Students' Reasoning and the Level of Interactivity in Science Content Courses for Future Elementary Teachers

Dean A. Zollman, Kansas State University, dzollman@phys.ksu.edu

Mojgan Matloob-Haghanikar, Winona State University

Sytil Murphy, Shepherd University

Abstract

As part of a study of the science preparation of elementary school teachers, we investigated the quality of students' reasoning and explored the relationship between sophistication of reasoning and the degree to which the courses were measured to be interactive. First, we devised written content exam questions, which were open ended and required students to apply recently learned concepts in a new context. All the questions developed were based on a common template that required students to recognize and generalize the relevant facts or concepts and apply them. To evaluate students' answers, we developed a rubric based on Bloom's taxonomy as revised and expanded by Anderson *et al*. Along with analyzing students' reasoning, we visited 20 universities, observed the courses and used the RTOP to determine their level of interactivity. Statistical analyses indicate some relationship between the students' reasoning on the exams and the level of interaction in the class.

Keywords: elementary teachers, interactivity, reasoning, pre-service

Students' Reasoning and the Level of Interactivity in Science Content Courses for Future Elementary Teachers

As part of the National Study of Undergraduate Education in Science (NSEUS) which is investigating how active engagement in undergraduate science content courses taken by future teachers affects their teaching of science in elementary schools, we have investigated the relationship between students' learning of content knowledge and the degree that science courses utilize interactive (reformed) teaching-learning strategies.

The courses, which were investigated at several universities, covered a variety of science disciplines. Accordingly, a direct comparison on subject matter learning was impossible. Instead, we concentrated on comparisons of evidence of reasoning skills within the content that the students had learned. Because of the size of study, we were not able to interview the students and needed to rely on written responses to exam questions. Therefore, we wrote questions designed to elicit reasoning skills and developed a rubric for comparing the reasoning patterns in the students' written responses.

Framework

Previous research often emphasizes the significance of applying prior knowledge to construct knowledge in a new context. In agreement with researchers of schemata theory (Mayer, 2002) the level of understanding relates to the pieces of knowledge, cognitive abilities that students bring to a new context and the way they connect and organize pieces of information. In other words, reasoning can be defined in terms of the thought processes and knowledge types that students bring to a new context. As the student brings the knowledge pieces together, they create new knowledge through mental processes such as association, classification, combination and refinement (Dufresne, R., *et al.*, 2005). The National Science Educational Standards (National Research Council, 1996) also emphasize conceptual understanding, using various procedural skills to approach a problem and engage students in higher levels of thinking such as classifying, summarizing, inferring, comparing, explaining and applying their prior knowledge to a new context. According to the National Science Education Standards, successive statements that follow one another logically without gaps from statement to statement characterize a well-reasoned response. Evaluating correctness, use of controlling variables or measuring students' conceptual knowledge may not effectively assess the students' gains due to inquiry (Russ, R. *et al*. 2008). Russ *et al*. [4] emphasized not only the association of cause and effect, but also the underlying process that explains how the cause and effect are associated.

Using this previous research as a foundation, we selected our objectives from Bloom's taxonomy as revised by Anderson and Krathwohl, (2001). This taxonomy was developed for organizing and classifying instructional objectives. One of its main goals is to make objectives specific and clear for instructional plans and assessment design. Our objectives along the cognitive dimension, include *Compare, Infer and Explain* from the category ***Understand*** and the category ***Apply***. These characteristics of reasoning were used to design our rubric.

Design/Procedure

Rubric

Based on the assessment objectives above, we developed a scoring guide with four traits that distinguishes the different levels of reasoning in the cognitive domain. In contrast to holistic scoring in this type of approach, called analytical-trait scoring [6], the assessor judges students' performance several times, each time through the lens of different criteria. Based on the description provided for each trait, we determined whether evidence for the particular trait occurred in the written answer.

Understand /Compare and Contrast

We looked for the evidence of comparing those aspects and features that were fundamental for justifying cause and effect changes, or comparing variables that provide plausible evidence for why and how and what changed that caused the effect.

Understand/ Infer

We assessed if the answer recognized the patterns that connected series of the events and instances and plausible connections and relations between cause and effect.

Understand/ Explain

We looked for a cohesive and convergent argument that described the situation, predicted the outcome, and was well supported by showing why and how things are happening.

Apply

We sought evidence of the association among facts, concepts and procedures that were reconstructed in connection with the features of question scenario to present a plausible answer.

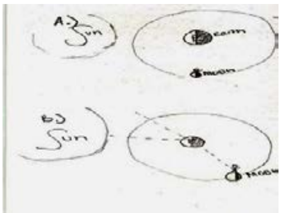
Analysis of Student Responses

For an example of questions that we used we present a question that was given to an astronomy course for elementary education majors at a small Midwestern university. This question is based on one developed for an inquiry-based physical science course. (McDermott *et al.* 1996) Fifty students completed this question on the final exam.

Question: You look outside and see a first quarter moon. Suppose that an astronaut were on the moon looking at Earth. Make a sketch of the Earth as seen by the astronaut. How will the illuminated portion of the Earth appear different three days later?

We have chosen two typical responses that are representative of the types of reasoning that show evidence or non-evidence for each trait of our rubric. The students submitted the drawing below each response as of the part of the response.

Response 1) *The astronaut would see a 3rd quarter, waning moon. The moon will have moved slightly more in its evolution, making earth see the moon as slightly more than 1st quarter. In contrast, the earth would appear less full to the astronaut on the moon.*



Response 2) *The earth-illuminated portion would decrease same, it would be a waning gibbous instead of a third quarter. It would be even a waning crescent almost a full earth, depending on the rotation*

Our analysis of the evidence for each component is:

- Understand/Compare and Contrast:
  The first response compared the moon's positioning and moon's and earth's sunlit portions in the sun-earth-moon model. The second response just compared the appearance of the sunlit portions of earth that are analogous to the moon. Although the first student's response shows more in-depth reasoning than the second, in both cases we see evidence for the cognitive process of compare.
- Understand/Infer:
  The first response showed an in-depth interconnection between a series of causes and effects including the changes in the location of the observer in two situations on the sun-earth-moon geometrical model that in effect causes a change in the appearance of the sunlit portions of the earth and the moon. However, the second response includes a series of disconnected concepts without any plausible connection between the described events.
- Understand/Explain
  The first student described his/her understanding of the situation and predicted the outcome with series of explanations that describes why and how the predicted outcome is true. However, the second student did not provide any additional explanation to support his/her predicted outcome.
- Apply:
  The first student reconstructed facts, concepts, and procedural skills from the perspective of the observer that is located on the moon, whereas, the second student did not discuss the question in the context of new situation.

Site Visits

To measure the degree to which a science class was interactive we used the Reformed Teaching Observation Protocol (RTOP) (Swada, D., *et al*. 2000). The characteristics in this observational protocol are organized into five categories: Lesson Design and Implementation, Propositional Pedagogical Knowledge, Procedural Knowledge, Communicative Interactions and Student/teacher Relationships. Each category includes five items which the observer ranks on a scale of 0-4. Observations for the RTOPs took place during site visits to each institution in the middle of the semester.

Findings & Analysis

To explore the relationship between the evidence of the students' reasoning and the measure of interactivity, we collected data from 904 students in 18 courses at 13 universities. For every student's response to the content questions we assigned four binary codes indicating whether the response showed evidence for each of the cognitive traits above. Because our data are categorical and our variables are dichotomous, we used a logistic regression model [9].

Odds Ratio Analysis

As a first step in the analysis of these data, we used a modified analysis which is based on logistic regression with two dichotomous variables. To obtain a dichotomous variable for the RTOP score, we used as a break point the average RTOP score for all of the classes observed. That average (65.5) was the boundary between high and low RTOP scores.

A common way to quote results from logistic studies involving two dichotomous variables is to use the "odds ratio." First we calculated the odds that students will show evidence for each trait of rubric if they were in a class with a higher than average RTOP. (i.e. RTOP > 65.5). Then, we calculated the odds for the students in a class with lower than average RTOP scores. The odds are given by:

$$Odds = \frac{Evidence}{No\ Evidence}$$

Then we calculated the odds ratio:

$$Odds\ ratio = \frac{High\ RTOP\ odds}{Low\ RTOP\ odds}$$

Table 1 shows the number of students who showed evidence or no-evidence for cognitive process of "Apply" for the two groups of RTOP<65.5 and RTOP>65.5. The odds ratio for this measure was 1.3, which implies that a student in a higher than average RTOP class is 1.3 times more likely to show evidence of using "Apply" than one in a low RTOP class. Table 2 gives the odds ratios for the cognitive processes we investigated in this way.

Table 1. Odds ratio for the "Apply" cognitive process

Apply	RTOP		Total
	<65.5	>65.5	
Evidence	191	254	445
No-Evidence	190	195	385
Total	381	449	730

Table 2. Odds ratio for all cognitive processes under investigation

Cognitive process	Odds ratio
Understand /Compare	1.84
Understand /Infer	1.42
Understand/Explain	1.00
Apply	1.30

Summary of General Analysis

Describing the full logistic analysis is beyond the scope of the present paper. Here we present a brief summary of the analysis to address two research questions:

- What is the relation between the quality of students’ thought processes and the degree to which course is considered to be reformed as measured by the RTOP?
- What is the relation between the conceptual structure of the responses and the degree to which course is considered to be reformed as measured by the RTOP?

We analyzed the data using two different models which we will call the simplified and generalized models. The difference between simplified and generalized version of logistic regression lies in the number of independent variables that were accounted for in the model. In the simplified version we ran the analysis with one independent variable in the model and in the generalized version we created a statistical model in terms of two independent variables.

In the generalized version of the analysis, we made use of two predictors. Using RTOP we initially considered five predictors including; the quality of lesson design, propositional knowledge, procedural knowledge, communication/interaction and student/teacher relationship. However, for the data we collected from 13 universities, the measures for lesson design, procedural knowledge, and communication/interaction and student/teacher relationship were highly correlated. Therefore, we defined a new predictor called “combination score” that was the average of these four variables with propositional knowledge a separate variable.

For the simplified version of analysis, we obtained that evidence for occurring knowledge and cognitive traits in most of the cases were depending on overall RTOP scores in the favor of higher RTOP scores. While, the results showed gradual increase for the traits factual, conceptual, compare and apply, there was no significant change infer and explain and there was a steep decrease for procedural knowledge. For the higher RTOP overall scores, students’ responses to exam question displayed a better understanding of facts and concepts; moreover, the answers improved in the use of facts and concepts in connection with the features of question as RTOP score increased. However, the responses did not show a positive correlation with evidence for higher cognitive processes such as infer and explain to present a plausible answer.

In sum, the variation of RTOP overall scores were between 20/100 and 90/100. We found a correlation between the RTOP overall score with the probability that the trait was displayed in students’ answers. As the RTOP overall score increased, factual, conceptual, compare and apply increased slightly. However, infer and explain remained constant and procedural knowledge decreased.

Acknowledgments

This work is supported by National Science Foundation Grant ESI 0554594. We thank other members of the NSEUS Team (See www.nseus.org) for helpful comments and suggestions on this research.

References

Anderson, W. and Krathwohl, D.R (2001) *Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives,* New York: Longman

Dufresne, R., *et al*., (2005) "Knowledge representation and coordination in the transfer process" in J. P. Mestre (Ed.), *Transfer of learning from a modern multidisciplinary perspective*. Greenwich, CT: Information Age

Hosmer, D. W. & Lemeshow, S. (2000) *Applied Logistic Regression,* New York: John Wiley & Sons

Mayer, R. E. (2002) "Understanding conceptual change: A commentary". In Limon, M., & Mason, L. (Eds.), *Reconsidering conceptual change: Issues in theory and practice,* Dordrecht, The Netherlands: Kluwer Academic Publishers.

McDermott, L.C. *et al.* (1996) *Physics by Inquiry,* New York: John Wiley & Sons

National Research Council (1996) *National Science Education Standard*, Washington, DC, National Academy Press.

Russ, R. *et al*. (2008) "Recognizing mechanistic reasoning in student scientific inquiry: framework for discourse analysis developed from philosophy of science." *Science Education*, **92**: p,499 – 525

Swada, D., *et al*. (2000) *Reformed Training Observation Protocol (RTOP) Training Guide*. Tempe, AZ: Arizona Collaborative for Excellence in the Preparation of Teachers

Wiggins, G. and McTighe, J. (1998).*Understanding by design*, Alexandria, Virginia: Association for Supervision and Curriculum Development