself in the early Greeks' position. To look towards the sky when one's mind is full of questions is common among all men, but it is those like Plato, Aristotle and Heracleides that, remarkably enough, found answers. Although the quest for an understanding of the motions of heavenly objects underlies the work of all three, each had previous information to either accept and build upon or reject and change in their own models. It is here, in these models that I come closest to seeing this as they did, a change in my present perspective. Their use of the model was to get them to see, to understand, to know what I have knowledge of today. My use of the model is to follow their progress, their reasoning and the evolution of heavenly knowledge; the ability to see what I cannot see with the achievements of man embedded in my mind; and perhaps, most importantly, to understand why I accept what I do today as truth." This student's first person exposition makes it clear that the writing and the ideas are hers. In some of the more descriptive passages this type of conclusion is less clear. In fact, in many places where the source of information is the reading material that we have assigned we simply see it copied or only moderately rephrased. This poses a problem, but one that can be addressed in a number of ways. Making sure that one sees samples of the students' own writing can be assured if a specific personal interpretation is requested, or, in a lab context, if descriptions of the students' own procedures and/or observations are part of the assignment.

In the third essay we asked students to summarize the beginnings of the modern mechanical view of the world and relate it to a series of associated lab exercises. Particularly, we wanted students to focus on the use of procedures to obtain data as a means of testing or validating "laws." In this context we see data as evidence, and the way in which it is handled as one means of observing a student's cognitive position. As indicated above, this is an occasion where we do get at the issue of seeing a student's own writing and thought processes because of our emphasis on specific lab-related observations. Interestingly, this is also the case where we often see discrepancies between the way in which students relate vicarious experiences encountered through words and the way in which they relate their own experiences which do not have a verbal component associated

with them. In some cases it would seem that vocabulary exercises would be extremely valuable, but we have not developed any of these.

(K3) A particularly articulate student started her paper with a lengthy discussion of her view of "modernité" which she conceived in very broad terms. She writes, "Frankly, I believe the Egyptians and the Babylonians were as modern in their view of acquiring knowledge about their world for practical, functional reasons as the civilizations that followed and benefited from them." Several paragraphs later, after referring to the astronomical observations of the Babylonians and Egyptians, and the functional role of their knowledge, she asks, "In this so-called modern, electronic age, do we always ask how a computer chip does its work in order to be able to use it? Our concept of what is modern may, in fact, only be a scientific "second-coming" or a scientific difference of opinion." She continues, "Having a "modern view" of our universe is neither better, nor worse, than not having one. It is merely another point of view and seems relative to one's cultural and/or philosophical values."

This student clearly writes well, and is familiar with a range of material that we do not cover in the course. Her conclusion is certainly not unique. What is disturbing is her lack of distinction between knowledge about the world, and about man-made devices, and the fact that the latter are designed to perform in certain ways on the basis of other knowledge. Further, in her last quoted sentence, despite its overt relativism, there is an implicit denial of the use of evidence in creating a scientific view of the world. Given this conclusion, it may not be coincidental that this student described the experiments and summarized the "desired" conclusions of confirmation, but neither gave nor discussed any data.

(D3) More to our liking was the discussion by a student introduced above. "Through this series of labs it became evident that the process of experimentation is an essential tool in verifying scientific theories. The fact that each person in the class (or each pair) was working independently and all arrived at similar or related answers, to me proved that the results were reproducible and verifiable. We also saw that they were reproducible by the fact that we often repeated portions of the experiments numerous times and they came out relatively the same each time. Before conducting the experiments, we worked out the expected results with help from the formulas in the theories we were testing. In almost every case, the results of the experiment came very close to our predictions. In cases where it did not, the discrepancies were mostly attributed to human error. If they hadn't agreed at all, I would have interpreted that to mean that either the hypothesis was faulty, or the experiment was not appropriate for testing it, or there were major mechanical problems with the scientific instruments of the experiment. In general, experimentation is an extremely valuable tool in making abstract concepts tangible, and the theories we

tested became more convincing once we saw them illustrated three dimensionally. It would be extremely difficult to digest these complex theories if we had only read written explanations of phenomena which are difficult to conceptualize, and it would be ironic to study motion in the classroom if we didn't incorporate some hands-on activity in our analysis, and were always standing still." Our feeling is that the kind of appreciation of experimentation and data as represented by this essay is a more appropriate goal for the sort of course we have put together than would be a line through a set of points, or a statistical test of significance for a particular null hypothesis.

(B3) This woman, whom we have also met before, concludes her essay, "In the end, I feel quite satisfied that each hypothesis we set out to test was confirmed through these experiments. My new understanding of the methods developed by Copernicus, Kepler, Galileo and Newton, and their theories of motion, certainly would assist me in reproducing these experiments with similar results. More importantly, this experience gives me a foundation from which I may approach future situations where the process of "scientific method" is imperative. I won't dare say for sure, but this knowledge may just have more potential energy taking me closer to the moon than the Bible does. But then again, who knows?" As in the previous example, it seems clear that this woman knows what she's come up with. Both women display a degree of humor in their writing, which is not typical for scientific prose, but in this case that doesn't really matter. The difference, which is one of style, and not related to degree of success in the course, is that the woman whose work is quoted in this paragraph seems to take the enterprise more personally.

(L3) Another example of growth is seen in this sample. While her treatment of specific data was not as strong as we would have liked, her discussion of errors, and of the development of science both show the beginnings of a personal understanding.

"The experience of labtime is helpful in having a "hands on" relationship with experiments. In the process of verifying laws/theories, I found that our experiments were reproducible but our calculations did not always agree with our observations. I learned that even though our results didn't match up with our expectations, that it didn't mean that our hypothesis was wrong. I learned that in experiment we had to allow for errors. Errors can be caused by reaction times as in the experiments in lab 6, or just human errors as well as mechanical error. Averaging is also important because we can never really get an exact result.

"In conclusion, I would like to note that integrating the experiments with the development of a new modern view has helped me to see the whole story as one big piece rather than bit by bit. We saw how gradual change led us to a new modern view. There was a sort of domino effect as

each scientist knocked down an idea of the one before. We saw how Copernicus got rid of Aristotle's idea of a geocentric universe, how Kepler got rid of circular spheres, and so on. Learning how we came to today's conclusions has allowed me to better understand the laws of motion. My only gripe is that just as soon as I become comfortable with one theory, I am bombarded with the next."

Despite the gripe, it is clear that this student is reflecting on the material being covered. She has introduced her own analogy for concept modification. So, even though she complains of the pace of the course, she was able to grasp both our intention and the historical development of these concepts.

The fourth essay deals with the modern origin of the study of gases and asks that the work of four men be considered jointly. In a sense it is a "compare and contrast" exercise, but in addition we want the use of evidence to be considered, as well as the validity of using the behavior of a spring as an analogy for the compressible nature of air. Although as in the third essay we are again asking students to use evidence in an historical context, the events discussed are not so obviously sequential. This makes the structure somewhat more complex. At this point – after three revised essays – it is often possible to note changes in a student's sophistication as seen by her reliance on evidence to justify conclusions.

Many students followed the text they were given almost to the letter. In this case we were able to provide translations of 17th century sources, and in narrating what each man did, the same archaic language of the time was presented to us, as though that was part of the story. We take this to be a sign of lack of understanding at several levels. But not all students reacted in this way.

The writing sample and previous comments about this student (G1) suggested that while she could describe well, she was nevertheless looking for some grandly simple goal for science. In this essay (G4) she begins by saying "Due to limited information, the earlier scientists formed many false assumptions [about] atmospheric pressure and a vacuum. Torricelli was a mathematician who was influenced by the writings of Galileo. He demonstrated the weight of air by experiments with mercury filled tubes, and correctly distinguished weight and pressure. Torricelli believed that air is a substance, air has weight, and that a vacuum is natural. He also believed that the nature of the atmosphere can be more or less 'dense.'" She goes on to describe some of Galileo's experiments, but she does not relate these to his beliefs. It is almost as though she were following and revising a text, but not internalizing it.

She concludes this paper, "It is very interesting to see just how far our study of science goes. I must say that I never would've thought of experimenting with tubes filled with mercury to observe barometric pres-

sure. I would not have thought of different levels of pressure changing with altitudes. I am pleased to say that although I came into the course thinking that I had a barrier to science in my brain, now I do not. I am very interested in the observations and the experiments we've conducted, and the conclusions that we, like the ancients, arrived at. Simply, now I know that I do appreciate the study of science and also that if one applies the mind, anything can be accomplished and easier to accept." This very personal statement displays a mind aware of itself and in transition. More than most of the writing samples we have given, this one, jointly with the previous selection from this student, may indicate the kind of growth that writing can encourage and display.

(H4) In her summary descriptions, this student avoided both the language and the detailed narration of failure that typified the original and many of her classmates' papers. For example: "Otto von Guericke (1602-1686), a German military engineer, was aware of Torricelli's experiment with a column of mercury and how a vacuum was produced, and wanted to further investigate the existence of a vacuum in nature. He used a qualitative approach in which his aim was to construct an air pump which would indicate the existence of a vacuum. He reasoned that if he could construct a vessel, fill it with water and pump the water out, then there would exist a vacuum since there would be no air inside. This procedure was a complicated one and he was unsuccessful in his early attempts due to the pressure that was exerted on his vessels. Either his vessels would fly apart from the pressure, or the wood that he used was too porous to resist the strength of the pressure which allowed air to seep in. He eventually succeeded in his attempt by constructing a perfect copper sphere. Von Guericke's approach was primarily a qualitative one in which he designed an experiment to further explore the existence of a vacuum. He had not used any numerical quantities in his approach and it is therefore considered qualitative."

She finally reaches the question of whether the analogy of a spring is appropriate. "Boyle used the analogy of air in the atmosphere behaving like a spring. He was aware of an elasticity in the air and stated that air particles were like little bodies, one piled on the other and may resemble fleece of wool which are flexible and may compress or expand with the weight of a force applied, like a spring. This analogy is conceptually similar (to the wool) yet if we attempt to quantify [it], we see it breaks down and does not hold true.

"The atmosphere seems like a spring because with a greater amount of atmosphere above us, it becomes more dense and more compressed at the bottom and less dense and more rarefied on top. Yet, by looking at the results obtained when measuring the different lengths obtained as increased amounts of force are exerted on a spring, we can see a propor-

tional extension that occurs with the amount of weight applied. Yet, the force which occurs in the atmosphere does not behave in such a quantitatively linear manner. A larger force is necessary to compress air as it becomes more confined in a volume, and through analysis it is evident that there is a difference in the behavior of the force required to compress the air. If we graph the results obtained when testing this relationship we can observe that the air does not expand proportionally (with decreased force) and disprove Boyle's analogy. It seems that the analogy of the air being like fleeces of wool would be a more appropriate model since the force of each layer of wool would compress the bottom layer of wool and would necessitate greater amounts of wool to compress the bottom layer while the top layer would remain less compressed. However, if we would want to have this as a suitable model, it would be necessary to test this relationship."

Compared with her previous papers, in which she had demonstrated an ability to describe, but still had a tendency to be somewhat florid in her view of science, this student has come a long way. Her writing might be edited, but her concepts are remarkably strong. She summarizes Boyle's work by saying "By instituting an 'if...then' approach to his experiment with the mercury he was able to verify what he had predicted." She then goes on to conclude: "As with all areas of science, we may look at the progression of the various observations made by the scientists of the 17th century and notice that each man's contribution was significant to the growth of knowledge obtained in pneumatics. However, it is through the use of hypothesis testing to obtain quantitative results, such as Boyle used, that we may develop evidential conclusions to substantiate the implications proposed by these men. With Boyle's carefully thought out controlled experiments, he successfully obtained knowledge about the substance of air. This enabled future scientists to explore this substance further..." To be sure, a number of her images do come from the reading, but these were available to the rest of the class as well. Not uniquely, but unusually, this student used the concepts that had been presented, and integrated them into her own work. But what bears repeating so that it is not lost is the growth that is indicated in comparison with the concluding sentence presented in (H2). There, only a few weeks earlier, science was heroic; here science has become more mundane in the sense that it has a procedure that enables its practitioners to deal with the stuff of the world. We see this development, and the way in which it is based on and grows out of descriptive evidence, as an exceptionally clear example of the way in which personal change can be captured.

(K4) In this paper the student again shows the wide range of her interests and knowledge, and her superior writing skills in the introduction. She starts by quoting a lengthy description of an ancient Greek

"water organ," and continues: "But, what does this have to do with the study and development of pneumatic devices? Perhaps nothing at all, except that it pleases me to know that one of the earliest applications of this aspect of the hard sciences (pneumatics) reflects the meaning of its Greek prefix: 'pneuma' – the soul or vital spirit. If music gives voice to that 'vital spirit,' then the fusion of science and aesthetics may have occurred far earlier than the 18th century." But even with all her verbal gifts going for her, she concludes, "In answer to part 'C' for the requirements of this paper, the use of the analogy of a spring to understand the behavior of gases and vacuums is an appropriate one. A spring action results when air is compressed or expanded." Here, again, we have the problem of how we should interpret a "wrong" – or incomplete – answer. Is it due to lack of understanding of the more quantitative material? Or just being turned off, or not being turned on, to or by the matter at hand? Or might it more simply be the result of having missed the lab sessions in which we worked with springs precisely so that we could compare their detailed behavior (Hooke's Law: F=-kx) with the behavior of an enclosed sample of a gas (Boyle's Law: PV=const.). Mathematically and graphically the first of these is linear, and the second is not. Our lab conclusion was that Boyle's use of this word was appropriate as an analogy, but not beyond that. We hoped that the evidence which we had developed would be used to justify an answer. This student, again, chose not to incorporate evidence and related inferences into her discussion.

Student (B) demonstrates a different approach. At the end of her paper on this topic, which she titled "From Voids to Vapors," she writes, "In conclusion, I think it is easy to see that, in a remarkably short period of time, the study of gases traveled quite a distance from Torricelli's theory of 'sea of air' to Boyle's discovery of a quantitative law. Understanding nature in order to advance craft-based phenomena required 17th century scientists to improve experimental techniques and record their data systematically. Experimental approaches ranged from metal pipes to glass tubes, exploding spheres to controlled instruments, sea level to mountain tops, outdoors to indoors, and most importantly, from voids to vapors. Clearly, the accomplishments of Torricelli, Pascal, von Guericke, and Boyle established the foundation for the modern study of gases by confirming air as a substance."

A last sample (J4) shows yet another way in which reflection can be pursued. In this case, the writer is self conscious about what she has done, even as she is actively engaged in comparisons. Her use of other sources is evident and acknowledged, although it was not required. Nevertheless, it is most likely a skill she has mastered, and in context it is appropriate. "I have compared the scientific work of Otto von Guericke to that of Robert Boyle. Actually, I think Robert Boyle's contributions to

science can, in many ways be more aptly compared to those of Galileo. Both of these men were indefatigable experimenters who were interested in many fields of science. Galileo contributed much more new knowledge to science, particularly with his study of motion. Additionally, Galileo and Boyle played similar roles in discrediting currently accepted erroneous scientific theory. Galileo, with his Dialogue on the Two Principal World Systems, exposed the weakness of Ptolemaic and Aristotelian astronomic theory. Boyle, with his Sceptical Chemist, also in the form of a dialogue, showed the fallacies of the Aristotelian 'four basic elements' theory and of Paracelsus' 'sulfur, mercury and salt' theory ['Landmarks in Science,' Robert B. Downs]. In the same way that Galileo prepared the way for Newton, so Boyle prepared the way for Lavoisier and Dalton, among others. With his definition of an element as a substance which cannot be further broken down or decomposed, he helped lay the foundations for one new science – physics. Boyle also established the importance of another new field of science based on the analysis of substances – chemistry. Indeed, he coined the word 'chemist.' The work of Galileo and Newton had revolutionized our view of the world. Boyle's work, along with that of Torricelli, Pascal and von Guericke, was also instrumental in bringing about a transformation in our understanding of the world."

Journals

We have noted in some of the writing samples above that a major goal of this course (which does not factor into our grading scheme) is for our students to transfer their understandings of principal course concepts to personal experiences and reflections. As a specific means of examining this, besides the comments students make in their essays, we asked students in two lab sections to keep journals. They were directed to make weekly entries which in some way related fundamental concepts developed in the lab to phenomena, experiences, or reflections outside the lab. This exercise was further intended to help the students generalize concepts. To facilitate reflective responses, journal entries were neither graded nor criticized for content, style or grammar; however, written responses were entered next to the students' entries to reinforce their efforts, to provide them with clarification and relevant examples, and to complete their ideas and generalize from them. These responses to their individual musings, as well as the personal challenge of finding appropriate topics, seem to have provided enough motivation for them to continue, as seen by their nearly impeccable regularity.

The journals, like other modes of writing, reveal misconceptions where we anticipate them. But more significantly, the freer, more personal format for expression afforded students opportunities to discuss areas where they were aware of cognitive dissonance. These discussions were

especially instructive to us in enabling us to amend instruction, and for insights they provided into our students' constructions of knowledge.

One concept that troubles students on their way to understanding theories of evolution is the idea of homology – characteristics uniquely shared by groups of organisms due to their presence in the last common ancestor of those groups. As phrased, this concept is based upon, as well as supports, the notion of descent with variation from a common ancestor, which is often visually summarized with a "family tree." The following entry shows that a student does not yet grasp that there are different levels of homology. In other words, she is not yet nesting or superordinating groups hierarchically, as is done in the Linnean system of classification.

"The idea that we (humans) and animals have derived from one common ancestor is hard to accept for me. If that is the case, why don't we have birds with pouches, fish with utters, or cats with antlers? [This student might have been as much at home with a medieval bestiary, which depicted the creatures she describes, as with a visit to a zoo.]

"Did the common ancestor have all these characteristics? Where did the common ancestor originate from? I can understand how animals can adapt to certain conditions, but is adaptation the only reason for the change in certain species? Perhaps these questions will be answered by semester's end."

Five weeks later the same student is much more at ease with concepts of evolution, and though she leaves much unsaid, there is an implicit understanding not only of homology, but of the nature of paradigm shifts in science. On her own initiative, and perhaps searching for journal material, she has read a popular science magazine and comments:

"I was fascinated by the controversy about the origin of birds. I can't help but sympathize with Sankar Chatterjee. It seems that history keeps repeating itself. New controversial hypotheses are always bombarded if they don't serve the purpose of reinforcing the main stream idea. The same happened to Copernicus, Galileo, Lamarck, Mendel, etc. I wonder if Chatterjee is really correct about the origins of birds. It would mean quite a big twist in the evolutionary theory for some. But, does it matter when species branch out? Does everything have a time table? Why can't they accept the fact, or thought, that an ancient bird could come from a reptile versus a dinosaur??"

While she sympathizes with the underdog (Chatterjee), she still cannot judge the evidence herself – as so few can. Nevertheless, she does appear to see that the evidence (one very crushed partial skeleton) may be plausible, though not strong, for the reptilian vs. dinosaurian origin of birds. She now seems to have grasped different levels of homology and different times and events of origin.

Another student considers characteristics of kangaroos and rabbits to answer for herself which are unique to each group and which are analogs, and whether kangaroos and rabbits share a recent common ancestor. Her familiarity with the animals is superficial, yet she partitions their attributes in order to consider evidence for homology and analogy. She does not discuss the evidence, and likely is not prepared to do so, though she may be prepared to read a discussion. Her writing is in outline form.

"Kangaroo Homologies:
Pouched (for carrying young)
Fur
Mammaries
"Analogous Traits of Kangaroos & Rabbits:"
Large hind legs and paws for jumping
Erect quadrupedal posture.
"Non Analogical Traits of Kangaroos & Rabbits:"
Rabbits born with fur (kangaroos are not)
Length, structure & function of tails are different
<Too many differences>
"Are kangaroos in own class because of pouch?"

Her conclusion is correct, that is, kangaroos and rabbits are not closely related, but the issue is not a particularly challenging one, and her use of evidence is spotty. Nevertheless, on her own she has constructed both the problem or hypothesis, and the tests.

One last example of transferring knowledge from the classroom deals with the interaction of ionic salts and polar molecules. As one experiment in the laboratory, salt was applied to samples of gelatin, and it was observed that the salt "pulls" water from the gelatin.

"I was trying to understand one of my childhood's greatest mysteries: why snails (the ones without a shell) melt away when salt is poured on them. My mom killed them that way. I don't think it was such a nice [thing] to get rid of them, but it worked.

"After our latest lab experiment ... I could identify the snail with the gelatin. When salt was poured on top of it, the salt "pulled" the water out of the gelatin. Whatever snails are composed of, it must be very polar because the way the salt diminishes it. My snail research is not over yet. This is just food for thought."

In sum, because it is non-directive, journal writing is less effective than other assignments in allowing the instructor to finger a student's abilities to select and use evidence. But also because it is non-directive, journal writing compels students to find personal relevance where they may otherwise not, and this often happens in apparently unlikely places.

Students frequently remarked that writing in their journals was the most difficult requirement of the course (which is also the most difficult and demanding of their courses), and in fact they spent days thinking over the one or two paragraphs they were to write. This was their part of the course, and these comments reflect a new-found pride in the possession of knowledge. They also claimed that they had never been asked to actually apply knowledge from one context to another, and especially not scientific knowledge.

Conclusions

In a recent discussion of science concepts among adolescents, adults and experts, Lewis and Linn (1994) note that in contrast to adolescents and adults, experts typically see the world more "holistically," integrating their formal knowledge with their intuitions and everyday experience. We believe that we have seen such an effect in the writing of a number of our students. This has been referred to above with regard to the Journals (which were collected during the second semester of the one year course), and also in other aspects of the second semester, as when they discuss the role of polarity in being able to account for the beading of water on a car's windshield, or in doing laundry. We believe that without writing as a vehicle for self expression, we would neither have encouraged, nor have been able to see, this aspect of intellectual growth.

The collective view of the many examples of student writing presented and discussed above all serve to demonstrate one theme. Beyond correctness and completeness of content, and beyond demonstration of "writing ability," student writing provides an invaluable window on cognitive position. Different modes of writing do this in different ways, and the way that is preferred in any given context should probably be determined with a clear idea of how this information will be used. Most simply, as opposed to errors of fact, students reveal the depths of their misconceptions when they are given the opportunity to express themselves, and these can be addressed on a one-to-one basis in comments on papers. Where students also show that their approach is out of touch with that of a course, this too can be addressed on a personal basis, either in writing, or possibly more easily in discussion. Finally, when one sees the spectrum of a class's response to certain questions it is possible to address particular issues in a number of ways. It may be necessary to create text material for the class covering particularly thorny issues, open letters to the class may be written which will give the students something to mull over, or class presentations may be adjusted to reflect the nature of the class, as opposed to the knowledge of the instructor. Any of these pedagogical responses, however, requires that the instructor first have a valid assessment of the character of a student or a class, and we believe that this can

be achieved by assigning appropriate writing exercises, and reading them attentively.

Acknowledgments

We wish to thank our many students, those whose writing samples have found their way into this paper and all the others, who continually challenge us to understand what they are thinking, and why. We also want to thank Karen Kitchener, Tony Lawson and Bill Moore, whose work and whose comments have helped us to find what understanding we have in this area.

References

Hays, Janice N., and Brandt, Kathleen S. (1992). Socio-cognitive development and students' performance on audience-centered argumentative writing. In Marie Secor and David Charney (Eds.), *Constructing Rhetorical Education.* Carbondale, Illinois: Univ. S. Illinois Press.

Herron, J. D. (1975). Piaget for chemists. *J. Chem. Ed. 52*:146-150.

Hunt, K. W. (1965). *Grammatical Studies Written at Three Grade Levels*. Nat. Council of Teachers of Eng., Res. Rep. No. 3

King, Patricia M., and Kitchener, Karen Strohm. (1994). *Developing reflective judgment: understanding and promoting intellectual growth and critical thinking in adolescents and adults.* San Francisco: Jossey-Bass Publishers.

Kitchener, Karen S. and King, Patricia M. (1990a). The reflective judgment model: Transforming assumptions about knowing. In J. Mezirow (Ed.), *Fostering critical reflection in adulthood: A guide to transformative and emancipatory learning* (pp. 159-176). San Francisco: Jossey Bass

Kitchener, Karen Strohm and King, Patricia M. (1990b). The reflective judgment model, ten years of research. In M.L. Commons, C. Armon, L. Kohlberg, F.A. Richards, T.A. Grozer, & J.D. Sinott (Eds.), *Adult Development: Vol. 2. Models and methods in the study of adolescent and adult thought* (pp. 63-78). New York: Praeger.

Kleinsasser, Audrey M., Collins, Norma Decker, and Nelson, Jane. (1994). Writing in the disciplines: Teacher as gatekeeper and as border crosser. *JGE: The Journal of General Education 43*(2), 117-133.

Lawson, A. E., (1985). A review of research on formal reasoning and science teaching. *J. Res. Science Teaching 22*, 569-617

Lawson, Anton E., and Shepherd, Gene D. (1979). Written language maturity and formal reasoning in male and female adolescents. *Language and Speech, 22*(2), 117-127.

Lewis, Eileen L., and Linn, Marcia C. (1994). Heat energy and temperature concepts of adolescents, adults, and experts: Implications for curricu-

lar improvements. *J. Res. Science Teaching, 31*(6), 657-677.
Moore, William S. (1991). The Perry scheme of intellectual and ethical development: An introduction to the model and major assessment approaches. Presented at the American Educational Research Association meeting, Chicago, April 3-7.
Perry, William G. (1970). *Forms of Intellectual and Ethical Development in the College Years: A Scheme*. New York: Holt, Rinehart and Winston.
Piaget, Jean (1959). *Judgment and Reasoning in the Child*. Paterson NJ: Littlefield, Adams & Co. (first published in 1928).
Piaget, J. (1972). Intellectual evolution from adolescence to adulthood. *Human Develop 15*, 1-12.
Shahn, Ezra (1988). On science literacy. *Educational Philosophy and Theory, 20*(2), 42-52.
Shahn, Ezra (1990). Foundations of science: A lab course for nonscience majors. In Don Emil Herget (Ed), *More History and Philosophy of Science in Science Teaching Proceedings of the First International Conference*. Tallahassee: Florida State University.

Contact Information:

Ezra Shahn's Home Page: http://sonhouse.hunter.cuny.edu/faculty/es/shahn.html

Ezra Shahn's Email: eshahn@hunter.cuny.edu