Paper presented at the IHPST 2009 Biennial Meeting (Notre Dame, IN)

A Study on Using the History of Industrial Melanism to Teach the Nature of Science

DAVID W. RUDGE[1] and ERIC M. HOWE[2]
[1] *Biological Sciences & The Mallinson Institute for Science Education, Western Michigan University (Email: david.rudge@wmich.edu)*
[2] *Department of Education, Assumption College (Email: emhowe@assumption.edu)*

Abstract: Rudge (2004) presented an innovative approach to using the history of research on industrial melanism to help students learn issues associated with the nature of science (NOS). The phenomenon of industrial melanism is presented to students as a "mystery phenomenon" in which they are provided with examples of observations that led past investigators to discover the phenomenon. Students are asked to consider why the phenomenon is occurring and predictably develop explanations that represent well-known alternative conceptions of evolutionary phenomena once proposed by past scientists. Students are then challenged to think through how they might test among these alternatives, with their predictable responses being used to motivate discussions of the results of similar experiments by past scientists. Throughout the three day unit, issues associated with the nature of science are addressed using an *explicit* and *reflective* approach (c.f. Abd-El-Khalick & Lederman 2000). Rudge et. al. (2007) presented the results of a pilot study (19 participants) aimed at evaluating the efficacy of this unit with reference to a targeted set of NOS issues including the nature of theories and experiments, theory change, how the results of experiments are interpreted, and what role imagination and creativity play in science. In the present paper we present the results of a more extensive study involving a total of 130 participants. The efficacy of the unit is once more assessed by means of open-ended surveys (VNOS) (Lederman Abd-El-Khalick, Bell & Schwartz 2002) and follow-up interviews. A collective comparison of pre- and post- responses of students to survey questions and interviews with 17 participants provides some evidence of improvement. The essay concludes with a discussion of several lessons learned regarding the usefulness of the VNOS instrument as a measure of students' understandings of issues associated with the nature of science.

Key words: evolution learning; evolution teaching; history of science; nature of science; philosophy of science

1. Introduction

National science standards emphasize that students should learn both *of* and *about* science (AAAS 1990, NRC 1996). While scholars recognize the importance of learning about the nature of science (NOS), many disagree regarding how best to promote this aim. In numerous publications, Michael Matthews has emphasized the necessity of using history and philosophy of science for learning about the nature of science (e.g. Matthews 1994). Norm Lederman, in contrast, has openly questioned the usefulness of using history of science to promote NOS at all (Lederman 1998). Lederman's concern is fundamentally empirical: there are relatively few empirical studies that have documented that use of history of science in the science classroom promotes NOS and indeed some that suggest otherwise (e.g. Abd-El-Khalick 1998; Abd-El-Khalick & Lederman 2000; but see Kim & Irving 2009).

Part of the difficulty establishing whether and to what extent history can help students learn about the nature of science is that naive conceptions about these issues are particularly tenacious and resistant to change (Abd-El-Khalick & Lederman 2000). To address this, Fouad Abd-El-Khalick and his associates have advocated an *explicit* and *reflective* pedagogical approach (c.f. Abd-El-Khalick & Lederman 2000) to the teaching of NOS (e.g. Akerson et. al. 2000, Khishfe & Abd-El Khalick 2002). By 'explicit' these authors mean that treatment of nature of science is a planned instructional activity involving direct examination and discussion by students, in contrast to implicit approaches in which the student is presumed to 'pick up' an understanding of the nature of science merely by learning the conceptual material or engaging in the process of science. By 'reflective' they draw attention to the importance of providing students with opportunities to reach their own insights into the nature of science issue as a result of their own deliberations rather than didactically telling them what to believe.

In this paper we present results of a study on a three class instructional unit developed with reference to the history of research on industrial melanism, which likewise uses Abd-El-Khalick's recommended explicit and reflective approach to the teaching of issues associated with the nature of science using the VNOS instrument (Lederman Abd-El-Khalick, Bell & Schwartz 2002). As with our previous studies, history is used *instrumentally* to place students in a similar (but not identical) problem solving context to that faced by past scientists (Howe 2004; Rudge 2007). This study was intended to address two questions:

(1) Did the Mystery Phenomenon Unit improve student understandings of a targeted set of issues associated with the nature of science? and,
(2) If so, what was it about the unit that led to these changes?

2. Method

2.1 PARTICIPANTS

The study was initiated under the auspices of a Human Subjects Institutional Review Board at a large Midwestern university during the Fall 2007 and Spring 2008 terms. It took place in the context of an introductory biology course (BIOS 1700 *Life Science* *for Elementary Educators I*) taken by future elementary school teachers. At the time of the study the course was taught in a lecture lab format, with students meeting once a week for a 2 hour 'lecture' session and twice a week for a 1½ hour small section (24 student) lab.[1] The lecture section was taught by the first author of the present paper; a total of 11 lab sections (6 during Fall 2007; 5 during Spring 2008) were taught by multiple graduate student instructors, each of whom was responsible for either one or two lab sections in a given term. The instructional sequence of interest was taught in lab over three successive sessions as part of a final unit devoted to evolutionary biology. All of the instructors of lab sections had previously taught the unit of interest at least twice.

The potential pool of participants was 221. The actual number of students who agreed to participate in the study and completed both pre and post surveys was 130, and among them 17 agreed to be interviewed. 93% of the participants self identified as Caucasian, 85% were female. Participants ranged in age from 18-30 years old; the mean age was 21. A small number of students (17%) indicated they had previously taken a philosophy course.

2.2 CONTEXT AND INTERVENTION

BIOS 1700 *Life Science for Elementary Educators I* is an introductory biology course, one of six science content courses taken at the university by elementary education majors. Each of these courses features an open-ended problem solving environment in which students are encouraged to take charge of their own learning. BIOS 1700 has four units devoted to taxonomy, anatomy and physiology, ecology and evolution. A companion course, BIOS 2700 *Life Science for Elementary Educators II*, contains three units devoted to more advanced topics (genetics, molecular and cell biology, and a capstone unit that invites students to study a single phenomenon from the standpoint of all of these subdisciplines.

The three day instructional sequence of interest in the present study is based upon the history of research on industrial melanism, presented to students as 'The Mystery Phenomenon'. (A brief review of this intervention is provided below, additional detail regarding how each class was conducted may be found in Rudge (2004).) During the first lab session, the instructor shares a series of observations made by naturalists during the mid to late nineteenth century that led them to notice a pattern in nature. Students are put in the position of identifying the pattern and coming up with their own ideas for why it is happening. The instructor asks students to explicitly discuss not only why they think the phenomenon is occurring, but also the reasons that suggest their ideas are plausible and worthy of further

---

[1] Lab sections were taught by inquiry (i.e. laboratory instructors taught primarily by means of carefully worded questions aimed at facilitating student driven discussions of and about biological topics and the process of science). Lecture sessions were devoted to practicing example problems of the sort that would appear on the final exam, with students attempting to solve problems on their own and with the help of the person seated next to them before the class as a whole discussed their answers.

consideration. In the next class students learn that the highly predictable ideas they came up with are similar to three distinct hypotheses put forth by past scientists who studied the phenomenon. Attention is drawn to the fact that in the case of the scientists, their hypotheses were developed in the context of theoretical frameworks, specifically Darwin's Theory of Natural Selection, Larmarck's Theory of the Inheritance of Acquired Characteristics and De Vries' Mutation Theory. The instructor then asks students to *explicitly* and *reflectively* (c.f. Abd-El-Khalick & Lederman 2000) discuss what theories are in general and how scientists choose among them using still other examples. Students discuss how the three proposed hypotheses for the mystery phenomenon might be tested by means of observation and experiment. Student responses are entirely predictable[2], and as such can be used to motivate discussions of the results of similar experiments by past scientists. Students are invited to consider the results from the perspective of both advocates and critics of each theory. Attention to the important role that observations and experiments play in testing hypotheses provides a natural opportunity for the instructor to lead an *explicit, reflective* discussion on what experiments are in science and whether they are always necessary. In the third and final class, students view a film (*Evolution in Progress*) that appears to conclusively demonstrate that the phenomenon should be understood in terms of natural selection. The instructor then reveals 'the rest of the story', additional details that draw attention to the fact that the mystery phenomenon is much more complicated than textbooks would have us believe. A concluding discussion asks participants (all of whom are future elementary school teachers) to consider how they will help their students recognize how very misleading textbook accounts can be with reference to both the process and nature of science.[3]

2.3 PROCEDURE[4]

Open-ended Surveys

A list of the questions used before and after the instructional sequence of interest is provided as Appendix A. (The actual surveys administered to students differed only in the inclusion of large amounts of space between questions to provide students with sufficient room for their responses.) The survey includes six separate questions, each question addressing a distinct issue associated with the nature of science that has been emphasized in recent reform documents (AAAS 1990, NRC 1996). Question 1 was developed in the context of a previous similar research project (Howe 2004). Questions 2-6 were taken from the Views of Nature of Science version C (VNOS-C), discussed in detail in Lederman et. al. (2002). (The design of the VNOS survey and the method for its validation are discussed in detail in Lederman & O'Malley 1990.)

Semi-structured Interviews

Lederman and O'Malley (1990) demonstrated that the validity of open-ended surveys is enhanced when used in conjunction with semi-structured interviews. Interviews allow the investigator to clarify students are interpreting questions as intended and conversely clarify the investigator is interpreting students responses as they intended. They are also potentially valuable in providing further insights into why a participant's views have changed, because during the course of the interview the interviewer can pointedly ask students to clarify what factors led them to change their answers.[2]

All 130 students who completed both the pre- and post- surveys were invited to participate in half hour interviews at the conclusion of the unit for a small financial reimbursement ($20) in order to ensure that at least 13 would take part. Seventeen students ultimately agreed to be interviewed. Six were interviewed by the first author of the present paper, who had been trained in a similar research project by the second (Howe 2004); the remaining eleven were interviewed by a single graduate student. (The first author trained the graduate student interviewer by having her reading several relevant articles in the literature (e.g. Abd-El-Khalick & Ledermann, 2000; Palmquist & Finley, 1997; Ginsberg 1997), copies of student written responses to pre- and post- surveys, and mock interviews.)

During the interviews of the present study the interviewer asked each student to reread both the survey questions aloud and their responses. The interviewer asked follow up questions aimed at clarifying whether the student understood the original question and also the specific wording of the students' answers. Students were then invited to paraphrase their responses and, when change occurred, discuss what led them to change their answers.

2.4 DATA ANALYSIS

Data was analyzed in a manner similar to established protocols in similar research into student conceptions of the nature of science (Abd-El-Khalick 2001, Abd-El-Khalick et. al. 1998; Akerson et. al. 2000, Howe 2004, Palmquist & Finley 1997). The first author read through all the pre- and post- surveys to identify emergent themes amongst student responses to each of the questions considered in isolation. For each question he was able to identify 4-10 themes, which provided a basis for subsequent coding of student responses. The first author then considered each response once more with reference to the student's response to other questions in the hope this would clarify the student's intended answer to the question of interest.

The reliability of this coding of data was assessed by having the second author independently categorize 23.1% of the completed surveys from all 130 participants (randomly chosen from both pre- and post- surveys) using the first author's coding scheme (60 surveys total). The inter-rater reliability, or degree to which the independent coding of the second author agreed with that of the first author was 70%.

Survey data was used primarily to identify and track the frequency of nature of science conceptions held by the participants, some of which can be identified as representing naïve and more sophisticated views of the nature of science. Aggregate comparisons of pre- and post- survey responses (discussed below) provide evidence of change within the study population as a whole.

---

[2] Interviews were only conducted post instruction owing to the brevity of the unit (3 lab classes held over the course of two weeks).

3. Results

3.1 SURVEY DATA

Data collected pre- and post- instruction from the one hundred thirty participants who completed both surveys are summarized in Table 1. Taken as a whole, students' pre-instructional conceptions of targeted issues associated with the nature of science were relatively naïve. This is consistent with previous findings by Abd-El-Khalick (1998) and Howe (2004). A collective comparisons of pre- and post-responses indicates some change occurred as a result of the unit.[5]

Question 1 invited participants to share their understandings of what theories are, how they are created, and provide an example of when they have used a theory. Prior to instruction, most student responses (71.5%) revealed relatively naïve views about the nature of theories, with most identifying a theory as either an unproven claim or an educated guess. A relatively small percentage (28.5%) held a more sophisticated view of theories as explanatory frameworks that provide a basis for prediction. By the end of the instructional sequence this number had increased (38.5%), as did the number of those that were able to provide an example that demonstrated their claim (63.9% vs. 43.1%). Seventeen students used examples drawn from the Mystery Phenomenon Unit in their post survey responses. These results are comparable to those found in Howe's previous 2004 study, which documented a similar increase in understanding of the explanatory role of theories among student responses post instruction (from 50% (n= 42) to 70% (n = 57) [N =81]).

As an additional step in the analysis, the pre- and post- instruction responses of each participant were compared to one another. With regard to the question of what theories are (Question 1a), many exhibited no appreciable change (43.8%), but those who shifted from less to more sophisticated points of view (33.8%) exceeded those whose views shifted in the opposite direction (22.3%). A similar shift was also observed among responses to the request for an example (Question 1c). Most participants (56.2%) exhibited no change in their responses pre and post instruction; but over twice as many appeared to have improved compared to those whose views declined (31.5% vs. 12.3%)

Question 2 followed up on the first by asking students to consider whether theories ever change, and if so, why. A naïve conception often held by students is that scientific knowledge once discovered never changes, it is either wholly accepted or wholly abandoned (Cotham & Smith 1981). A more sophisticated understanding includes recognition that not only do theories change over time (as the result of new findings or technology), but that such changes may result from the reinterpretation of existing data. Prior to instruction, most of the student responses (91.5%) indicated that the student recognized that theories do change, most identifying this as a consequence of new evidence (73.8%). Only a small fraction (7.8%) recognized theories might change as a result of the reinterpretation of data.

Table 1. Change in students' nature of science conceptions (N = 130)

<table>
<tr><td><i>Theme</i></td><td colspan="5"><i>Description</i></td><td colspan="2"><i>Pre Instruction</i></td><td><i>Post Instruction</i></td><td><i>How Sophisticated</i></td></tr>
<tr><td colspan="10">Question 1a: What is a theory?</td></tr>
<tr><td>1</td><td colspan="5">Explanation</td><td colspan="2">28.5% (n=37)</td><td>38.5% (n=50)</td><td>Most</td></tr>
<tr><td>2</td><td colspan="5">Claim</td><td colspan="2">33.8% (n=44)</td><td>39.2% (n=51)</td><td></td></tr>
<tr><td>3</td><td colspan="5">Hypothesis/Guess</td><td colspan="2">24.6% (n=32)</td><td>11.5% (n=15)</td><td></td></tr>
<tr><td>4</td><td colspan="5">Nonsensical/unclear/no answer</td><td colspan="2">13.1% (n=17)</td><td>10.8% (n=14)</td><td>Least</td></tr>
<tr><td colspan="9">Change Pre->Post Instruction</td><td>% Change</td></tr>
<tr><td colspan="9"><i>Improvement</i></td><td><i>33.8% (n=44)</i></td></tr>
<tr><td>2->1</td><td>3->1</td><td>4->1</td><td>3->2</td><td>4->2</td><td>4->3</td><td></td><td></td><td></td><td></td></tr>
<tr><td>10.0% (n=13)</td><td>9.2% (n=12)</td><td>3.1% (n=4)</td><td>4.6% (n=6)</td><td>5.4% (n=7)</td><td>1.5% (n=2)</td><td></td><td></td><td></td><td></td></tr>
<tr><td colspan="9"><i>No Change</i></td><td><i>43.8% (n=57)</i></td></tr>
<tr><td>1->1</td><td>2->2</td><td>3->3</td><td>4->4</td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>16.2% (n=21)</td><td>18.5% (n=24)</td><td>6.2% (n=8)</td><td>3.1% (n=4)</td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td colspan="9"><i>Backsliding</i></td><td><i>22.3% (n=29)</i></td></tr>
<tr><td>1->2</td><td>1->3</td><td>1->4</td><td>2->3</td><td>2->4</td><td>3->4</td><td></td><td></td><td></td><td></td></tr>
<tr><td>10.8% (n=14)</td><td>1.5% (n=2)</td><td>0.0% (n=0)</td><td>2.3% (n=3)</td><td>3.1% (n=4)</td><td>4.6% (n=6)</td><td></td><td></td><td></td><td></td></tr>
<tr><td colspan="10">Question 1b: How are theories created?</td></tr>
<tr><td>1</td><td colspan="5">Reflecting on prior knowledge</td><td colspan="2">21.5% (n=28)</td><td>14.6% (n=19)</td><td>Most</td></tr>
<tr><td>2</td><td colspan="5">Direct result of empirical research</td><td colspan="2">57.7% (n=75)</td><td>63.1% (n=82)</td><td></td></tr>
<tr><td>3</td><td colspan="5">Coming up with an idea</td><td colspan="2">14.6% (n=19)</td><td>13.1% (n=17)</td><td></td></tr>
<tr><td>4</td><td colspan="5">Nonsensical/unclear/no answer</td><td colspan="2">6.2% (n=8)</td><td>9.2% (n=12)</td><td>Least</td></tr>
<tr><td colspan="9">Change Pre->Post Instruction</td><td>% Change</td></tr>
<tr><td colspan="9"><i>Improvement</i></td><td><i>20.0% (n=26)</i></td></tr>
<tr><td>2->1</td><td>3->1</td><td>4->1</td><td>3->2</td><td>4->2</td><td>4->3</td><td></td><td></td><td></td><td></td></tr>
<tr><td>6.2%(n=8)</td><td>2.3% (n=3)</td><td>0.0%(n=0)</td><td>6.9%(n=9)</td><td>3.8% (n=5)</td><td>0.8% (n=1)</td><td></td><td></td><td></td><td></td></tr>
<tr><td colspan="9"><i>No Change</i></td><td><i>52.3% (n=68)</i></td></tr>
<tr><td>1->1</td><td>2->2</td><td>3->3</td><td>4->4</td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>6.2%(n=8)</td><td>40.8%(n=53)</td><td>3.8% (n=5)</td><td>1.5% (n=2)</td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td colspan="9"><i>Backsliding</i></td><td><i>27.7% (n=36)</i></td></tr>
<tr><td>1->2</td><td>1->3</td><td>1->4</td><td>2->3</td><td>2->4</td><td>3->4</td><td></td><td></td><td></td><td></td></tr>
<tr><td>11.5%(n=15)</td><td>3.1%(n=4)</td><td>0.8% (n=1)</td><td>5.4% (n=7)</td><td>5.4% (n=7)</td><td>1.5% (n=2)</td><td></td><td></td><td></td><td></td></tr>
</table>

Table 1. Change in students' nature of science conceptions (con't)

Theme	*Description*	*Pre Instruction*	*Post Instruction*	*How Sophisticated*
Question 1c: Give an example when you have created or used a theory.				
1	Reflective use of example from MPU	0.8% (n=1)	13.1% (n=17)	Most
2	Reflective use of other scientific example	42.3% (n=55)	50.8% (n=66)	
3	Vague reference to example from MPU	0% (n=0)	0.8% (n=1)	
4	Vague reference to other scientific example	14.6% (n=19)	9.2% (n=12)	
5	Invalid reference/non-scientific example	26.9% (n=35)	14.6% (n=19)	Least
6	Nonsense/no example	15.4% (n=20)	11.5% (n=15)	

Change Pre->Post Instruction									**% Change**
Improvement									*31.5% (n=41)*
4->1	5->1	6->1	4->2	5->2	6->2	5->4	6->4		
1.5% (n=2)	2.3% (n=3)	1.5% (n=2)	4.6%(n=6)	11.5%(n=15)	7.7%(n=10)	0.8% (n=1)	1.5% (n=2)		
No Change									*56.2% (n=73)*
1->1	2->2	4->4	5->5	6->6	2->1	5->6	6->5		
0.8% (n=1)	26.9% (n=35)	4.6%(n=6)	10.0%(n=13)	3.8%(n=5)	6.9%(n=9)	2.3% (n=3)	0.8% (n=1)		
Backsliding									*12.3% (n=16)*
2->3	2->4	2->5	2->6	4->5	4->6				
0.8% (n=1)	2.3% (n=3)	3.1%(n=4)	2.3% (n=3)	0.8% (n=1)	3.1%(n=4)				

Theme	*Description*	*Pre Instruction*	*Post Instruction*	*How Sophisticated*
Question 2a: Explain why you think that scientific theories do (or do not) change.				
1	Yes, if the theory is reinterpreted	7.7% (n=10)	6.2% (n=8)	Most
2	Yes, as a result of new evidence	73.8% (n=96)	80.8% (n=105)	
3	Yes, everything is subject to change	10% (n=13)	5.4% (n=7)	
4	No, theories are discarded/replaced	7.7% (n=10)	7.7% (n=10)	Least
5	Inconsistent (yes and no)	0.8% (n=1)	0% (n=0)	

Change Pre->Post Instruction									**% Change**
Improvement									*18.5% (n=24)*
2->1	3->1	4->1	3->2	5->2					
3.8%(n=5)	1.5% (n=2)	0.8% (n=1)	5.4%(n=7)	6.9%(n=9)					
No Change									
2->2	3->3	5->5	5->3						
63.1%(n=82)	0.8% (n=1)	1.5% (n=2)	1.5% (n=2)						
Backsliding									*14.6% (n=19)*
1->2	1->3	1->5	2->3	2->5					
5.4%(n=7)	0.8% (n=1)	1.5% (n=2)	4.6%(n=6)	2.3(n=3)					

Table 1. Change in students' nature of science conceptions (con't)

Theme	*Description*	*Pre Instruction*	*Post Instruction*	*How Sophisticated*
Question 2b: Defend your answer with examples.				
1	Reflective use of example from MPU	0% (n=0)	13.1% (n=17)	Most
2	Reflective use of other scientific example	53.8% (n=70)	39.2% (n=51)	
3	Vague reference to example from MPU	0% (n=0)	1.5% (n=2)	
4	Vague reference to other scientific example	7.7% (n=10)	6.2% (n=8)	
5	Invalid reference/non-scientific example	3.8% (n=5)	1.5% (n=2)	Least
6	Nonsense/no example	34.6% (n=45)	38.5% (n=50)	

Change Pre->Post Instruction									**% Change**
Improvement									*18.5% (n=24)*
4->1	5->1	6->1	4->2	5->2	6->2	6->3	6->4		
0.8% (n=1)	0.8% (n=1)	4.6% (n=6)	2.3% (n=3)	0.8% (n=1)	7.7% (n=10)	0.8% (n=1)	0.8% (n=1)		
No Change									*60.8% (n=79)*
2->2	4->4	6->6	2->1	5->6					
28.5% (n=37)	2.3% (n=3)	20.8% (n=27)	6.9% (n=9)	2.3% (n=3)					
Backsliding									*20.8% (n=27)*
2->3	2->4	2->5	2->6	4->6					
0.8% (n=1)	3.1% (n=4)	1.5% (n=2)	13.1% (n=17)	2.3% (n=3)					

Theme	*Description*	*Pre Instruction*	*Post Instruction*	*How Sophisticated*
Question 3a: What is an experiment? (What is the point of an experiment?)				
1	Test of a hypothesis	36.1% (n=47)	23.8% (n=31)	Most
2	Test of a hypothesis or theory	10.0% (n=13)	10.8% (n=14)	
3	Test of a theory	16.1% (n=21)	26.2% (n=34)	
4	Test (unspecified)	14.6% (n=19)	19.2% (n=25)	
5	Nonsensical/unclear/no answer	23.1% (n=30)	20.0% (n=26)	Least

Change Pre->Post Instruction										**% Change**
Improvement										*23.8% (n=31)*
2->1	3->1	4->1	5->1	5->2	5->3	5->4				
0.8% (n=1)	3.1% (n=4)	4.6% (n=6)	3.1% (n=4)	3.1% (n=4)	3.8% (n=5)	5.4% (n=7)				
No Change										*46.2% (n=60)*
1->1	2->2	3->3	4->4	5->5	2->3	2->4	3->2	3->4	4->2	4->3
12.3% (n=16)	0.8% (n=1)	8.5% (n=11)	3.8% (n=5)	7.7% (n=10)	3.8% (n=5)	1.5% (n=2)	1.5% (n=2)	2.3% (n=3)	0.8% (n=1)	3.1%(n=4)
Backsliding										*30.0% (n=39)*
1->2	1->3	1->4	1->5	2->5	3->5	4->5				
4.6% (n=6)	6.9% (n=9)	6.2% (n=8)	6.2% (n=8)	3.1% (n=4)	0.8% (n=1)	2.3% (n=3)				

Table 1. Change in students' nature of science conceptions (con't)

Theme	*Description*	*Pre Instruction*	*Post Instruction*	*How Sophisticated*
Question 3b: What is an experiment? (What does the conduct of an experiment involve?)				
1	Manipulating system and observing change	28.5% (n=37)	43.1% (n=56)	Most
2	Collecting evidence	22.3% (n=29)	19.2% (n=25)	
3	Use of the scientific method	0.8% (n=1)	0.0% (n=0)	
4	Nonsensical/unclear/no answer	48.5% (n=63)	37.7% (n=49)	Least

Change Pre->Post Instruction									**% Change**
Improvement									*36.2% (n=47)*
2->1	3->1	4->1	4->2						
10.0% (n=13)	0.8% (n=1)	14.6% (n=19)	10.8% (n=14)						
No Change									*45.4% (n=59)*
1->1	2->2	4->4							
17.7% (n=23)	4.6% (n=6)	23.1% (n=30)							
Backsliding									*18.5% (n=24)*
1->2	1->4	2->4							
3.8% (n=5)	6.9% (n=9)	7.7% (n=10)							

Theme	*Description*	*Pre Instruction*	*Post Instruction*	*How Sophisticated*
Question 4a: Does the development of scientific knowledge REQUIRE experiments?				
1	No, science can develop by other means	13.1% (n=17)	16.2% (n=21)	Most
2	No, experiments not always appropriate	0.8% (n=1)	9.2% (n=12)	
3	Yes, only experiments can prove	19.2% (n=25)	23.1% (n=30)	
4	Yes, data collection is necessary	47.7% (n=62)	39.2% (n=51)	
5	Yes (nonsense/no reason)	0.8% (n=1)	3.8% (n=5)	Least
6	Unclear OR no answer	18.5% (n=24)	8.5% (n=11)	

Change Pre->Post Instruction										**% Change**
Improvement										*36.2% (n=47)*
3->1	4->1	6->1	3->2	4->2	6->2	4->3	5->3	6->3	6->4	
2.3% (n=3)	4.6% (n=6)	2.3% (n=3)	3.1% (n=4)	3.1% (n=4)	2.3% (n=3)	9.2% (n=12)	0.8% (n=1)	2.3% (n=3)	6.2% (n=8)	
No Change										*48.5% (n=63)*
1->1	3->3	4->4	6->6	1->2						
6.9% (n=9)	10.0% (n=13)	25.4% (n=33)	5.4% (n=7)	0.8% (n=1)						
Backsliding										*15.4% (n=20)*
1->3	1->4	1->6	2->4	3->4	3->5	4->5	5->6			
0.8% (n=1)	3.8% (n=5)	0.8% (n=1)	0.8% (n=1)	3.1% (n=4)	0.8% (n=1)	3.1% (n=4)	2.3% (n=3)			

Table 1. Change in students' nature of science conceptions (con't)

<table>
<tr><th><i>Theme</i></th><th colspan="5"><i>Description</i></th><th><i>Pre Instruction</i></th><th colspan="2"><i>Post Instruction</i></th><th><i>How Sophisticated</i></th></tr>
<tr><td colspan="10">Question 4b: Defend your position with examples.</td></tr>
<tr><td>1</td><td colspan="5">Reflective use of example from MPU</td><td>0.0% (n=0)</td><td colspan="2">4.6% (n=6)</td><td rowspan="2">Most</td></tr>
<tr><td>2</td><td colspan="5">Reflective use of other scientific example</td><td>35.4% (n=46)</td><td colspan="2">20.0% (n=26)</td></tr>
<tr><td>3</td><td colspan="5">Vague reference to example from MPU</td><td>0.0% (n=0)</td><td colspan="2">0.8% (n=1)</td><td rowspan="2"></td></tr>
<tr><td>4</td><td colspan="5">Vague reference to other scientific example</td><td>21.5% (n=28)</td><td colspan="2">16.9% (n=22)</td></tr>
<tr><td>5</td><td colspan="5">Invalid reference/non-scientific example</td><td>0.8% (n=1)</td><td colspan="2">1.5% (n=2)</td><td rowspan="2">Least</td></tr>
<tr><td>6</td><td colspan="5">Nonsense/no example</td><td>42.3% (n=55)</td><td colspan="2">56.2% (n=73)</td></tr>
<tr><td colspan="9">Change Pre->Post Instruction</td><td>% Change</td></tr>
<tr><td colspan="9"><i>Improvement</i></td><td><i>15.4% (n=20)</i></td></tr>
<tr><td>4->1</td><td>6->1</td><td>4->2</td><td>6->2</td><td>6->4</td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>0.8% (n=1)</td><td>1.5% (n=2)</td><td>3.1% (n=4)</td><td>6.9% (n=9)</td><td>3.1% (n=4)</td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td colspan="9"><i>No Change</i></td><td><i>51.5% (n=67)</i></td></tr>
<tr><td>2->2</td><td>4->4</td><td>6->6</td><td>2->1</td><td>4->3</td><td>5->6</td><td>6->5</td><td></td><td></td><td></td></tr>
<tr><td>10.0% (n=13)</td><td>6.9% (n=9)</td><td>30.0% (n=39)</td><td>2.3% (n=3)</td><td>0.8% (n=1)</td><td>0.8% (n=1)</td><td>0.8% (n=1)</td><td></td><td></td><td></td></tr>
<tr><td colspan="9"><i>Backsliding</i></td><td><i>33.1% (n=43)</i></td></tr>
<tr><td>2->4</td><td>2->6</td><td>4->5</td><td>4->6</td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>6.9% (n=9)</td><td>16.2% (n=21)</td><td>0.8% (n=1)</td><td>9.2% (n=12)</td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td colspan="10">Question 5: How are DIFFERENT CONCLUSIONS possible from SAME SET OF DATA?</td></tr>
<tr><td>1</td><td colspan="5">Differences in interpretation</td><td>53.8% (n=70)</td><td colspan="2">60.8% (n=79)</td><td>Most</td></tr>
<tr><td>2</td><td colspan="5">Data is insufficient to decide between them</td><td>32.3% (n=42)</td><td colspan="2">32.3% (n=42)</td><td rowspan="2"></td></tr>
<tr><td>3</td><td colspan="5">Both might be true</td><td>6.1% (n=8)</td><td colspan="2">2.3% (n=3)</td></tr>
<tr><td>4</td><td colspan="5">Nonsensical/unclear/no answer</td><td>7.7% (n=10)</td><td colspan="2">4.6% (n=6)</td><td>Least</td></tr>
<tr><td colspan="9">Change Pre->Post Instruction</td><td>% Change</td></tr>
<tr><td colspan="9"><i>Improvement</i></td><td><i>23.1% (n=30)</i></td></tr>
<tr><td>2->1</td><td>3->1</td><td>4->1</td><td>4->2</td><td>4->3</td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>13.8% (n=18)</td><td>3.8% (n=5)</td><td>1.5% (n=2)</td><td>2.3% (n=3)</td><td>1.5% (n=2)</td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td colspan="9"><i>No Change</i></td><td><i>63.8% (n=83)</i></td></tr>
<tr><td>1->1</td><td>2->2</td><td>3->3</td><td>4->4</td><td>3->2</td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>41.5% (n=54)</td><td>16.9% (n=22)</td><td>0.8% (n=1)</td><td>2.3% (n=3)</td><td>2.3% (n=3)</td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td colspan="9"><i>Backsliding</i></td><td><i>13.1% (n=17)</i></td></tr>
<tr><td>1->2</td><td>1->4</td><td>2->4</td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>10.8% (n=14)</td><td>1.5% (n=2)</td><td>0.8% (n=1)</td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
</table>

Table 1. Change in students' nature of science conceptions (con't)

<table>
<tr><th>Theme</th><th colspan="6">Description</th><th>Pre Instruction</th><th>Post Instruction</th><th>How Sophisticated</th></tr>
<tr><td colspan="10">Question 6a: Do scientists use their creativity and imagination during their investigations?</td></tr>
<tr><td>1</td><td colspan="6">Yes, all stages; explanation</td><td>20.8% (n=27)</td><td>20.8% (n=27)</td><td>Most</td></tr>
<tr><td>2</td><td colspan="6">Yes, all stages; no explanation</td><td>2.3% (n=3)</td><td>6.2% (n=8)</td><td></td></tr>
<tr><td>3</td><td colspan="6">Yes, P&D, data collection; explanation</td><td>3.1% (n=4)</td><td>4.6% (n=6)</td><td></td></tr>
<tr><td>4</td><td colspan="6">Yes, P&D, data collection; no explanation</td><td>2.3% (n=3)</td><td>0.8% (n=1)</td><td></td></tr>
<tr><td>5</td><td colspan="6">Yes, P&D only; explanation</td><td>46.9% (n=61)</td><td>42.3% (n=55)</td><td></td></tr>
<tr><td>6</td><td colspan="6">Yes, P&D only; no explanation</td><td>4.6% (n=6)</td><td>3.1% (n=4)</td><td></td></tr>
<tr><td>7</td><td colspan="6">Yes, no specific stage; explanation</td><td>6.2% (n=8)</td><td>7.7% (n=10)</td><td></td></tr>
<tr><td>8</td><td colspan="6">Yes, no specific stage; no explanation</td><td>3.8% (n=5)</td><td>6.2% (n=8)</td><td></td></tr>
<tr><td>9</td><td colspan="6">No</td><td>8.5% (n=11)</td><td>6.2% (n=8)</td><td></td></tr>
<tr><td>10</td><td colspan="6">Nonsensical/unclear/no answer</td><td>1.5% (n=2)</td><td>2.3% (n=3)</td><td>Least</td></tr>
<tr><td colspan="9">Change Pre->Post Instruction</td><td>% Change</td></tr>
<tr><td colspan="9">Improvement</td><td>26.2% (n=34)</td></tr>
<tr><td>2->1</td><td>5->1</td><td>7->1</td><td>9->1</td><td>4->2</td><td>5->2</td><td>8->2</td><td>5->3</td><td>9->3</td><td></td></tr>
<tr><td>0.8% (n=1)</td><td>6.9% (n=9)</td><td>1.5% (n=2)</td><td>0.8% (n=1)</td><td>0.8% (n=1)</td><td>0.8% (n=1)</td><td>1.5% (n=2)</td><td>3.1% (n=4)</td><td>0.8% (n=1)</td><td></td></tr>
<tr><td>5->4</td><td>6->5</td><td>7->5</td><td>8->5</td><td>9->5</td><td>9->6</td><td>10->7</td><td>10->8</td><td></td><td></td></tr>
<tr><td>0.8% (n=1)</td><td>1.5% (n=2)</td><td>1.5% (n=2)</td><td>0.8% (n=1)</td><td>2.3% (n=3)</td><td>0.8% (n=1)</td><td>0.8% (n=1)</td><td>0.8% (n=1)</td><td></td><td></td></tr>
<tr><td colspan="9">No Change</td><td>46.2% (n=60)</td></tr>
<tr><td>1->1</td><td>5->5</td><td>6->6</td><td>7->7</td><td>8->8</td><td>9->9</td><td></td><td></td><td></td><td></td></tr>
<tr><td>10.8% (n=14)</td><td>28.5% (n=37)</td><td>0.8% (n=1)</td><td>1.5% (n=2)</td><td>0.8% (n=1)</td><td>3.8% (n=5)</td><td></td><td></td><td></td><td></td></tr>
<tr><td colspan="9">Backsliding</td><td>27.7% (n=36)</td></tr>
<tr><td>1->2</td><td>1->5</td><td>1->6</td><td>1->7</td><td>1->8</td><td>2->3</td><td>2->5</td><td>3->5</td><td>3->7</td><td>4->5</td></tr>
<tr><td>3.1% (n=4)</td><td>3.1% (n=4)</td><td>0.8% (n=1)</td><td>1.5% (n=2)</td><td>1.5% (n=2)</td><td>0.8% (n=1)</td><td>0.8% (n=1)</td><td>2.3% (n=3)</td><td>0.8% (n=1)</td><td>1.5% (n=2)</td></tr>
<tr><td>5->6</td><td>5->7</td><td>5->8</td><td>5->9</td><td>6->8</td><td>6->9</td><td>6->10</td><td>7->10</td><td>8->9</td><td></td></tr>
<tr><td>0.8% (n=1)</td><td>3.1% (n=4)</td><td>2.3% (n=3)</td><td>0.8% (n=1)</td><td>0.8% (n=1)</td><td>0.8% (n=1)</td><td>0.8% (n=1)</td><td>1.5% (n=2)</td><td>0.8% (n=1)</td><td></td></tr>
</table>

Table 1. Change in students' nature of science conceptions (con't)

<table>
<tr><th>Theme</th><th colspan="5">Description</th><th colspan="2">Pre Instruction</th><th>Post Instruction</th><th>How Sophisticated</th></tr>
<tr><td colspan="10">Question 6b: Defend your answer with examples.</td></tr>
<tr><td>1</td><td colspan="5">Reflective use of example from MPU</td><td colspan="2">0.0% (n=0)</td><td>8.5% (n=11)</td><td rowspan="2">Most</td></tr>
<tr><td>2</td><td colspan="5">Reflective use of other scientific example</td><td colspan="2">33.8% (n=44)</td><td>24.6% (n=32)</td></tr>
<tr><td>3</td><td colspan="5">Vague reference to example from MPU</td><td colspan="2">0.0% (n=0)</td><td>0.0% (n=0)</td><td rowspan="2"></td></tr>
<tr><td>4</td><td colspan="5">Vague reference to other scientific example</td><td colspan="2">2.3% (n=3)</td><td>2.3% (n=3)</td></tr>
<tr><td>5</td><td colspan="5">Invalid reference/non-scientific example</td><td colspan="2">1.5% (n=2)</td><td>0.8% (n=1)</td><td rowspan="2">Least</td></tr>
<tr><td>6</td><td colspan="5">Nonsense/no example</td><td colspan="2">62.3% (n=81)</td><td>63.8% (n=83)</td></tr>
<tr><td colspan="9">Change Pre->Post Instruction</td><td>% Change</td></tr>
<tr><td colspan="9">Improvement</td><td>18.5% (n=24)</td></tr>
<tr><td>5->1</td><td>6->1</td><td>4->2</td><td>5->2</td><td>6->2</td><td>6->4</td><td></td><td></td><td></td><td></td></tr>
<tr><td>0.8% (n=1)</td><td>6.2% (n=8)</td><td>0.8% (n=1)</td><td>0.8% (n=1)</td><td>9.2% (n=12)</td><td>0.8% (n=1)</td><td></td><td></td><td></td><td></td></tr>
<tr><td colspan="9">No Change</td><td>61.5% (n=80)</td></tr>
<tr><td>1->1</td><td>2->2</td><td>6->6</td><td>2->1</td><td>6->5</td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>0.0% (n=0)</td><td>13.8% (n=18)</td><td>45.4% (n=59)</td><td>1.5% (n=2)</td><td>0.8% (n=1)</td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td colspan="9">Backsliding</td><td>20.0% (n=26)</td></tr>
<tr><td>2->4</td><td>2->6</td><td>4->6</td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>1.5% (n=2)</td><td>16.9% (n=22)</td><td>1.5% (n=2)</td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
</table>

Table 2. Summary of change in students' nature of science conceptions (N = 130)

Question	*Improvement*	*No Change*	*Backsliding*
Question 1a. What is a theory?	33.8% (n=44)	43.8% (n=57)	22.3% (n=29)
Question 1b. How are theories created?	20.0% (n=26)	52.3% (n=68)	27.7% (n=36)
Question 1c. Give an example when you have created or used a theory.	31.5% (n=41)	56.2% (n=73)	12.3% (n=16)
Question 2a. Explain why you think that scientific theories do (or do not) change.	18.5% (n=24)	66.9% (n=87)	14.6% (n=19)
Question 2b. Defend your answer with examples.	18.5% (n=24)	60.8% (n=79)	20.8% (n=27)
Question 3a. What is an experiment? (What is the point of an experiment?)	23.8% (n=31)	46.2% (n=60)	30.0% (n=39)
Question 3b. What is an experiment? (What does the conduct of an experiment involve?)	36.2% (n=47)	45.4% (n=59)	18.5% (n=24)
Question 4a. Does the development of scientific knowledge REQUIRE experiments?	36.2% (n=47)	48.5% (n=63)	15.4% (n=20)
Question 4b. Defend your position with examples.	15.4% (n=20)	51.5% (n=67)	33.1% (n=43)
Question 5. How are DIFFERENT CONCLUSIONS possible from SAME SET OF DATA?	23.1% (n=30)	63.8% (n=83)	13.1% (n=17)
Question 6a. Do scientists use their creativity and imagination during their investigations?	26.2% (n=34)	46.2% (n=60)	27.7% (n=36)
Question 6b. Defend your answer with examples.	18.5% (n=24)	61.5% (n=80)	20.0% (n=26)

After instruction, a somewhat larger proportion of students agreed theories can change (92.4%), with a slightly smaller proportion (6.2%) indicating this might be the result of the reinterpretation of existing evidence. There was no change in the relative proportion of students who denied theories can change pre and post instruction (7.7%). Seventeen students specifically drew upon examples from the Mystery Phenomenon Unit in their post survey responses. These results contrast with those found in Howe's previous study, which documented a decline in student post instructional responses denying that theories can change (from 12% (n= 10) to 2% (n = 2) [N =81]) and a slight increase in student recognition that change could be a consequence of the reinterpretation of data (from 4% (n= 3) to 7% (n = 9) [N =81]).

Question 3 asked student to define what they think an experiment is. It was analyzed twice, first in terms of what the response said regarding what the point of an experiment is (3a), and second in terms of what the conduct of an experiment involves (3b).

With regard to the question of what the point of an experiment is, a naïve conception identifies experiments as tests. A more sophisticated understanding draws attention to the direct role experiments play in testing hypotheses, the indirect bearing of results of such tests on theories. Prior to instruction, 36.1% identified experiments are tests of hypotheses and 16.1% identified experiments as tests of theories. After instruction, 23.8% identified experiments as tests of hypotheses and 26.2% identified experiments as test of theories. With regard to the question of what does the conduct of an experiment involve, a naïve view identifies experiments with the collection of evidence. A more sophisticated understanding includes recognition that experiments in the context of biology involve perturbing the system involved and comparing results with a system not so perturbed (the experimental and control arms of the study, respectively). Prior to instruction 28.5% of the responses indicated recognition that experiments in biology are distinct from other forms of data collection in that they involve manipulation. After instruction, 43.1% of responses made this distinction. This shift was also evident in comparisons of individual responses pre and post: 36.2% demonstrated improvement in their understanding compared to 18.5% whose views as indicated in the surveys declined.[6] One student provided an unsolicited example from the Mystery Phenomenon Unit in response to this question:

> "In a controlled experiment (not natural environment) that involve running types of tests or trials in support of trying to figure something out. Example is the test (experiment) of Harrison's theory by giving certain leaves to moths." (Student 15, post survey)

Question 4 invited students to consider whether experiments were essential for the development of scientific knowledge. A scientifically literate person recognizes that scientific knowledge can change for other reasons (e.g. observations and the reinterpretation of existing data). Prior to instruction, while most of the participants claimed experiments were required for the development of scientific knowledge (67.7%), many responses identified experiments with data collection (47.7%). Only a relatively small proportion denied experiments were necessary

(13.9%). After instruction, a relatively higher proportion denied experiments were necessary (25.4%). (Comparisons of individual responses pre and post demonstrated a similar shift: 36.2% demonstrated improvement in their understanding compared to 15.4% whose views as indicated in the surveys declined.) These results are comparable to those found in Howe's previous study, which documented a similar decline in student post instructional responses denying that scientific knowledge can develop in the absence of experiments (from 52% (n= 22) to 17% (n = 8) [N =42][7]). His study also demonstrated an increase in student recognition that observational evidence alone can lead to the development of scientific knowledge (from 5% (n= 2) to 43% (n = 21) [N =42]).

Students were asked to defend their position on the necessity of experiments with an example. While six students specifically drew upon examples from the Mystery Phenomenon Unit in their post survey responses when asked for an example, the relative proportions of students able to provide a reflective example pre and post instruction declined (35.4% vs. 24.6%). This contrasts with the findings of Howe's previous study, in which an increase was observed (17% (n=7) vs. 36% (n=15) [N=42]).

Question 5 asked students to account for how it is possible that two groups of scientists looking at the same data could reach different conclusions. A naïve conception is one that attributes this to limits in available data, i.e. we don't know enough yet to decide between them. A more sophisticated conception includes recognition that the same data might be interpreted differently by virtue of differences in scientists' theoretical and experimental frameworks. Prior to instruction, a majority recognized the possibility that scientists might interpret the data differently (53.8%). After instruction a greater proportion of the responses provided some evidence that they recognized disagreements might be due to differences in how one interprets data (60.8%). These results are comparable to those found in Howe's previous study, which documented a similar increase in student post instructional responses mentioning the possibility that scientists could interpret the same data differently (from 54% (n= 44) to 64% (n = 59) [N =81]). Comparisons of individual responses pre and post demonstrated a similar shift: 23.1% demonstrated improvement in their understanding on this issue compared to 13.1% whose views as indicated in the surveys declined. Four students provided unsolicited examples from the Mystery Phenomenon Unit in response to this question:

> "Even if both groups have the same set of data, they are not going to have the same ideas of how the dinosaurs became extinct. For example, the mystery phenomon [sic] everyone in the class had different ideas of how the butterfly was dark colored." (Student 46, post survey)
>
> "Two different conclusions are possible because there can be different ways to interpret the same set of data. For example, a decrease in the number of dark moths could be due to lack of camouflage in a non-polluted forest, but it could also mean that they have a second predator in the environment. One set of data doesn't necessarily yeild [sic] one answer." (Student 55, post survey)

> "It is possible because neither group of scientist [sic] have discredited the other. And it is possible to have more than one theory for event it does not make either of them wrong example the three theories for the light moths and the theories for the silver box." (Student 95, post survey)

> "It's just like when we did the Mystery Phonomon [sic], we come up with 3 theories that fit the information and some we didn't look at as well. People's background & beliefs and education will affect a theory. Is one better than the other, it depends on the tests and expirments [sic] done to provide evidence to get to a conclusion to that answer." (Student 126, post survey)

Finally, Question 6 invited students to consider whether imagination and creativity play any role in scientific investigations. A naïve conception may acknowledge that imagination plays some role, but often portrays this as an aberration to be avoided. A more sophisticated understanding draws attention to the role of imagination and creativity in all stages of the investigation, from coming up with the design to figuring out how to collect data and interpret it. Prior to instruction, most recognized that imagination and creativity play some role in scientific investigations (86.2%), but most limited it to planning and design only (46.9%). After instruction, a slightly greater proportion indicated that imagination and creativity play at least some role in scientific investigations (91.5%).[8]

Only about a third of the pre and post responses included valid examples of creativity that gave some indication of reflectivity (33.8% vs. 33.1%). Eleven students offered examples taken from the Mystery Phenomenon Unit. Other responses to this question were difficult to interpret because many students either failed to share anything or gave such an abstract answer that it fell short of providing a specific example:

> "Scientist [sic] have to use creativity and imagination during investigations. By planning and designing theres creativity in the design and inner workings of the expirement [sic]. Once data is collected imagination must be relied on to determine what the numbers could possibly conclude. Other reasons for creativity is [sic] sometimes scientists cannot see anything they are hypothesizing about such things in space or the past. Therefore they rely on their own imagination to fill in the blanks." (Student 96, post survey)

Summary

The foregoing analysis of the survey data suggests that some changes amongst participants' understandings of a select set of issues associated with the nature of science did occur as a result of the instructional sequence (see Table 2).

As a group the students developed a better sense of the explanatory role theories play in science and more were able to come up with valid examples of theories. They also appeared to have a greater awareness of the distinction in biology between experiments and other forms of inquiry, as well as the fact it is possible for scientific knowledge to develop in the absence of experiments. Finally, they appear to have developed a greater sense of the role of interpretation in data analysis.

3.2 INTERVIEW DATA

The foregoing analysis of the aggregate survey results indicates the unit was more effective in addressing some nature of science issues than others. In this section we will explore the results of interviews with 17 participants for insights into what it was about the unit that either aided or detracted from student understanding of the targeted issues.

It is important to recognize that the interviews were conducted prior to the coding of survey responses, and as such, there was no way of knowing antecedently whether any of the volunteers actually exhibited the desired change. During the interviews each participant was directly asked to compare his/her pre and post instruction responses. For each question, interviewees were specifically asked whether they noticed a difference between their pre- and post responses, and if so, what led them to change their views. The analysis below will focus on responses to questions 1a, 3b, 4a, and 5 for which some improvement was noted in the aggregate population as a whole.

Question 1a. What is a theory?

Six of the seventeen students who agreed to be interviewed exhibited some improvement in their understanding of what theories are, as indicated by a comparison of how their pre- and post- instructional responses were coded (see examples in the first row of Table 3). Four of the six agreed their views had changed between the pre- and post- surveys; two of these four ascribed it vaguely to an outside class reading or something the interviewee "learned in class." It is striking that the two interviewed students who denied their views changed in response to the direct question nevertheless ascribed some activity in class as accounting for why their views had changed:

> "I: Okay. Good. And, um, on the pre-survey, you said, "Theories are answers or opinion or even beliefs about – that a person might have." And in the post-survey, you said, "Theory is someone's belief or understanding of an event that takes place in an environment." So, um, when you compare these two, does it seem to you that they're saying pretty much the same thing, or do you think that there might be a difference between the two?
> S: Uh… Pretty much the same thing.
> I: Mmm-hmm. Mmm-hmm. Okay. Um, in the one, you mention that you think it's an answer or an opinion and in the other, you mention that it's a belief or an understanding. Are those basically all the same things?
> S: No, as far as the answer, I believe that before I, the section, my impression was that the answer, whatever they believe, they naturally assumed, was the correct answer without conducting any type of experiments or observations.
> I: Mmm-hmm.
> S: But after the section that we reviewed and finding out that a different scientist actually had to go out and do observations and do experiments to get the correct answer, which changed my answer for the post answers." (Student 95 interview, 3->2)

Table 3. Representative quotations from interviewed students

Nature of Science Issue	*Less Sophisticated NOS Views*	*More Sophisticated NOS Views*
1a. What is a theory?	“A theory is something people believe to be true” (Student 129, pre-survey, 3)	“A theory is an explanation of why or how something happens for which there is evidence that suggest that the theory is correct but no concrete evidence that proves it.” (Student 129, post-survey, 1)
1b. How are theories created?	“Theories are created by any one or by anything with a valuable reason yet it has to make sense” (Student 92, pre-survey, 3)	“Theories are created from someone wanting to find out the answer to something they have been noticing or an observation” (Student 92, post-survey, 1)
1c. Give an example.	“We are using and studying math theories to look at areas of shapes” (Student 111, pre-survey, 5)	“Use theory of evolution to talk about changes in organisms” (Student 111, post-survey, 2)
2a. Do theories change? Why?	“Scientific theories do change that is why they are theories and not laws” (Student 112, pre-survey, 5)	“Yes, theories constantly change. A scientist will explain what they have found, then another scientist will prove that wrong or to be inaccurate” (Student 112, post-survey, 2)
2b. Give an example.	“The theory that the heart has a open circulatory system. It then was changed to a closed circulatory system.” (Student 112, pre-survey, 4)	“The theory that betularia [peppered moths] are changing from light to dark because of mutation. This would mean <u>every</u> betularia born would go through the mutation.” (Student 117, post survey, 1)
3a. What’s the point of an experiment?	“An experiment is a way to test a theory or a hypothesis” (Student 16, pre-survey, 2)	“An experiment is a test of a hypothesis to attept [sic] at proving or supporting a theory” (Student 16, post-survey, 1)
3b. What does an experiment involve?	“An experiment is when you test something out. You may have a theory or hypothesis and you can do an experiment to provide evidence for or against your original beliefs” (Student 129, pre-survey, 4)	“An experiment is when you are testing something, and observing to see what the outcomes are. Generally you have a control group and an experimental group” (Student 129, post-survey, 1)
4a. Does the development of scientific knowledge require experiments?	“Yes, because in order to prove some kind of scientific knowledge you have to test it to see if it is in fact a theory.” (Student 16, pre-survey, 4)	“Not always. Like the example you just gave us in class with scientist looking at if there was another planet outside of Uranus. The scientist made calculations and found Neptune. No experiment required.” (Student 16, post-survey, 4)

4b. Defend your answer with an example.	“[Y]ou can also gain scientific knowledge through observation.” (Student 129, pre-survey)	“Scientific knowledge can be gained through observation as well. There are not experiments to figure out what the composition of the earth is. Scientist just observe and take note of what they find” (Student 129, post-survey)
5. How are different conclusions possible from the same data?	“These different conclusions are possible because they can only use the evidence that is left behind and have no way of knowing which is correct. The two ideas may be closely related.” (Student 90, pre-survey, 2)	“Scientist may read the data differently and therefore come to different conclusions about what caused the extinctions.” (Student 90, post-survey, 1)
6a. Do scientists use creativity and imagination in their investigations?	“Scientist have to use their imaginations because they didn’t they would not know how to create the settings and supply what is needed to carry out the experiment.” (Student 95, pre-survey, 5)	“Yes because if they don’t use their imagination and creativity they can’t come up with the appropriate experiments or even a sufficient theory or way to collect the data.” (Student 95, post-survey, 3)
6b. Defend your answer with an example.	“It’s like cooking: you may prepare a baked chicken the same way everytime, but the next you might say, hey I have some fruit and I’m going to add that and see what happens” (Student 92, pre-survey, 5)	“For example, we learned that LaMarck [Harrison] cheated on his theory and therefore used his imagination and creativity to alter results” (Student 92, post-survey, 1)

> "I: Um, so if you look at your, um, if you look at your pre-survey and your post-survey, I noticed that you wrote a little bit more on the post-survey and it sounds familiar from what you were saying but if you could just let me know – this is the pre-survey and this is the post-survey – and Part A, if you could just let me know if you agree that there is a difference there or do you believe that it's pretty similar. Or… just what do you think?
> S: I think they're pretty similar, I just think with the post-survey I went into more explanation of what I thought about it –
> I: Okay.
> S: Because we had talked about it more in class, so…" (Student 129 interview, 3->1)

Question 3b. What is an experiment? (What does an experiment involve?)
Six of the seventeen students who agreed to be interviewed exhibited some improvement in their understanding of the distinction between experiments and other types of empirical inquiry used by scientists, as indicated by a comparison of how their pre- and post- instructional responses were coded (see examples in the seventh row of Table 3). During the interviews all of the respondents whose views changed in the favored direction agreed their views changed; two identified the cause of their changed opinions to be in class activities:

> "S: Ok. Um, I um, I remembered learning more about the experiment, the control group and experimental group in here. Like you have a control arm and an experimental arm of the, of it. So I think I, that kind of added to my, because we learned that a long time ago, but refreshing it while we were in class made me think about those again." (Student 39 interview, 4->1)

> "I: Ok. And why do you think you changed your answer a little bit to include the control and experimental group and it wasn't mentioned in the pre-survey. Do you recall why that is?
> S: I think this probably, like my post-survey, probably has more to do with in class, just the way we were talking about things, and just the different things we've gone over. So it might just be I had more information or…
> I: Ok, I see…
> S: That may be why…to be completely honest I don't know why I changed my answer, but I, I'm also positive we talked about this, my post survey, in class." (Student 71 interview, 2->1)

Question 4a. Does the development of scientific knowledge REQUIRE experiments?
Four of the seventeen students who agreed to be interviewed exhibited some improvement in their understanding that scientific knowledge can develop by means other than experiments, as indicated by a comparison of how their pre- and post- instructional responses were coded (see examples in the eighth row of Table 3). Two of the students were particularly forthcoming in why their views had changed, and, as with the previous pilot study (Rudge et. al. 2007), both students drew upon examples from the earth sciences:

> "I: Good, good. All right, now the next question asks you to consider whether it's essential, necessary for a scientist to conduct an experiment in order to um learn more about [inaudible]. So, does the development of scientific knowledge require experiments?

S: Um, from what I recall from class, you said it doesn't [laughs].
I: [Laughs] Well our goal here isn't for you to remember what was said in class…
S: Ok.
I: I just want to know what you believe today.
S: Um…
I: You can be honest and it won't hurt my feelings if you disagree with me.
S: Ok. Well, no I really do think there's some examples where scientists can just stumble upon something, and find new information that will help them along.
I: Ok, do you have an example in mind?
S: Um, I suppose like geology and studying like rocks or dirt or the earth. Um, there's really not too much there that you have to do experiments with. Um, you know you find a new rock, you identify it and it helps you find out about the past.
I: Ok, good. Again, let's see here. The first survey said "yes it does require experiments because in order to prove some kind of scientific knowledge you have to test it to see if it is in fact a theory. With the theory of gravity we couldn't just simply say what it is, we have to prove it through experiments." And in the second set, "Not always, like the example you just gave us in class with scientists looking at if there was another planet outside of Uranus. The scientist made calculations and found Neptune. No experiment was needed." So, do you see a difference between these two or do they say pretty much the same thing?
S: Um, there's definitely a difference. With the second one um I kind of took what we, what we were learning in class and saying that…Well, before I had I believe that all, you always had to have some experiment to prove to have science knowledge in an areas of science. With this, with the second time I wrote it down I just believed it wasn't always, you don't have to have an experiment to find science.
I: Now, it strikes me that the key thing here in both of your answers, and you tell me if I'm mistaken in this, is this idea of a test. So, in essence what you've opened up is the possibility that you can test a theory without experiments, is that sort of along the lines of what you had in line?
S: Yeah." (Student 16 interview, 4->1)

I: Sure. Alrighty. So… I'm going to ask this question once more. "Does the development of scientific knowledge require experiments?"
S: Uh… I don't – I want to say yes because that's, I mean, that, just for me, that's how I would want to do it. You know, like, that I would feel more confident if I actually did an experiment than if I just had a theory. I'm trying to think of a different way I could… you know, concretely feel like my, my theory was accurate if…
I: Um, okay. Um, in the – excuse me, I just have to look at the notes to see what you said – in the pre-survey, you said, "Yes. The development of scientific knowledge does require experiments. If you don't test your theories, you can only assume them to be accurate. An example is the relationship between the *Brassica rapa* (yawns) and the cabbage white butterfly. If you did not test the relationships, you could only assume that there was a relationship." And in the post, you said, "No. Someone may base their theories on observation. If you watch animals in the wild, you can come up with conclusions for their behavior without testing the accuracies." Oh –
S: I guess, I guess what I was – I mean, I still think both of those. I guess my thing is whether an observation is a type of experiment.
I: Good. Let's say for the moment that it isn't… Then it's – so we have two different ways to test a hypothesis. One might involve controls and experimentals and the other might involve just observations. Is it possible that the development of

> scientific theory could proceed with observations? Or would they have to have, use a theory?
> S: No, I think that they could with observations because you can't, I don't feel like you could completely use an experiment to test everything. Test all theories.
> I: Okay. Are there any instances where you just can't do experiments?
> S: Um… I don't know, like, something, like a theory about the world. I mean it's not like you can, you have two worlds and you can alter one. I mean maybe you can (laughs). I mean, someday that might, might happen where they could do that. But it seems like there are theories about the earth, you know, that you can't, you can't take the earth and put it in a little room and –
> I: Do you think biology's like that too?
> S: I think so.
> I: Alright.
> S: I think so." (Student 90 interview, 4->1)

Question 5. How are DIFFERENT CONCLUSIONS possible from the SAME SET OF DATA?

Five of the seventeen students who agreed to be interviewed exhibited some improvement in their understanding of how scientists might reach different conclusions from the same set of data, as indicated by a comparison of how their pre- and post- instructional responses were coded (see examples in the tenth row of Table 3). As with other responses, one of the interviewees minimized the import of written differences between his/her pre and post survey responses to this question:

> " I: Good. Ok. In the pre survey you said "because since no one really was there when the dinosaurs became extinct, scientists can take the same data as someone else and turn it into a hypothesis." And the second says, "No one can say for sure what happened to the dinosaurs. This example is just the scientists interpreting the information differently." Now, I know this seems a little redundant but I just need to check with you, do you think of these as pretty much the same thing or do you see a difference between them?
> S: Um, generally they're the same, I just kind of hit on different points each time. With the second one I just kind of hit more on um the idea of scientists interpreting the information differently. Well, I didn't really mention that over here. But, for both I said that nobody was there so its easy to say something different with the same information." (Student 16 interview, 2->1)

4. Lessons Learned

This began as a study of the efficacy of an instructional unit, but as our investigation progressed through the pilot and this final stage, it has increasingly become an abject lesson about the inadequacy of the VNOS protocol as a means of assessing student conceptions of issues associated with the nature of science, and in particular, whether these conceptions have changed.[3]

---

[3] Elby and Hammer (2001) have raised fundamental questions regarding the adequacy of the VNOS protocol as a single measure, pointing out that movement towards acceptance of a particular claim about the nature of science may not itself indicate the participant has become more sophisticated. Nagasawa (2004) has also raised concerns about the adequacy of the VNOS protocol as a single measure.

Student Intellectual Involvement

We began using the VNOS instrument in one section of BIOS 1700 as part of the evolution unit in the Fall of 2003 under the auspices of our local Human Subjects Institutional Review Board. Our first attempt as using the survey was a dismal failure owing to low participation. A third of the students didn't attend class that day, which we interpreted as reflecting a perception on the part of absent students that because they had just completed the third exam for the course they could miss the first day of the evolution unit without falling behind. We also made the mistake of identifying participation in the study with filling out the surveys. We think several students declined to participate simply out of intellectual laziness. To address these problems we started awarding points for attendance and made filling out the surveys a required activity (restricting our request for volunteers to the act of giving us permission to publish their written comments anonymously). Even with these changes, as noted above and in our previous pilot study, student responses to survey questions are often too brief and inarticulate to be particularly revealing (Rudge et. al. 2007). Survey responses could be graded for credit, but we are inclined to think such a procedure would compromise the validity of the instrument in that it would constitute an incentive for students to share what they think their instructor wants them to write, rather than their sincere beliefs, even if the grade was assigned for effort alone. While a lack of engagement by participants is a potential problem for any survey instrument, we are inclined to think it is particularly pernicious when it comes to the VNOS, because the process of coding depends crucially upon interpreting student responses and it is very easy to misread brief responses by students as more or less sophisticated than they should be (as was often revealed in the interviews, see below).

Abstract nature of VNOS questions

The VNOS instrument is also problematic in terms of the nature of the questions being asked. With the exception of question 5, the questions on the abbreviated instrument used here are all decontextualized. We are inclined to think even professional philosophers of science would have a difficult time defining what a theory is in the abstract and cannot help but wonder if it is unrealistic to expect students to share sophisticated understandings to such abstract prompts. Certainly part of the problem in interpreting the results of the VNOS survey has to do with the fact that a single question can be interpreted in multiple ways, as we found with respect to Question 3. What is an experiment? As noted above, many students in our study interpreted this as a question about what the point of doing an experiment is, whereas others interpreted this as a question about what the conduct of an experiment involves. While many students responded in a way that revealed their answers to both ways of interpreting the question, others did not. Related to this is a more general problem regarding how to interpret blank responses to questions, such as requests for examples. We'd like to think that if a student is unable to come up with an example in the pre-survey but does so in the post-survey, this represents improvement. Perhaps. But when the reverse happens, do we really want to say the student's ability to come up with an example has diminished?

Interview Results

Perhaps the most disconcerting result of our study's attempt to use the VNOS instrument has to do with what was revealed during the interviews. The written responses of interviewed students to the VNOS questions were initially coded by the first author. They were then recoded by the same author in light of the interview data (see Table 4). As may be seen from a casual inspection of the table, the initial coding of written participant responses proved to be largely unreliable. During the course of interviews students were often able to share more sophisticated views than provided in the surveys. We found interviewees were routinely able to provide examples despite the fact that many did not on the written survey, which suggests that the absence of an answer does not imply the participant could not come up with one. And conversely, we found many written responses to questions were coded as more sophisticated than revealed during the interviews. Several of the interviewees who wrote that experiments are always necessary for scientific knowledge to develop backed away from the claim when the interviewer asked them if an observation of animal behavior in nature represented an increase in knowledge, which we interpret as indicative of a simple semantic difference regarding what the word experiment means (an empirical test of any sort vs. a systematic observation of a system that has been perturbed). Below is just one example of the discrepancies noted, in this instance with reference to Question 6 regarding the role of imagination and creativity in science:

> "Yes, the scientists have to be so creative, if it were not for people thinking out of the box there would not have been as many new discoveries" (Student 116, pre-survey)
>
> "Yes, they use creativity, the [sic] must think out side of the box at times. One never knows what will lead to evidence." (Student 116, post-survey)

The above were scored 7->7 and 6->6 with regard to questions 6a and 6b, because in each case the student vaguely claims that scientists use creativity and imagination but does not specify which stages and provides no examples. Now compare the foregoing with the interviewee's responses during the interview:

> "I: I think we've beat it up (laughter). Okay, the last question is, um, based on creativity and imagination. It says, "Scientists perform experiments/investigations when trying to find answers to the questions they put forth. Do scientists use their creativity and imagination during investigations?" What do you believe today?
> S: Most definitely.
> I: Definitely? Okay, and in what way do they use their creativity? If you look at, um, you can look at this question and kind of refer to that or you can just tell me…
> S: Well.
> I: …in what way?
> S: For example, this bottle. Is that bottle half full or half empty?
> I: Um… what do you think? (Laughter)
> S: You know what I mean? It's like, you look at it here and you see it it's two-thirds of the way full but if I laid it over here and looked at it, you're going to see a different perspective. So, you have to be creative in that way that you present your evidence or how you see things.

Table 4. Comparison of How Interviewed Students Responses Were Coded

No.		1a	1b	1c	2a	2b	3a	3b	4a	4b	5	6a	6b
16	Surveys	3->2	2->2	5->2	2->3	6->6	2->1	4->4	4->2	2->2	2->1	1->8	6->2
	Interviews	“	“	4->4	2->2	6->2	“	4->1	4->1	4->2	“	1->1	“
39	Surveys	2->1	2->2	2->2	2->5	2->2	3->3	4->1	4->3	2->6	1->1	5->5	6->6
	Interviews	“	“	“	2->2	“	“	“	4->4	4->4	“	“	“
71	Surveys	2->2	3->3	2->5	5->5	2->2	3->3	4->1	1->1	2->2	1->1	1->1	2->2
	Interviews	“	3->1	“	5->2	“	“	2->1	“	“	“	“	“
90	Surveys	1->1	1->1	2->2	2->2	2->2	3->2	4->1	4->1	2->2	2->1	5->6	6->6
	Interviews	2->1	3->3	“	“	“	“	2->2	“	“	“	6->1	6->2
92	Surveys	3->1	3->1	2->1	2->2	2->2	2->5	2->1	6->6	4->5	4->4	5->5	5->1
	Interviews	1->1	“	2->2	“	“	4->4	2->2	3->4	6->2	3->1	“	6->1
95	Surveys	3->1	1->1	6->1	2->2	5->6	1->4	2->2	4->4	4->6	1->2	5->3	6->1
	Interviews	3->2	3->1	5->1	2->1	5->2	4->3	2->1	3->4	“	1->1	5->1	“
104	Surveys	1->1	1->2	2->2	2->2	2->2	1->1	2->4	4->4	6->6	1->1	3->7	2->6
	Interviews	“	1->1	2->4	“	“	“	2->2	4->1	“	“	3->3	4->6
107	Surveys	2->2	3->3	5->2	2->2	2->2	5->2	1->1	3->3	2->6	1->1	5->5	2->6
	Interviews	“	“	“	“	4->4	“	“	4->1	4->6	“	5->1	4->2
110	Surveys	1->1	1->1	2->1	2->2	2->1	1->3	1->1	3->3	6->6	1->1	1->5	2->6
	Interviews	1->2	“	“	“	“	5->3	2->2	“	6->1	“	1->1	2->2
111	Surveys	1->2	1->1	5->2	2->2	2->2	5->3	1->1	3->3	6->6	2->2	1->5	2->2
	Interviews	2->1	“	“	“	4->4	“	4->1	4->3	6->4	“	1->1	“
112	Surveys	2->1	1->2	5->2	5->2	4->6	3->4	2->4	4->4	6->6	1->1	7->1	2->6
	Interviews	“	1->1	“	5->5	4->4	“	2->1	3->3	6->2	“	6->1	4->2
116	Surveys	3->3	2->4	5->5	2->2	2->6	5->4	1->4	6->1	6->4	2->2	7->7	6->6
	Interviews	3->1	4->4	“	“	4->4	4->4	2->2	6->4	“	2->1	8->1	“
117	Surveys	1->1	2->1	2->1	2->2	2->1	1->3	4->2	4->4	2->1	4->2	1->1	2->1
	Interviews	2->1	3->2	“	“	4->1	1->1	“	4->3	4->1	2->1	2->1	“
121	Surveys	1->2	1->2	2->2	2->2	2->6	3->3	1->1	3->3	6->1	1->1	9->9	6->6
	Interviews	2->2	1->1	“	“	4->4	3->4	2->4	4->1	“	“	“	6->2
124	Surveys	1->1	1->1	2->1	2->2	2->2	3->3	1->1	3->4	6->6	2->1	5->5	6->6
	Interviews	“	“	5->1	“	“	“	“	4->4	“	“	6->5	6->1
125	Surveys	3->3	2->2	2->2	2->2	2->2	1->1	1->1	3->3	6->6	1->1	9->5	6->2
	Interviews	“	“	“	“	“	“	2->2	4->4	6->2	“	“	6->6
129	Surveys	3->1	2->2	2->2	2->2	2->2	2->4	4->1	1->1	4->2	2->2	5->5	2->6
	Interviews	“	“	“	“	“	“	2->1	“	6->2	2->1	6->6	4->2

> I: Okay. And is that the only stage that creativity and imagination is necessary or would you say that creativity is needed throughout the whole research process?
> S: Well, it's needed throughout the whole process. I mean, for one, you have to be creative in your question, on what it is you want or are looking for and two, you have to be creative in how you go about gathering the research, you know, you don't just want one, you know, internet stuff, resources. You want resources from books, magazines, articles, journals… so that takes creativity. Uh, you have to be able to express it so you have to be able to transfer your thoughts onto paper, so that takes creativity, organization, planning, collection, uh. And as far as publishing it, granted you can go out and say, "Yeah, this is it" and hold up a piece of paper but if you come up and give out, like, a motivational speech, "Yeah, this is it!" you know what I mean, that's going to have more impact than that one, per say, or a hard copy or a piece of paper compared to a powerpoint or what have you.
> I: Okay.
> S: So, yes.
> I: Okay…
> S: In all aspects, all the time.
> I: Okay. Did you want to add to that or, um, did you want to add any examples here? I noticed in your pre- and post-survey, I don't see any specific examples…
> S: The one I just listed is fine." (Student 116, interview)

What is particularly telling is the interviewee's claim, at the end of the above transcript, that an example has been given during the interview, when in point of fact no example has been given. Again, there appears to be a semantic issue at work here. At least some of the participants apparently interpreted a request for an example as a request for an abstract example rather than a specific example. And indeed, for some of the questions, such as the challenge posed by question 4 regarding whether experiments are always necessary, it is altogether understandable why students would provide a vague example, such as a scientist might make an observation, as their counter-example.

Interestingly, we discovered on multiple occasions that interviewees, whose pre- and post- instructional responses were coded differently (suggesting change had occurred), outright denied their views had changed. While one might argue that such a student might not consciously be aware of a change in his or her views, this completely undercut the follow up question of most interest to us, namely, "What led you to change your views?" In a couple of instances the interviewer pressed the student by asking him/her to clarify why the wording on the response was changed, pre vs. post. But this invariably led students to focus on what they were thinking at the time they filled out the post survey, not what happened in class that might have led them to change their views.

This brings us to a fundamental problem with the potential usefulness of the VNOS survey as an assessment instrument for the improvement of instruction. While the VNOS survey itself may have the potential to track whether change has occurred, it is extremely limited in providing insight into why change has occurred. None of the questions prompt the student to reflect why his/her views have changed, which is understandable in that the exact same instrument is administered both pre- and post- instruction. So the only way the written survey responses alone could reveal anything about why a student's views changed is by his/her choice of examples, and in particular, if the example was one mentioned in class. This underscores the importance of the interviews, because it is during the interviews that the interviewer gets to press the student on why his or her views changed. But as noted above in the Results subsection Interview Data, our experience in this latest study has led us to conclude that the interviews do not serve this function well. Many of the interviewed participants denied their views had changed, and even when they did acknowledge a change had taken place, they were unable to make more than vague references to what happened in class.

Was the VNOS instrument useful at all? During the interviews some students volunteered that filling out the survey was helpful to them in drawing their attention to issues associated with the nature of science.

> "I: Okay. Okay. Um, is there anything you wanted to add to this question? Anything else? How 'bout in terms of to the whole survey? Did you have any comments or any opinions you wanted to…
> S: I actually liked it.
> I: Oh, did you? Why?
> S: It's, it's kind of long but it, uh, it definitely made me, like, think a lot more, like, during this unit in our, in my class, it just like, I guess opened my eyes to a few more things like, kind of like, maybe subtle things, but like, that uh…
> I: Taking the pre, taking the pre-survey kind of opened your eyes to the beginning of that…
> S: Yeah.
> I: … unit or…
> S: Yeah, taking the pre one opened my eyes and then like, and then I was able to kind of like, think about, like, a few of those questions. Like, figure out the unit.
> I: Okay. Good." (Student 110 interview)

> "I: Anything else you wanted to just comment on in terms of the whole survey itself?
> S: Um, I thought it was good. I thought that um, I know like, the… in the beginning it seemed like my answers were a little bit more specific. I was only thinking of certain examples in my mind – I can't think of them now. And then in the post-survey, I seemed to like, broaden my understanding a little bit more and kind of think of things in more general terms and different cases than just one in particular.
> I: And why do you think that is?
> S: Um… I don't know, I think like, that over the semester, testing different things, I had more things in my mind that just maybe like one thing, the first thing that came to mind, I wrote about in the pre-survey. But then, in the post-survey, it was kind of like a conclusion thing in my mind and so I could think of like, different cases, and how those were tied together.
> I: And when you say you were, like, testing different things, what sort of things were you testing?
> S: Um, like, just like, thinking about, like, the experiments and stuff and taking stuff from other classes and stuff, like math and other stuff. I don't know… I just felt like I had a lot more to go on then.
> I: Okay. Well, thank you…" (Student 111 interview)

Did the study reveal anything that would potentially improve the conduct of the Mystery Phenomenon Unit? The responses to questions 1b and 2a suggest widespread confusion among students regarding how to distinguish between theories and hypotheses. Day 2 of the unit invites students to come up with experiments that would allow them to choose between three alternative accounts of the mystery phenomenon, and one can appreciate in light of this why the insight that experiments test hypotheses directly, but theories only indirectly would be lost on participants. Student comments in this study, and particularly in connection with the previous pilot study, also suggest that use of multiple examples of how observations in other historical sciences promote the development of scientific knowledge is helpful in drawing attention to the fact that experiments are not always necessary in science. We are nevertheless suspicious that use of the VNOS was necessary to provide us with these insights, and indeed doubtful that they justified the time, energy and financial resources necessary to carry the study out.

Acknowledgements

This material is based on work supported by the National Science Foundation Grant No. 0202923. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. The authors here acknowledge the assistance of Phyllis Haugabook Pennock, who administered surveys and conducted several of the interviews; and also Katherine L. Block and Brandy Ann Skjold, who wrote up transcripts from the interviews. Thanks also to the participants at the IHPST 2009 conference who attended our session, during which a preliminary report of our findings was shared.

Notes

[1] See Rudge & Howe (2009) for a discussion of the advantages of this instrumental approach to the use of history in the science classroom.

[2] For example, when asked to consider how one might test whether moths become darker because they need to, students invariably suggest one could test this by observing whether pale moths placed in a confined container having dark surfaces become darker to match their surroundings.

[3] Only one lecture session took place during the course of this three lab sequence. Students practiced problems that required them to distinguish evidence for natural selection from evidence for common descent. They also practiced explaining microevolutionary phenomena in terms of natural selection and did some concept mapping.

[4] A graduate student (who was not responsible for evaluating students in the affected class) administered surveys and conducted interviews in the absence of the first author, who analyzed data with assistance from the second author.

[5] In the text that follows, results of the present study will be compared with those of the second author's previous study as reported in Howe & Rudge (2005). The results of Abd-El-Khalick's (1998, 2001) previous study were not reported in such a way to allow a question by question comparison with the results of the present study.

[6] Question 3 was not included in surveys used in Howe (2004).

[7] The total population of students asked this question in Howe's (2004) study was 42 rather than 81 because of a change in its wording between semesters when the survey was administered.

[8] Question 6(a &b) was not included in surveys used in Howe (2004).

[9] Student 1 mentioned it with reference to her answers to the first three questions of the surveys, Student 2 with reference to the second question, Student 3 with reference to the fourth and sixth questions, Students 4 and 6 with reference to the first two questions, and Student 5 with reference to the third question.

References

Abd-El-Khalick F (1998) The influence of history of science courses on student conceptions of the nature of science. Unpublished doctoral dissertation, Oregon State University, Corvalis, OR

Abd-El-Khalick F (2001) Do history of science courses influence college students' views of the nature of science? Paper presented at the 6th Annual Meeting of the International History, Philosophy and Science Teaching Group, Denver, CO

Abd-El-Khalick F, Bell R, Lederman N (1998) The nature of science and instructional practice: Making the unnatural natural. Sci Educ 82:417-436

Abd-El-Khalick F, Lederman N (2000) The influence of history of science courses on students' views of nature of science. J Res Sci Teach 37:1057-1095

Akerson V, Abd-El-Khalick F, Lederman N (2000) Influence of a reflective explicit activity-based approach on elementary teacher's conceptions of the nature of science. J Res Sci Teach 37:295-317

American Association for the Advancement of Science (1990) Science for all americans. Oxford University Press, New York

Cotham J, Smith E (1981) Development and validation of the conceptions of scientific theories test. J Res Sci Teach 18:387-96

Elby A, Hammer D (2001) On the substance of a sophisticated epistemology. Sci Educ 85:554-567

Ginsberg HP (1997) Entertaining the child's mind: The clinical interview in psychological research and practice. Harper Collins, New York

Hagen J, Allchin D, Singer F (1996) Doing biology. Harper Collins, New York

Howe EM (2004) Using the history of research on sickle-cell anemia to affect preservice teachers' conceptions of the nature of science. Unpublished doctoral dissertation, Western Michigan University, Kalamazoo, MI

Howe EM, Rudge DW (2005) Recapitulating the history of sickle-cell anemia research: improving students' NOS views explicitly and reflectively, Sci & Educ 14:423-441

Khishfe R, Abd-El-Khalick F (2002) Influence of explicit and reflective versus implicit-oriented instruction on sixth graders' views of nature of science, J Res Sci Teach 39:551-578

Kim SY, Irving KE (2009) History of science as an instructional context: Student learning in genetics and the nature of science Sci & Educ - forthcoming

Lederman N (1998) The state of science education: Subject matter without context. Elec J Sci Ed (3) http://unr.edu/homepage/jcannon/ejse/lederman.html
Lederman N, Abd-El-Khalick F, Bell R, Schwartz R (2002). Views of nature of science questionnaire: Toward valid and meaningful assessment of learners' conceptions of nature of science. J Res Sci Teach 39:497-521
Lederman N, O'Malley M (1990). Students' perceptions of tentativeness of science: Development, use, and sources of change. Sci Educ 74:225-239
Lederman N, Wade P, Bell RL (1998) Assessing understanding of the nature of science: A historical perspective. In McComas WF (ed) The nature of science in science education. Kluwer, Netherlands
Matthews MR (1994). Science teaching: The role of history and philosophy of science. Rutledge, New York
Nagasawa P (2004) Epistemological relevations in students' science writing. Paper presented at the National Association for Research in Science Teaching, Vancouver, Canada
National Research Council (1996). National science education standards. National Academy Press, Washington D.C.
Palmquist B, Finley F (1997) Preservice teachers' views of the nature of science during a post baccalaureate science teaching program. J Res Sci Teach 34:595-615
Rudge DW (2004) Using the history of research on industrial melanism to help students better appreciate the nature of science; The mystery phenomenon: Lesson plans, In Metz D (ed) Proceedings of the seventh international history, philosophy science teaching group meeting. Winnipeg, Canada
Rudge DW, Howe EM (2009) An explicit and reflective approach to the use of history to promote understanding of the nature of science. Sci & Educ, 18(5):561-580.
Southerland SA, Johnston A, Sowell S, Settlage J (2005) Perhaps triangulation isn't enough: A call for crystallization as a methodological referent in NOS research. Paper presented at the AERA 2005 Annual Meeting (Montreal, Canada) http://digitalcommons.uconn.edu/terc_docs/1

Appendix A: VNOS Survey

1. Often in science we hear words like "theories" to describe scientific knowledge.
    - (a) What is a theory?
    - (b) How are theories created?
    - (c) Give an example of when you have created or used a theory?

2. After scientists have developed a scientific theory (e.g. atomic theory, theory of gravity), does the theory ever change?

    If you believe that scientific theories do change:
    - (a) Explain why.
    - (b) Defend your answer with examples.

    If you believe that scientific theories do <u>not</u> change:
    - (a) Explain why.
    - (b) Defend your answer with examples.

3. What is an experiment?

4. Does the development of scientific knowledge **require** experiments?
    - If yes, explain why and give an example to defend your position.
    - If no, explain why and give an example to defend your position.

5. It is believed that about 65 million years ago the dinosaurs became extinct. Of the hypotheses formulated by scientists to explain the extinction, two enjoy wide support. The first, formulated by one group of scientists, suggests that a huge meteorite hit the earth 65 million years ago and led to a series of events that caused the extinction. The second hypothesis, formulated by another group of scientists, suggests that massive and violent volcanic eruptions were responsible for the extinction. How are these **different conclusions** possible if scientists in both groups have access to and use the **same set of data** to derive their conclusions?

6. Scientists perform experiments/investigations when trying to find answers to the questions they put forth. Do scientists use their creativity and imagination during their investigations?

    If you believe <u>yes</u>, scientists do use imagination and creativity,
    - (a) Explain why, indicating which stages this occurs (planning and design, data collection, after data collection).
    - (b) Defend your answer with examples.

    If you believe no, scientists do not use imagination and creativity,
    - (a) Explain why.
    - (b) Defend your answer with examples.