UNDERSTANDING STUDENTS' CONCEPTIONS OF THE NATURE OF SCIENCE THROUGH MULTIPLE ASSESSMENT TOOLS

The importance of teaching the nature of science (NOS) to students in K-12 settings is highlighted in the benchmarks and standards. Many assessments used to measure students' and teachers' conceptions of the NOS have been criticized. Thus far, a qualitative and quantitative comparison of data collected from different assessment formats (forced choice, open ended-survey, and interviews) taken by the same students has not been undertaken, which was the goal of this study. Three assessments were chosen to test the creative, tentative, and empirical aspects of the NOS: the Modified Nature of Scientific Knowledge Scale (MNSKS), an Open-Ended Survey (OES), and follow-up interviews developed based on the work of previous researchers. Analysis was based on responses from thirteen 6^{th} grade students. Quantitative analysis showed that students had significantly higher scores on the tentative subscale of the MNSKS and that students made significantly fewer responses about the creative and empirical inferential NOS on the OES. Qualitative analysis revealed differences in the percentage of students holding informed NOS views between the MNSKS and OES. Inconsistencies in student responses were found within and between the assessments, creating difficulties in making clear comparisons and determining students' NOS views.

Dana Gnesdilow, University of Wisconsin- Madison

Sarah A. Sullivan, University of Wisconsin- Madison

Anushree Bopardikar, University of Wisconsin- Madison

Sadhana Puntambekar, University of Wisconsin- Madison

The Nature of Science (NOS) "typically refers to the epistemology of science, science as a way of knowing, or the values and beliefs inherent to scientific knowledge and its development (Lederman, 2007, p. 883) Despite disagreements between philosophers of science concerning the NOS (Alters, 1997), there is sufficient agreement between educators and researchers pertaining to the aspects of the NOS appropriate for teaching K-12 students (Lederman, 2007). The importance of teaching the NOS to students in K-12 settings is highlighted in the benchmarks and standards (AAAS,1993; NRC,1996), and in other international documents (McComas, Clough, & Almazroa (1998). It is believed that understanding of the NOS is key for achieving science literacy- promoting citizens who are better able to sort through complex science and technology issues related to every day life and make more informed decisions (McComas *et al.,* 1998).

In recent years, many assessments used to measure students' and teachers' conceptions of the NOS have been highly criticized. One of the main critiques has been aimed at the forced choice nature of earlier NOS assessments, which restricted students' responses to the view of the NOS promoted by the creator of the assessment (Lederman, Wade, &

Bell, 1998; Sandoval, 2005). In response to this drawback, many researchers have developed open ended surveys (e.g., Kafishe & Abd-El Kahlick, 2002; Lederman, Abd-El Kahlick, Bell, & Schwartz, 2002) and interview protocols (e.g., Carey, Evans, Honda, Jay, & Unger, 1989; Moss, Abrams, & Robb, 2001) designed to allow students to express their own views about this topic. However, open-ended measures and interviews require far more time to administer and even more time to "score" and sometimes contain complex questions that are abstract. Also, determining the meaning of student's often terse and poorly elaborated responses can be challenging (Liang, Chen, Chen, Kaya, Adams, Macklin, & Ebenezer, 2005; Roth & Roychoudhury, 1994; Sandoval, 2005).

While there have been many critiques of past NOS assessments, questioning their value in actually measuring students' NOS views, there has not been a qualitative and quantitative comparison of data collected from these different assessment formats taken by the same students. This study conducted such an examination and presents the consistencies and inconsistency of student responses within and between different types of NOS measures. Such an examination is likely to add another layer to NOS assessment research by providing a more comprehensive picture of students' thinking in this area which may have implications for designing future NOS assessments.

McComas, Clough, & Almazroa (1998) provided a list of twelve overlapping NOS objectives gathered from eight international science education standards documents. Three of these categories were examined in this study: a) Scientists are creative; b) Scientific knowledge is empirical and impinges heavily, on observations, experimental evidence, logical arguments, and doubt; and c) Science is tentative, or subject to change based on new evidence. These three NOS categories, (creative, empirical [distinguishing between observations and inferences], and tentative) were chosen for examination based on the prior work of Khishfe & Abd-El-Khalick (2002) due to their correspondence with the stated goals for teaching the NOS in the benchmarks and standards for the age group of students who participated in this study (AAAS, 1993; NRC, 1996).

Method

Participants

This study was conducted with nineteen 6th grade students at a private mid-western middle school. These sixth graders were selected because they were already participating in the CoMPASS research project, developed by the fourth author of this study. Of the 19 students in the study, data from thirteen students (nine male and four female) were complete and included in the analysis.

Measures and Scoring Procedures

Two assessments were chosen to test the creative, tentative, and empirical aspects of the NOS for 6th graders. The first measure, the Modified Nature of Scientific Knowledge Scale (MNSKS) (Miechtry, 1992) was chosen because it was developed and tested for use with 6th-8th grade students. The MNSKS consists of 32 statements related to the creative, testable (empirical), developmental (tentative), and unified dimensions of the NOS. The unified subscale was omitted for this study since it did not fall into the three categories of interest identified above. This Modified Nature of Scientific Knowledge Scale consisted of the same positive and negative statements, eight per subcategory, presented in a five choice, Likert-scale format as the original MNSKS. A higher score on this measure indicates that the student's NOS views are more consistent with the premise of the scale. Forty points represents the highest sub-scale score, eight the lowest, and a score of twenty four reflects a more neutral position. One sample statement from each of the three MNSKS subscales is listed in Table 1 in order to provide examples of the types of statements presented to students.

Table 1

Sample statements from each of the MNSKS subscales.

MNSKS Subscale:	Statement:
Creative	Producing scientific knowledge involves human imagination.
Empirical	Scientific laws, theories, and concepts are based upon repeated observations.
Tentative	Scientific beliefs do not change over time.

The second measure, an Open-Ended Survey (OES) related to the NOS, was created based on two different NOS open-ended measures: 1) "Nature of Science Survey" (Khishfe & Abd-El-Khalick, 2002), and 2) the "Perspectives on Scientific Epistemology" (POSE) Questionnaire (Abd-El-Khalick, 2002). The OES in this study consisted of six main questions, four of which contained sub-questions for a total of eleven questions, aimed at eliciting the creative, tentative, and empirical aspects of the NOS. For example, a three part question that we used was: "Scientists believe that the dinosaurs lived more than 65 million years ago. How do scientists know that the dinosaurs really existed? How do scientists determine what the dinosaurs actually look like (for example, the texture and color of their skin, the shape of their eyes)? How certain do you think the scientists are about how the dinosaurs actually looked?

Finally, follow-up interviews were conducted to ensure that researcher's interpretations of the students' open-ended responses on the OES were accurate and reliable. Interviews also provided a means to qualitatively compare the students' interview responses to their responses on MNSKS and OES. Interviews were semi-structured and based on protocols of Khishfe &Abd-El-Khalick (2002) and Lederman, Abd-El Kahlick, Bell, and Schwartz, (2002). The semi-structured interviews were conducted by two of the authors. Each researcher read the question from the OES, followed by the student's written response to it, and then asked the student to elaborate and clarify his or her answer.

The three different assessments were chosen because we believed that they would provide different glimpses of the students' NOS conceptions for qualitative and quantitative comparison to attempt to determine the affordances and constraints of each.

Data Sources and Analysis

We analyzed student responses on all three assessments. Student responses on the MNSKS subscales were determined as follows: for positively stated items, a response of "strongly agree" was assigned 5 points and a "strongly disagree" was assigned a 1. Negatively stated items were scored in the reverse manner. The highest score for each subscale was forty points, eight the lowest. A score of twenty-four reflected a neutral position. Higher scores indicated that the students' NOS views were more consistent with the premise of the scale and more informed.

Coding the OES and Interviews

We analyzed the OES and interview responses using a coding scheme that examined each of the identified NOS aspects of interest in this study (creative, empirical, and tentative). A student was said to hold informed NOS views for each aspect of the NOS if all of his or her responses were consistent with currently accepted NOS views. Student responses were coded as either being informed or uninformed within each NOS category. Two raters independently analyzed the responses to the eight questions from the OES and follow-up interviews and coded the responses as discussed in the bulleted section below. Examples of student responses for each category on the OES are presented in Table 2. Inter-rater reliability for coding the OES (93%) and interviews (83%) were established.

• Creative- Informed views of the creative NOS refer to an understanding that scientists use their imaginations and creativity in designing investigations, solving problems, and in developing scientific explanations and models.

• Tentative- Informed views of the tentative NOS acknowledge that science is an ever changing body of knowledge based on scientists' best attempts at understanding the world- scientific facts, theories, and laws are subject to change based on new, relevant information.

• Empirical- Informed views of the empirical NOS identify that scientific evidence is based on observation and scientists make inferences based on scientific evidence. Thus, responses were placed into two empirical categories, inferential and observational to see how many students saw inference as a part of the empirical NOS in addition to empirical observation.

Table 2

Examples of Students' Informed and Uninformed NOS Views from the OES

NOS Coding Category:	Creative	Tentative	Empirical Observation	Empirical Inference
Example of Informed Student Responses:	"Yes I do (think scientists use imagination and creativity in conducting research/ experiments) because doing in experiment is being creative"	"I think they are about 50% certain because they have never really seen a dinosaur."	"(Scientists know that the dinosaurs really existed because) there are dinosaur bones all over the world today & fossils in the ground"	"They just make scientific guesses because it is impossible for any one to know what a dinosaur looked like because no one is that old"
Example of Uninformed Student Responses:	"no (I don't think scientists use imagination and creativity in con-ducting research/ experiments) because every-thing that scientists do is supported by knowledge."	"I think scientists know what they (dinosaurs) look like because when they put the bones together they can tell."	none	none

Quantitative Analyses: Within-Measure Comparisons

When examining students' OES responses in reference to the NOS categories (creative, empirical observations and inferences, and tentative), the frequency of responses for each aspect was determined by counting the number of responses that a student made in reference to the identified NOS category. Since each student had eight question opportunities to freely express his or her ideas about the NOS, frequencies were calculated by dividing the number of responses related to a particular category by eight. For example, out of the eight opportunities to respond, one student made two statements (25%) about the creative aspect of the NOS.

We decided to conduct a quantitative statistical analysis to test if there was a significant difference in the number of responses that students made in each category of the NOS measured on the OES. We thought that this might be informative for a couple of key reasons: 1) So that the reader would know the frequency of responses that all of our data analyses were based on, and 2) Some readers might be interested in knowing how students responded when given the chance to freely express their views on the OES. A Friedman test was conducted to evaluate differences in the frequency of responses to the different NOS categories of interest on the OES (creative: median = 1, tentative: median = 3, empirical observation: median = 4, and empirical inference: median = 1) (see Figure 1). The test was significant, $\chi 2$ (3, N = 13) = 26.363, p = .000.
Follow up pairwise comparisons were conducted using the Wilcoxon test and controlling for the Type I errors across these comparisons at the .05 level using the Least Significant Difference (LSD) procedure. The median number of empirical observation responses on the OES was significantly greater than both the number of creative responses made, p = .002 and the number of empirical inference responses, p = .003, but the number of empirical observation responses did not differ significantly from the number of tentative, p = .070. Additionally, the median number of creative responses made on the OES was significantly less than the number of tentative responses made, p = .002.

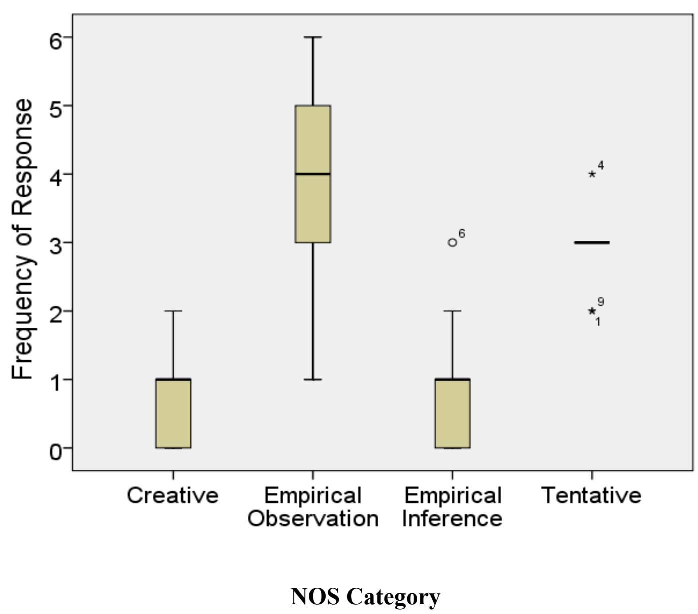

Figure 1. Frequency of Responses on the OES

A second within measure test was conducted in order to find out if there was a significant difference in students' MNSKS subscale scores in order to find out if students' were more or less informed in the different NOS categories. A Friedman test to evaluate differences in the median values of the NOS for the creative (median = 30), tentative (median = 31), and empirical (median =26) subscales scores on the MNSKS was conducted (see Figure 2). The test was significant, $\chi 2$ (2, N = 13) = 16.286, p =.000.

Follow up pairwise comparisons were performed using the Wilcoxon test, controlling for the Type I errors across these comparisons at the .05 level using the LSD procedure. The median score for the tentative subscale was significantly greater than both the creative subscale median score, p = .005 and the empirical subscale median score, p = .001, but the creative sub-scale median score did not differ significantly from the empirical subscale median score, p = .223.

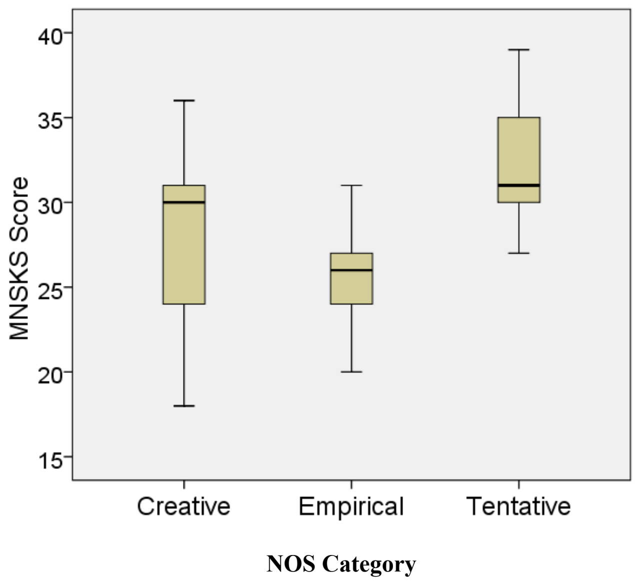

Figure 2. Scores on MNSKS Subscales

Qualitative Analyses

The MNSKS and OES were designed by different researchers with different ideas about how to assess students' views of the nature of science, which made them difficult to compare in clear cut ways. In attempting to make a comparison between the continuous, quantitative student scores on the MNSKS and responses on the qualitative, discrete, open-ended survey, we have converted the data into percentages of informed responses for each NOS category of interest.

Between-Measures Qualitative Analysis.

For the OES, the percentage of students holding informed views were calculated by dividing the number of students' holding informed views in each category by the total number of students. Additionally, in order to make a clearer comparison between the MNSKS and the open-ended survey in the empirical category (since the OES delineated between students' understanding of empirical observation and inference, unlike the MNSKS), a student was said to hold informed views on the empirical NOS on the open-ended survey if she held informed NOS views of both the observational and inferential empirical NOS. In many cases, students did not discuss the inferential aspect of the empirical NOS, thus they could not be said to hold informed views.

Since a score of 24 on the MNSKS represents a neutral response, students who scored a 25 or above (up to 40 points) on each subscale can be said to hold more informed views of the tentative, creative, and empirical NOS (Meichtry, 1992). Consequently, for this qualitative between-measures analysis we calculated the percentage of informed responses on the MNSKS for each NOS subscale by dividing the number of students who scored above 24 points by the total number of students.

When comparing the two different measures in this way, we found that more students held informed views of the empirical NOS on the MNSKS (61.54%) versus the OES (46.15%). For the creative NOS, the percentage of students holding informed views on the MNSKS (61.54%) was higher than on the OES (53.85%). The percentage of students found to hold informed views of the tentative NOS on the MNSKS (100%) was much higher than on the OES (30.77%).

Qualitative Examination of Consistencies and Inconsistencies.

In further examining students' responses within and between the different measures, we found that most students made inconsistent responses. An inconsistency in response meant that students made statements or responded showing an informed view related to one of the aspects of the NOS and then expressed views that were contradictory to this response within the same measure, or on one of the other NOS measures. For example, while 100% of students were found to hold informed views of the tentative NOS on the MNSKS, six of these students agreed with the statement "Scientific knowledge is true beyond a doubt," and five of them responded neutrally. (Interestingly, 29.48% (92 out of 312 in total) of all student MNSKS responses were neutral.) Agreement with this statement is a clear contradiction to the notion that science is tentative.

The discrepancies between the percentage of students with informed views of the tentative NOS on the MNSKS and OES (reported in previous paragraph) further displays inconsistencies in responses between measures. Other examples of between test inconsistencies can be seen by examining some of the responses made by a student who scored 39 out of 40 points on the tentative subscale on the MNSKS. While this student had the highest tentative MNSKS score, s/he was deemed not to hold informed views on the OES, stating, "I would say that they are 99.9% sure (that scientists are sure that an atom looks the way it does in a picture presented on the OES) because they probably used a very, very accurate microscope." Additionally, this student made no reference to the inferential nature of the empirical NOS on the OES, but did discuss this NOS aspect three times when elaborating upon OES responses during the follow-up interview.

Discussion

The quantitative analysis showed that students made significantly fewer responses related to the creative and empirical inference aspects of the NOS than the empirical observation aspect on the open-ended survey. They also made significantly fewer creative NOS responses than tentative NOS responses on the OES. This fact leads us to wonder: 1) Why were there fewer responses in these NOS categories in this group of students? Did question construction affect their responses or did this group of students as a whole simply make fewer considerations about the creative and inferential NOS? 2) If we used the same OES in other educational settings, would we find the same lack of creative and empirical inference responses? If so, perhaps researchers and educators should examine why students are less likely to conceive science as a creative and inferential process. 3) Can we say that a student holds an informed view of the NOS if they have only given one response? It seems to us that the evaluation of students' NOS ideas based on only one response is shaky at best. 4) Is it feasible to administer an open-ended test that does not require students to answer questions related to each aspect of the NOS of interest and, further, can we decide that students' views are less informed on a NOS category if they make no mention of it within their responses? We ask this two-part question to highlight the possible drawbacks in accurately measuring students' NOS views when employing an open-ended survey. Many students did not discuss all of the NOS aspects of interest when answering questions on the OES. While we might hope that these issues will be resolved through the semi-structured follow-up interviews, our experience indicates that there is no guarantee that this will occur.

While students' MNSKS tentative NOS subscale scores were significantly higher than the creative and empirical, and our qualitative analysis showed 100% of students having an informed tentative view on the MNSKS, almost half of the students in the study agreed and five students made a neutral response to the MNSKS statement that "Scientific knowledge is true beyond a doubt." Agreement with this statement is clear evidence that many students' conceptions of the tentative NOS may in fact be less informed than MNSKS results indicate. In this case, perhaps the results of the OES (which better takes account of students' inconsistent statements), showing that only 30.77% of the students showed informed tentative NOS views, gives a more accurate picture of the students' conceptions.

Questions about what score constitutes an informed view on the MNSKS have come to the forefront in conducting this analysis. If we are trying to make comparisons to the OES, where responses in a category were coded as either informed or not for a student to be categorized as holding informed views, shouldn't we hold the scores on the MNSKS to the same criteria to be more consistent? In this case, we would have only counted an MNSKS subscale score as informed if the maximum score was achieved; 40 points. With this stringent cut off, none of the students in this study would have been found to hold informed views on any of the NOS aspects that we examined in accordance with their scores on the MNSKS. However, we felt that this all-or-nothing approach to the MNSKS was inappropriate. MNSKS scores are based on a student's responses to several similar items within one dimension of interest. Scores on this measure are therefore meant to represent a continuum of views that are either more or less informed. Additionally, since students are forced into making a choice based on the views of the MNSKS test constructor, would we expect that students' views would conform 100% to the views of those who constructed the test? This is a much different situation than the OES, where students are able to freely express their ideas. Choices of what actually constitutes an informed score on the MNSKS seem to become somewhat arbitrary when attempting to make these types of analytical comparisons. For example, if a researcher decided to categorize any score above a 33 (the score just above the middle point between 24 and 40) on any MNSKS subscale as informed (in attempt to capture scores that are farther from neutral and closer to informed), the percentage of informed views might be drastically different from the ones we reported in our results section. This leads us to further question the value of using such a measure if it is going to be used for this type of comparison.

Aside from making students conform to the NOS notions of the designer of the scale, forced choice scales such as the MNSKS reveal nothing about why students might choose one response over another or how confident students are in their choices. Additionally, what do neutral responses on the MNSKS mean? Were students unable to interpret the questions or did they actually take a neutral position in regard to the statement? MNSKS neutral responses provide sparse information about students' thinking. Open-ended surveys and interviews can give us a more detailed picture of the students' thinking; however, many younger students lack the communication skills necessary to clearly express their thoughts. In a few cases, student responses on the OES were nebulous and were not elucidated during follow-up interviews. So far, the qualitative analysis of the inconsistencies in students' responses within and between measures reveals that they may not provide a full or accurate picture of students' NOS conceptions because, even when the measures are combined, it is not easy to make clear comparisons and determine students' NOS views.

Conclusions / Future Implications

Understanding the role of the NOS in science education, science learning, and science literacy has been a primary interest of researchers and science educators for decades. The creation of NOS assessments that are reliable, valid, and easy to administer and score are essential for further research on how students' and teachers' conceptions may affect science teaching and learning and scientific literacy. It is extremely important that we develop reliable, valid, and usable measure to assess whether or not we are meeting these goals in all science classrooms. This study provided a comparison of three measures. More research is needed to further explore the affordances of each to develop more authentic nature of science measures.

Acknowledgement

The research described in this paper is supported by an IERI grant (#0437660) from the National Science Foundation to the fourth author.

References

Ab-El-Khalick, F. (2002, April). *The development of conceptions of the nature of scientific knowledge and knowing in the middle and high school years: A cross-sectional study*. Paper presented at the annual meeting of the National Association for Research in Science Teaching, New Orleans, LA.

Alters, B. J., (1997). Whose nature of science?. *Journal of Research in Science Teaching, 34*(1), 39-55.

American Association for the Advancement of Science. (1993). Benchmarks for Science Literacy. New York: Oxford University Press.

Carey, S., Evans, R., Honda, M., Jay, E., & Unger, C. (1989). "An experiment is when you try it and see if it works": A study of grade 7 students' understanding of the construction of scientific knowledge. *International Journal of Science Education, 11*, 514-529.

Khishfe, R., & Abd-El-Khalick, F. (2002). Influence of explicit and reflective versus implicit inquiry-oriented instruction on sixth graders' views of nature of science. *Journal of Research in Science Teaching, 39*(7), 551-578.

Lederman, N. G. (2007). Nature of science: Past, present, and future. In S. K. Abell & N.G. Lederman (Eds.), *Handbook of Research in Science Education,* 831-879. Englewood Cliffs, NJ: Erlbaum Publishers.

Lederman, N. G., Abd-El-Khalick, F., Bell, R.L., & Schwartz, R.S. (2002). Views of nature of science questionnaire: Toward valid and meaningful assessment of learners' conceptions of nature of science. *Journal of Research in Science Teaching*, *39*(6), 497-521.

Lederman, N. G., Wade, P. D., & Bell, R. L. (1998). Assessing understanding of the nature of science: A historical perspective. In W.F. McComas (Ed.), *The Nature of Science in Science Education: Rationales and Strategies,* (pp.331-350). Dordrecht: Kluwer Academic Publishers

Liang, L., Chen, S., Chen, X., Kaya, O.N., Adams, A.D., Macklin, M., & Ebenezer, J. (2005). *Student understanding of scientific inquiry (SUSI): Development and validation of an assessment instrument.* Paper presented at Eighth International History, Philosophy, Sociology & Science Teaching Conference. University of Leeds, UK

McComas, W. F., Clough, M., & Almazroa, H. (1998). The role and character of the nature of science in science education. In W.F. McComas (Ed.), *The Nature of Science in Science Education: Rationales and Strategies,* (pp.3-40). Dordrecht: Kluwer Academic Publishers

Meichtry, Y. J. (1992). Influencing student understanding of the nature of science: Data from a case of curriculum development. *Journal of Research in Science Teaching, 29*(4), 389–407.

Moss, D.M., Abrams, E.D., & Robb, J. (2001). Examining student conceptions of the nature of science. *International Journal of Science Education, 23*(8), 771-790.

National Research Council. (1996). National science education standards. Washington, DC: National Academic Press.

Roth, W-M., Roychoudhoury, A. (1994). Physics students' epistemologies and view about knowing and learning. *Journal of Research in Science Teaching, 31*(1), 5-30.

Sandoval, W. A. (2005). Understanding students' practical epistemologies and their influence on learning through inquiry. *Science Education, 89*(4), 634-656.