Integrating Inquiry-Based Multimedia Learning Outcomes into Educational Accountability Systems

Steven McGee, Debra Panizzon, John Pegg, & Bruce C. Howard

Introduction

In the ongoing battle over school budgets in the United States, the word *accountability* has become the battle cry for those dissatisfied with current education. Most states in the U.S. mandate high-stakes testing of all students as a primary mechanism of school accountability (Kronholz, 1999). Some would argue that these tests are not a valid measure of student learning and should be abolished. However, in many states, the very constitutionality of the statewide-funding system hinges upon these high-stakes tests. There is no question that high-stakes testing is here to stay. Therefore, the Learning Sciences (LS) community is left with two alternatives for helping schools be accountable for participation in LS projects. We need to demonstrate that LS projects improve performance on high-stakes tests or, alternatively, we need provide better evidence that high-stakes tests do not provide accurate measures of students' higher-order thinking. This issue concerning the relationship between classroom teaching/learning and high-stakes testing has been identified by the Department of Education's National Center for Research on Evaluation, Standards, and Student Testing (CRESST) as a top priority goal (Baker, 1999).

Researchers at the NASA Classroom of the Future program in Wheeling, WV in collaboration with researchers at the University of New England in Armidale, Australia are directly addressing this issue. Of interest is the need for different types of assessment approaches that are aligned to both specific curriculum innovations and high-stakes tests. Furthermore, these approaches are guided by cognitive theories of learning and development, and are able to be administered efficiently by classroom teachers. This paper takes up this theme by describing and comparing two approaches to assessment that were used to identify the growth in students' understandings of planetary processes. The first approach belonged to the qualitative paradigm and sought to describe the underlying structure of students' responses in terms of a cognitive development model. The second approach encompasses the quantitative paradigm as a means of measuring conceptual understanding and problem solving. Consequently, this dual approach offers a unique lens through which to view student understanding in relation to classroom instruction.

Instructional Context

This study involved Years 5 and 6 students in Australia who participated in the *Astronomy Village®: Investigating the Solar System*™ project over a two-week period in August 1999. Through *Astronomy Village* students were transported to a virtual village in Hawaii where they investigated what the surface of Pluto might look like when the first NASA mission arrives in 2015. The program is designed such that a virtual mentor guides students in completing multiple investigation cycles that mirror the phases of scientific inquiry. In the first investigation cycle, students were introduced to the core research question concerning the surface of Pluto. During the exploration phase of the investigation, students were exposed to aerial images of various planetary features to prepare them for analyses of planetary images. In the background research phase, students read library articles and listened to lectures to help them understand key background concepts. During the data collection and analysis phases, students drew conclusions about the research question. This core investigation cycle lasted one week. In the second week, students followed the same sequence of phases as they did in the core investigation when they undertook a follow-up investigation. In one school, students investigated whether icy volcanoes could exist on Pluto by examining the surfaces of other icy bodies in the Solar System. In the other school, students explored whether plate tectonics might exist by investigating its existence on other rocky planets.

SOLO: A framework for response codings

SOLO, an acronym for the **S**tructure of the **O**bserved **L**earning **O**utcome, is a response model developed by Biggs and Collis in the late 1970s (Biggs & Collis, 1982). Since that time more than one hundred studies have been undertaken, both to apply and extend the model (Biggs and Collis, 1991; Pegg, 1992). In summary, SOLO provides a framework upon which the underlying structure of the answer to a stimulus question can be inferred from the response given. This theoretical perspective evolved through the identification that learners show a consistent pattern of development, over different topic areas and in different learning environments. Important variables that determine the quality of the response have much to do with the working memory available, the amount of information that can be retained, and features specific to the task.

The SOLO model has much in common with the broad raft of neo-Piagetian frameworks, such as those proposed by Case (1992) and Fischer and Knight (1990). Nevertheless, there are some significant differences between the ideas behind SOLO and those of Piaget (1954). The most important difference concerns the classification of a student's response to some stimulus using SOLO does not carry with it an implication that the response is typical of their stage of cognitive development, nor that this is necessarily age related. In particular, cognitive understanding is seen to have a more individual characteristic that is both content and context dependent.

Coding a student's response using the SOLO model depends on two features. The first feature concerns the nature or abstractness of the response and is referred to as the *mode of thinking*. This describes the type of intellectual functioning that is required to address a particular stimulus. As such, each mode has its own identity—its own specific idiosyncratic character. There are five modes of thinking. (1) *Sensorimotor* – a person reacts to the physical environment. For the very young child it is the mode in which motor skills are acquired. These play an important part in later life as skills associated with various sports evolve. (2) *Ikonic* – a person internalizes actions in the form of images. It is in this mode that the young child develops words and images which can stand for objects and events. For the adult this mode of functioning assists in the appreciation of art and music and leads to a form of knowledge referred to as intuitive. (3) *Concrete symbolic* – a person thinks through the application of a symbol system such as written language and number systems. This is the most common mode addressed in learning in the upper primary and secondary school. (4) *Formal* – a person considers more abstract concepts. This can be described as working in terms of 'principles' and 'theories.' Students are no longer restricted to a concrete referent. In its more advanced form it involves the development of disciplines. (5) *Post Formal* – a person is able to question or challenge the fundamental structure of theories or disciplines.

An implication of the model is that while the five modes of thinking are distinct and develop in the order provided above, the functioning in a later acquired mode (say, concrete symbolic) does not preclude the functioning in an earlier acquired mode (such as ikonic or sensorimotor). Also, for students in elementary and secondary education the target mode for instruction is primarily the concrete symbolic mode and teaching techniques are adopted generally to suit these learners (Collis & Romberg, 1991; Watson, Collis, Callingham, & Moritz, 1995). In general, students respond within the concrete symbolic mode although it is not uncommon for some students to respond in either the ikonic or formal modes. Because of this, it is these three modes that are the focus of this paper.

The second feature of the SOLO model depends on an individual's ability to handle relevant cues with increased sophistication. This feature is referred to as *levels of response*, which are seen to reside within cycles of learning that provide a hierarchical description of the nature of the structure of a response. While these levels occur within each mode, the specific nature of these levels is dependent on the particular mode targeted by the stimulus. Three levels make up a cycle of learning. (1) *Unistructural* – where the student focuses on the domain/problem, but uses only one piece of relevant data and so the response may be inconsistent. (2) *Multistructural* – where two or more pieces of data are used without any relationships perceived between them. No integration occurs of the data and some inconsistency may be apparent. (3) *Relational* – where all data are now available, with each piece woven into an overall mosaic of relationships. The whole has become a coherent structure. There is no inconsistency within the known system.

A particular focus of research using SOLO has explored the nature of student responses within a mode over an extensive range of questions within a topic. The result of several investigations (e.g., Campbell, Watson & Collis, 1992 Collis, Jones, Sprod, Watson, & Fraser, 1998; Levins & Pegg, 1993; Panizzon, 1999; Pegg, 1992) have identified at least two unistructural-multistructural-relational cycles within the concrete symbolic mode. Of particular interest within the intra-modal development is the characteristic that the relational response (R_1) in the first cycle becomes the unistructural element (U_2) in the second cycle.

Results and Discussion

Both assessment tasks were administered as pre- and posttests. The first assessment activity instructed students to, "Draw a picture in as much detail as possible showing what you think the surface of Pluto looks like." On the reverse side of the paper, students were then asked to, "In as much detail as possible, describe the processes that created the features that you drew." These written descriptions were coded by three researchers using the SOLO model. As the model is not applicable to visual responses, the drawings were not coded, however, they provided a useful context for interpreting some of the written descriptions.

The second assessment task was a multiple-choice test already developed for the *Astronomy Village* summative evaluation effort. Due to the large number of students involved, it was necessary to develop an assessment instrument that could measure inquiry skills in a cost-effective manner. Through a content analysis of *Astronomy Village*, it was determined that successful students were able to understand complex content as well as draw inferences about planetary processes from images of surface features. Item writers, not involved in the product development, designed multiple-choice items based on this task framework. Items were developed for each of the nine investigation cycles in *Astronomy Village*. The students in this study only completed the items related to the core and supplementary investigations.

SOLO Analysis

The responses provided by students were representative of both the ikonic and concrete symbolic modes. Within the responses designated in the ikonic mode, students provided descriptions of the various elements comprising their picture of Pluto including the color or the types of shapes that they used, e.g., “I drew a circle and then I coloured it with blue. Then I put white on it.” Clearly, this type of response is strongly visual where the student focused solely on explaining the physical attributes of their drawing with no attempt being made to incorporate scientific content in their responses.

In contrast, the majority of students demonstrated a higher degree of abstraction as they recognized that the symbols they had drawn on their diagrams represented real-world referents. This is characteristic of the concrete symbolic mode. It was possible to identify two cycles of levels of understanding within this mode.

First Cycle Responses

Within this cycle, students hypothesized the existence of a variety of planetary features on Pluto including the presence of volcanoes, impact craters, rivers, cracks or faults, folding mountains, glaciers, and sand dunes. In the first type of response, only one of these features was mentioned by the student, e.g., “I drew a cold deserty (desert) surface because I think it would be frosty and still on Pluto.” Within this response, the student focused on the cold and frosty nature of Pluto. This single emphasis is characteristic of a unistructural (U_1) response.

Alternatively, some students identified two or more features of Pluto that are believed to be distinctive of the planet. “The circular pictures show that I think there are craters and volcanoes. The lines around show a cold place and the blank lines show wind and cold.” This is a multistructural (M_1) response given that the student identified both landform (e.g., volcanoes and craters) and climatic features (e.g., cold and windy) of Pluto in the description. However, the student made no attempt to link the two features together.

In contrast, a small number of students used their knowledge of other planets as a means of comparing the features identified on Pluto, e.g., “I think Pluto is similar to Triton because they are both cold.” This response demonstrates a higher degree of understanding in that the student has recalled knowledge from one planet and used it as a model to organize the features found on Pluto. Responses of this type are denoted as relational (R_1).

Second Cycle Responses

A characteristic of the U_1-M_1-R_1 responses discussed above is that students merely focused on describing the various features of Pluto. However, it was possible to identify a series of responses in which students provided a causal explanation for each of the identified landform and/or climatic features of Pluto. Subsequently, a U_2 response is “Heaps of meadors (meteors) hit it with such force it created massive craters and gullies.” Clearly, this student recognized a cause and effect relationship between meteors and the craters found on the surface of Pluto.

However, other students discerned a number of these types of connections as a means of explaining how particular landform features arose where the connections for each of these features was dealt with independently. These were denoted as M_2, e.g., “I think Pluto will have craters, iced over rivers, mountains, volcanoes, holes, sand dunes, and more. The craters I think are created by showers of meteorites and cracks are created by wind. The sand dunes are created by wind erosion and iced over rivers are created by water that has frozen because of the weather.”

Finally, a small proportion of students used a planetary model already in their frame of reference to organize the causal explanations of the planetary features identified on Pluto. An example is “Comprehensive and extensive faults are scattered among the surface of Pluto, in a way that suggests that it is not plate tectonics as it does not form a jigsaw like pattern.

Volcanoes are scattered among its surface, caused by plume tectonics, on which lava flows up the cracks, and causes bulges in the planet's surface, thus causing volcanoes. Craters dot its surface, caused by meteorites' impact on the surface. It also has many glaciers as its surface is one of icy coldness, due to its distance from the sun." This ability to link the various components together using a unifying theme, which in this instance was plate tectonics as experienced on Earth, is characteristic of a relational (R_2) response.

The results for the SOLO pre- and posttest codings are summarized in Table 1.

Table 1 Summary of students' responses using SOLO categorizations

		Pretest	**Posttest**
No response		1	0
Ikonic mode		2	1
Concrete symbolic mode	U_1	10	3
	M_1	1	10
	R_1	0	1
	U_2	7	4
	M_2	10	10
	R_2	0	2

Most of the students' responses were coded into one of the six levels in the concrete symbolic mode. There were thirteen students (43%) who responded at the same level on both analyses, of these, eight were M_2 responses. Another, thirteen students (43%) improved their performance. Most of these students performed at the U_1 level on the pretest. On the posttest, seven students provided a M_1 response, and there was one in each of the R_1 and R_2 categories. Only four students (14%) gave lower-level responses on the posttest. A statistical analysis of pre/post performance using a Wilcoxon signed rank test revealed that although a significantly greater number of students improved than declined, the magnitude of the increase was not significant. It can be concluded that this implementation of *Astronomy Village* was effective at increasing many students' growth in cognitive reasoning as measured by the SOLO framework, but the magnitude of that growth is not as large as we would like to see.

Comparison Between SOLO and *Astronomy Village* Test

Student performance on the *Astronomy Village* test and the SOLO responses were compared using a Pearson product-moment correlation. To achieve this, the SOLO categories were encoded using 0 for ikonic mode responses and 1 to 6 for the levels in the concrete symbolic mode. The correlations were .438 between the pre SOLO category and the content pretest and .703 between the post SOLO category and the content posttest. These results suggest that the two assessment instruments are measuring highly related but separate constructs.

An analysis of pre/post differences provides further evidence that although these constructs are related, they have different developmental trajectories. The mean score for the content test increased from 37% at the pretest to 54% at the posttest. This difference was statistically significant ($t = 5.21$, $p < .01$). It can be concluded that this implementation of *Astronomy Village* effectively increased student conceptual understanding and problem-solving skills.

Conclusions

In this study, we compared two approaches to classroom assessment. Measuring cognitive reasoning using the SOLO framework provided mixed results from pre to posttest, whereas measuring conceptual understanding and problem solving using the *Astronomy Village* test showed significant results from pre to posttest. The results also indicate that there is a strong link between the coding of classroom assignments using SOLO and performance on the *Astronomy Village* test. We believe that these results provide a foundation for research on the relationship between classroom teaching and high-stakes testing. In the next step of our research, we plan to compare performance on the *Astronomy Village* test with performance on high-stakes testing. This research will allow us to either demonstrate improved performance from participation in *Astronomy Village* or it will allow us to provide strong evidence that high-stakes testing is not measuring higher-order thinking. Another direction for future research is an investigation of the influence and effectiveness of *Astronomy Village* on the development of cognitive reasoning. In the present study, there was improvement in conceptual understanding and problem solving but not cognitive reasoning. The next step is to expose students to a longer period of instruction by having them engage in four investigation cycles over a four-week period.

References

Baker, E. L. (1999, September). The current context for assessment: Standards, Quality, and Fairness. Presentation at the *CRESST National Conference*. Los Angeles, California.

Biggs, J. & Collis, K. (1982). *Evaluating the Quality of Learning: the SOLO Taxonomy*. New York: Academic Press.

Biggs, J. & Collis, K. (1991). Multimodal learning and the quality of intelligent behaviour. In H. Rowe (Ed.), *Intelligence, Reconceptualization and Measurement* (pp. 57-76). New Jersey: Laurence Erlbaum Assoc.

Campbell, K., Watson, J. & Collis, K. (1992). Volume measurement and intellectual development. *Journal of Structural Learning and Intelligent Systems, 11*, 279-298.

Case, R. (1992). *The Mind's Staircase: Exploring the conceptual underpinnings of children's thought and knowledge.* New Jersey: Lawrence Erlbaum Assoc.

Collis, K., & Romberg, T. (1991). Assessment of mathematical performance: An analysis of open-ended test items. In M.C. Wittrock & E. L. Baker (Eds), *Testing and Cognition* (pp. 82-130). Englewood Cliffs, New Jersey: Prentice-Hall Inc.

Collis, K., Jones, B., Sprod, T., Watson, J., & Fraser, S. (1998). Mapping Development in Student's Understanding of Vision using a Cognitive Structural Model. *International Journal of Science Education, 20* (1), 44-66**.**

Fischer, K.W. & Knight, C.C. (1990). Cognitive development in real children: Levels and variations. In B. Presseisen (Ed.), *Learning and thinking styles: Classroom interaction.* Washington: National Education Association.

Kronholz, J. (1999, October 29). Colorado School Districts Promise to Make the Grade. *Wall Street Journal*, p. A20.

Levins, L. & Pegg, J. (1993). Students' understanding of concepts related to plant growth. *Research in Science Education, 23*, 165-173.

Panizzon, D.L. (1999). Senior secondary and early tertiary science students' developmental understandings of diffusion and osmosis: A neo-Piagetian approach. Unpublished thesis for the degree of PhD, University of New England, Armidale, Australia.

Pegg, J. (1992). Assessing students' understanding at the primary and secondary level in the mathematical sciences. In J. Izard & M. Stephens (Eds), *Reshaping Assessment Practice: Assessment in the Mathematical Sciences Under Challenge* (pp. 368-385). Melbourne: Australian Council of Educational Research.

Piaget, J. (1954). *The Construction of reality in the child.* New York: Basic Books.

Watson, J., Collis, K., & Campbell, K. (1994). Developmental Structure in the Understanding of Common and Decimal Fractions. *Focus on Learning Problems in Mathematics, 17* (1), 2-25.

Watson, J., Collis, K., Callingham, R., & Moritz, J. (1995). A Model for Assessing Higher Order Thinking in Statistics. *Educational Research and Evaluation, 1* (3), 247-275.