Blurring the Lines: Learning *and* Assessing in Quadrant D
Some Key Lessons from Four Successive Years of Embedded Assessment Design in our Sustainable Earth Project

Ken Rose, Martin Block
Clemente Community Academy, Chicago Public Schools
The Center for Learning Technologies in Urban Schools, School of Education & Social Policy,
Northwestern University, 2115 N Campus Drive, Evanston, IL 60208
Tel: 847-491-7494 Fax: 847-491-8999
Email: rose@northwestern.edu, mblock@clementehs.k12.il.us

Abstract: This design paper describes some of the main characteristics of four successive iterations of an embedded assessment activity used in our annual "Sustainable Earth" project. The details shared here come mainly from a comparative analysis of our last two assessment design cycles, highlighting differences between two common activity structures: a culminating student presentation and a global summit debate. We found that reframing the task and attending to a few specific elements of the teachers' and students' roles both before and during the activity resulted in two significant impacts. First, the debate activity transformed the student learning that took place in terms of its rigor and relevance, making it more characteristic of what we describe as "Quadrant D learning." Second, the debate greatly increased our teachers' ability to assess their students' learning of desired content and skills, making it an even more effective embedded assessment activity than the final presentation alone.

Keywords: curriculum, assessment, instruction

Introduction/Framework

The work reported here is part of a larger research study that closely examines the complex decisions teachers make in the selection and interpretation of learning activities that can make their students' thinking more explicit for both formative and summative assessment. During this broader study, a team of teachers and researchers designed and piloted a series of four consecutive iterations of an embedded assessment activity for a long-term high school environmental science project called "Sustainable Earth." From 2000 to 2003, this project was conducted annually for 3-4 weeks each spring with ninth grade students at a large urban high school in the Midwest.

This particular design paper describes some of the main characteristics of the four successive iterations of the Sustainable Earth project, concentrating on a detailed comparative analysis of the last two assessment cycles, those taking place in the spring of 2002 and spring of 2003. It was in these final two years of implementation that the embedded assessment activity was most extensively changed—from a culminating student presentation about a specific country to a global summit of developed and developing countries. This global summit added a valuable component of structured debate to the final presentation, forcing the students to internalize the project's content, use the synthesis and communication skills they had been taught, and take on multiple perspectives of the countries they represented. Prior research has shown that such a debate, or argumentation, task can be worthwhile in increasing student learning (Snider & Schnurer, 2002; Bell, 1998), as varying perspectives actually strengthen the understanding of the content and its context. We claim here that such debate can be equally valuable as a technique for embedded assessment.

An explicit goal of our design research of novel instruction and assessment strategies has been to increase teachers' awareness and use of more types of formal and informal classroom assessment (Wilson & Sloane, 2000), and to determine how opportunities for such assessment can be improved. Black and Wiliam (1998) define classroom assessment as, "all those activities undertaken by teachers and by their students... that provide information to be used as feedback to modify the teaching and learning activities" in which they are engaged (p. 140). They conducted a survey of about 250 current assessment studies, published during or after 1988, and concluded that significant gains could be achieved by implementing full systems of classroom assessment that incorporate this kind of appropriate formative, as well as summative, function.

The Black and Wiliam review provides clear evidence that classroom assessments by teachers, combined with appropriate feedback to students, can have robust, positive effects on student learning and achievement. Learning gains from systematic attention to classroom assessment were found to be greater than most of those for any other educational intervention. Previous research supports these findings (Crooks, 1988; Fuchs & Fuchs, 1986).

Yet, as strong as those claims may be, classroom assessment practices are currently underdeveloped in most schools, and many unanswered questions remain. Teachers and teacher educators still wonder how to foster its use most effectively to promote student learning. To answer these kinds of questions, the specific design decisions that shape successful classroom assessments need to be examined much more closely in real-world contexts. This study does just that.

There is also a larger issue to be addressed. Researchers have indicated that these more informal classroom assessments are often not even viewed as a significant form of assessment by many teachers, principals, parents, or the general public (Atkin, Black, & Coffey, 2001). Developing it as a valid pedagogical skill is therefore seldom seen as a priority for overburdened classroom or professional development time. This research seeks to change that view.

In this paper, we describe the broader research context and provide a brief overview of the Sustainable Earth project. Then, we share some specific lessons learned from the design decisions we made, based upon our analysis of student learning compared across the last two instances. In general, each successive assessment design provided new constraints and affordances for us to gain insight about our students' understanding, and offered new opportunities for the activity to serve simultaneously as a valuable learning activity as well as an assessment activity. The findings from this ongoing assessment design project suggest specific improvements for our own future instruction as well as important ways to inform our ongoing study of classroom-based embedded assessment.

The design lessons presented here are important because culminating presentations continue to be a very common activity structure for teachers and curriculum designers to use for assessment in project-based learning. However, most final presentations generally lack an adequate amount of rigor and relevance in terms of the academic demands they place on students. In addition, they do not afford opportunities for continued student learning; that is, they are typically not embedded into the instruction, but are instead tacked on at the end of a project as a summative assessment.

We will demonstrate with this example that reframing the task and attending to a few specific elements of the teachers' and students' roles before and during the activity can have at least two significant impacts. It can transform the student learning that takes place in terms of its rigor and relevance, making it more characteristic of what we describe as "Quadrant D learning." It can also greatly increase a teachers' ability to assess what students know and can do, making it a more effective assessment activity.

Method

Research Context

The participants in our larger, ongoing study are part of an urban public high-school-within-a-school project known as the Mathematics-Science-Technology Academy (MSTA). MSTA teachers are collaborating with university researchers in the reiterative design and implementation of computer-supported, collaborative, inquiry-based projects for their students in two Chicago high schools (Kemeny & Kwon, 2002). This study of the assessment design in the Sustainable Earth project was taken from work done in one of those MSTAs each spring, from 2000 through 2003. The school is a medium-sized neighborhood high school, with approximately 800 students and 45 teachers, located in a large urban center in the Midwest.

Two classes of approximately 25 ninth-graders each participated in all four iterations of the project, with a total of 200 students over the course of four years. The freshmen teachers saw each group five times a week, for 45-minute periods, although the periods were sometimes blocked for 90 minutes.

Overview of the Sustainable Earth Project

The Sustainable Earth project grew out of earlier efforts by the Center for Learning Technologies in Urban Schools (LeTUS) and some MSTA teachers at nearby high schools. In the mid-1990s, LeTUS produced a

curriculum called Global Warming, which helped middle-school students learn about the effects of the differential heating and cooling of the earth's surface (see http://letus.org/globalwarming.htm for complete details on this curriculum). An MSTA teacher adapted that original curriculum for his ninth graders, adding more population content and deliberately bridging across disciplines to include more social studies. In our subsequent revisions of the curriculum, we added a driving question to kick the inquiry off. We began by posing the following broad problem to the students:

> Exponential human growth and development are threatening Earth's ability to support life. The unsustainability of natural systems and natural cycles is evident, and human development is the cause. Our reliance on nonrenewable sources of energy, our misuse of water and land resources, and our uncontrolled population growth all threaten the natural sustainability of the Earth. Do we control our fate? How can we recreate a more sustainable Earth?

After discussing the driving question, students were divided into groups of 2-4. Each group represented the governing body and heads of state for a particular country from different regions in the world (Mexico, Brazil, China, Japan, India, Germany, Switzerland, Kenya, South Africa, and Sudan). Each country has different needs, different degrees of development, and different problems. However, all countries will be affected if humans continue to disrupt the Earth's natural balances.

During the project the students completed several instructional activities to prepare them for a culminating assessment at the end of the project. They completed a pamphlet and a newsletter to describe different aspects of the country they were assigned. Their research focused upon three major topics: (a) Sources of Energy and Energy Consumption; (b) Water and Land Quality; and (c) Population Growth.

The Sustainable Earth project always had its focus in environmental science, but it involved the participation of the other core subjects—math, social sciences, and English—as well. For example, the math teacher helped the students use linear and exponential equations to project their countries' populations into the year 2050. With these data, they could represent their countries at the summit by making recommendations about the need for laws or programs to control population growth and making decisions about the use of natural or renewable resources over the next 50 years.

In the first three years, our students did oral presentations at the end of their investigation where they provided specific information about their countries. They were given lists of specific data to include in their presentations. This included type of government, demographics, economy/GNP, literacy rate, population growth rate, life expectancy, sources of energy, carbon emissions, and deforestation rates. Students focused on the three main themes of energy, land and water, and population growth in light of each nation's politics, economics, health, education, family life, religion, popular arts, and human rights.

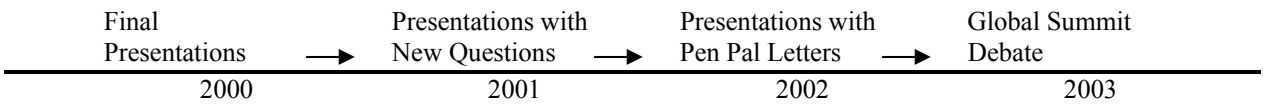

Figure 1: Activity Structure Redesign Timeline

Every year, after reflecting on the depth of student learning evident in the project presentations, we decided we needed to do some revisions to the culminating activity (see Figure 1). We were struggling to find a way to get the students to engage with rigorous content and take some ownership over the ill-defined problems being studied. We were trying to find out how much scaffolding they needed—and what kind—in order to successfully see the connections and interrelationships in the issues and nations represented. In the second year, therefore, they were given additional lists of thought-provoking questions to which they should be prepared to respond during their presentations. These questions were very specific, such as: "At the current rate of growth, how large will your population be in 20 years?" or "What is being done in your country to protect natural resources?" In the third year, a new component was added to the final assessment that provided students with a chance to take on more of their country's point of view. We asked that each group write five pen pal letters where they pretended to be a person

from the country they were assigned to represent. In these letters, they talked about the homes (building materials, type of structures), education system, transportation, tourist sites, recreation, climate, currency, average income and occupations.

In the fourth year, still frustrated at the lack of deep learning going on, both before and during the presentations, we changed the entire activity structure to that of a global United Nations summit, where students had the opportunity to share and debate their findings with students representing two or three other countries. The summit consisted of students' responding to larger open-ended questions for each of the three major topics. First, each group described its country's current practices. Second, it stated how its country's practices harm its populace and the global community. Third, it developed a plan to effectively solve any environmental threats. Finally, after each group had a chance to speak on all three of the important environmental topics, all countries in the summit had to come to a consensus and agree upon a piece of environmental legislation that will help restore the Earth's sustainability.

There was a rubric used to assess the students' understanding of the issues, but we found that often the activity eventually seemed to be less about the formal assessment of what the students knew and much more about continued learning about the connections and interrelationships between the issues and the countries. Indeed, this summit activity turned out to be a valuable embedded assessment task, and much more of a true learning experience than anything that had come before. We decided to look more closely at the summit activity, to find a way to describe the changes as well as to isolate some of the important design elements that made it a more successful embedded assessment.

Data Collection and Analysis

Our comparative cases are based on primary data collected during each of the project iterations, each spring from 2000 to 2003. We took field notes of classroom observations and our conversations with students. We also videotaped the final assessment activities and collected samples of student work. Finally, we have conducted interviews with participants and one-on-one design meetings between the participants, including the first two authors– a high school classroom teacher and a university researcher involved closely in the project.

One specific tool we found useful for describing the changing academic demands of the task is the Rigor/Relevance Framework (Daggett & Kruse, 1999). This Framework is based on traditional elements of education yet encourages movement to the application of knowledge instead of maintaining an exclusive focus on the acquisition of knowledge. It is based on two concurrent dimensions of higher standards and student achievement (see Figure 2).

First, there is the Knowledge Taxonomy. Located on the vertical axis, it is a continuum based on the six levels of Bloom's Taxonomy, which describes the increasingly complex ways in which we think. It describes the rigor demands of the activity. The low end involves acquiring knowledge and being able to recall or locate that knowledge. The high end labels the more complex ways in which individuals use knowledge, such as taking several pieces of knowledge and combining them in both logical and creative ways. The second continuum is known as the Application Model. Located on the horizontal axis, it is a spectrum of action or relevance. Its five levels describe putting knowledge to use. While the low end is knowledge acquired for its own sake, the high end signifies use of that knowledge to solve complex real-world problems and to create unique projects, designs, and other works for use in real-world situations.

The Rigor/Relevance Framework has four distinct quadrants. Each is labeled with a term that characterizes the learning or student performance at that level. A "Quadrant A" activity could be choosing the right definition for new vocabulary. A "Quadrant B" activity could be comparing a car lease to a car loan. A "Quadrant C" activity could be analyzing symbolism in a poem. A "Quadrant D" activity could be developing guidelines for publishing content on the Internet. This Rigor/Relevance framework provided us a systematic way to analyze curriculum, instruction and assessment across all of the subjects involved in the Sustainable Earth project. We were able to examine the different iterations of the Sustainable Earth assessment task, and characterize them in terms of the academic demands placed on students.

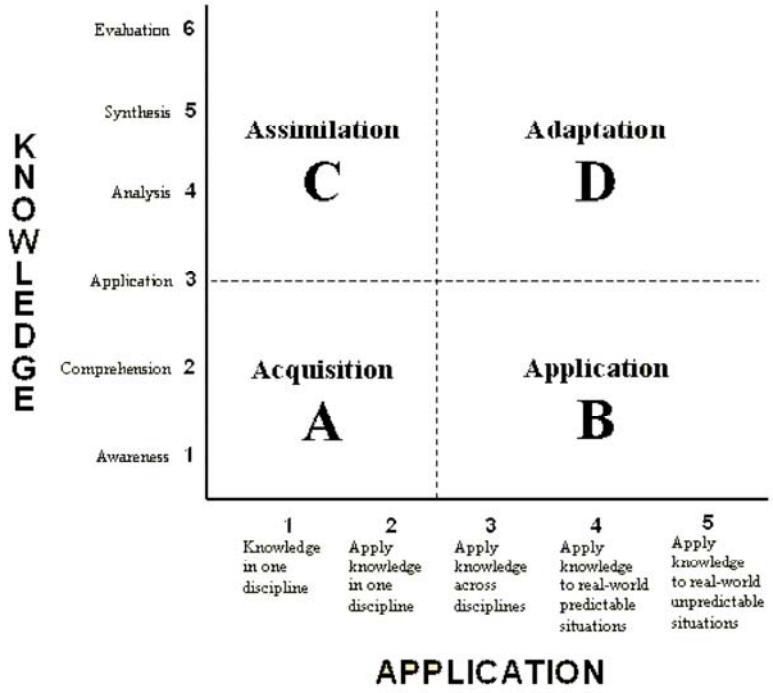

Figure 2: The Rigor/Relevance Framework

While analyzing the assessment activity data in terms of Quadrants A, B, C, and D, we found that we naturally began to attend to some other related dimensions of the project as well. We characterized levels of student ownership of the problem and involvement with the content during the final presentations in 2002 and the global summit debate in 2003. We then described the level of teacher scaffolding of the discussion, the effect of additional opportunities for students to practice beforehand and the changing role of the teachers during the debate. Finally, we looked closely at the effect of taking on the role of the country, or the effect of the students' own perspective.

Results and Discussion/Implications

In reviewing the video clips of the students' final assessments from the last two implementations, one thing stands out clearly. The students who took part in the global summit debate demonstrated a deeper and more contextualized understanding of their country's role in creating a more sustainable earth than the students who made the more formal final oral presentations. The use of the summit debate clearly encouraged more rigorous and relevant learning both before and during the assessment activity. As we looked for a way to describe and explain the differences we saw, we uncovered some rather important design elements.

During the presentations, the students demonstrated learning that we characterized as Quadrant A Learning. They had gathered and stored bits of knowledge and information for their presentations. They were primarily expected to remember or understand this knowledge, but not to use it. The final presentations consisted of lists of facts about each country, memorized and shared with the audience. The presenting students, without exception, followed the order in the guidelines that had been set for them in the task; they had prepared powerpoint slides and often read directly from the slides or from notecards. Questions from audience members, when asked, were surface-level questions like "What is the population of your country?" or "Who is the leader?" In this context, it was the teacher who directed all of the learning, and who had the stake in the learning.

However, during the global summit debates, students clearly demonstrated what we termed Quadrant D Learning. They had the competence to think in complex ways and to apply their knowledge and skills. Even when confronted with perplexing unknowns, they were able to use their extensive knowledge and skills to create solutions that further developed their knowledge. Unlike the final presentations, in which the teachers asked students for

specific information so that they could check it off on our rubrics, the summit was much more like an open-ended conversation, grounded in the main issues presented in the project, but broad enough for students to take ownership of the learning.

The move to the debate structure and the transfer of this ownership did not take place automatically. It required a conscious effort on our part to let go. And this was difficult to do. The first few debates we facilitated were in fact very similar to the final presentations, only now the students were seated instead of standing in front of an audience and they were in a different formation with two or three other countries. Not much else was different in those early debates. It seemed very necessary for our teachers to probe for information with questions like "What about when you burn the fossil fuels, what kind of problem does that cause?" We found that, the more the teacher held back with these types of direct questions, though, the more student-to-student discussion took place. After a while, we realized that encouraging students to negotiate and discuss their respective needs provided richer discussions.

The teacher had to play a much more subtle role during the debate than during the presentations. She had to keep reminding the students how they compared with the other nations, modeling the connections between them. This is the most crucial role for the adult in the summit task. In the end, this is what makes it a learning activity as well as an assessment activity. In later debates, we did start to see evidence of students' referring to and building on other students' comments. They responded to what was being said and used information from other groups. Yet this skill had to be modeled deliberately first.

Another less subtle role for the teacher was to remind students to pay attention to what the other country's representatives were saying. The students were encouraged to take notes during the debate, and in subsequent years, were asked to write a formal agreement following the debate. We also added practice sessions so that they could be coached before the summit took place. These sessions were very helpful for the students to gain confidence and flexibility in using their data. We wish we had videotaped them during the practice sessions to compare with the final sessions.

In terms of taking on the role of the country, there were marked differences between the two sets of assessment activities in the number of times the students personalized the content. Those students in the global summit referred to their country's challenges with phrases such as "our population" and "our energy sources." They discussed their solutions with language that showed this perspective; they owned the problems. It was not uncommon to hear them admit: "we have to work on that." However, the students doing formal oral presentations spoke almost completely in the third person. They reported on the current conditions of Germany and China, for example, without taking on the roles or perspectives they were given. They said "Germany should look into alternative fuels" or "China has a policy of..."

As we continue to design and pilot more truly embedded assessments that measure the type of Quadrant D learning we desire, we will strive to keep many of these simple lessons in mind. Above all, we must remember that in order for this kind of assessment to be successful, we must have a real trust in our students; we must believe that they can take ownership of their own learning and demonstrate the deep content understanding and skills we would like to see, as long as we provide them both the freedom and the scaffolding they need to learn.

References

Atkin, J.M., Black, P., & Coffey, J. (Eds.) (2001). *Classroom assessment and the National Science Education Standards*. Washington, DC: Committee on Classroom Assessment and the National Science Education Standards, Center for Education, National Research Council.

Bell, P. (1998). Designing for students' science learning using argumentation and classroom debate. Berkeley, CA, University of California

Black, P., & Wiliam, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan, 80*(2), 139-148.

Crooks, T. J. (1988). The impact of classroom evaluation practices on students. R*eview of Educational Research*, *58*(4), 438-481.

Daggett, W., & Kruse, B. (1999). *Taming the educational dinosaur*. Rexford, NY: Leadership Press.

Fuchs, L. S., & Fuchs, D. (1986). Effects of systematic formative evaluation: A meta-analysis. *Exceptional Children*, *53*(3), 199-208.

Kemeny, V., & Kwon, S. (2002). MSTA (Mathematics, Science, and Technology Academy) model of teacher-researcher collaboration. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.

Snider, A.C., & Schnurer, M. (2002). *Many Sides: Debate across the curriculum*. Burlington, VT: IDEA Books.

Wilson, M., & Sloane, K. (2000). From principles to practice: An embedded assessment system. *Applied Measurement in Education, 13*(2): 181-208.