INSTRUCTIONALLY SUPPORTIVE NCLB TESTS IN SCIENCE:

AN ILLUSTRATION

W. James Popham
University of California, Los Angeles

During the 2007-08 school year, at the latest, science achievement tests are to be administered at three grade ranges in all U.S. public schools according to the No Child Left Behind Act (NCLB). Whether those new science tests will actually enhance students' science skills and knowledge, however, depends almost entirely on the nature of the tests selected to satisfy this significant federal requirement.

Instructionally Supportive Accountability Tests

In late 2001, an independent national commission issued a report describing the nature of large-scale accountability tests that could (1) provide credible and accurate evidence of educators' instructional effectiveness and (2) stimulate improved instruction. That report, prepared by the Commission on Instructionally Supportive Assessment, is available on this web site.

In brief, the Commission identified three necessary attributes of an instructionally supportive accountability test. First, such a test had to measure only a modest number of genuinely significant skills (so as to not overwhelm teachers with too many assessment targets). Second, the skills to be assessed had to be described in teacher-palatable language so that teachers could understand clearly what was to be tested and, therefore, could aim their instruction at the skills and knowledge represented by a test, not at the actual items on that test. Finally, the results of an instructionally supportive accountability test had to be reported in a manner that would permit teachers to determine accurately which of their curricular aims had been achieved by students and, therefore, which parts of their instruction had/hadn't been effective (so that the ineffective instructional segments could be improved).

NSES Implications

One of the more influential documents in science education is the *National Science Education Standards* NSES) published by the National Academy Press (Washington, D.C.) in 1996. In that report, considerable attention is given to the importance of promoting students' mastery of science-as-inquiry. And, although NSES is so chock-full of curricular aspirations for the nation's science teachers that it tends to overwhelm even the zealous, the continuing emphasis on the significance of science-as-inquiry in NSES attests to the importance of this skill.

Actually, NSES breaks down science-as-inquiry into two sub-skills, that is, "abilities necessary to do scientific inquiry" and "understandings about scientific inquiry." Moreover, if one reviews the NSES document carefully at various grade ranges, there appear to be a set of evaluative criteria present that could be used to judge if a student had, indeed, mastered this skill, In other words, if a student were given some sort of constructed-response task intended to reveal the student's mastery of science-as-inquiry, these evaluative factors could be employed to determine how well the student had actually mastered the science-as-inquiry skill.

The availability of those NSES evaluative criteria seems to provide a useful contrast between how one might conceptualize and assess a student's mastery of science-as-inquiry for an NCLB science test and how one might measure a student's skill in composition for an NCLB language arts test.

Parallels Between Prowess in Composition and Prowess in Science-as-inquiry

Students who are good at written communication will generally be able to create original written communications that warrant our praise. Not only will students who possess prowess in written communication be capable of generating various types of written materials, such as expository or persuasive essays, those students will also possess a positive disposition toward the writing process itself. That is, most composition-skilled students will recognize that they are competent writers and, while perhaps not actually enjoying the act of writing, will be confident that, if required, they can write effectively.

A student who has mastered the writing process will have acquired a number of sub-skills that contribute to the quality of a final piece of writing. Not only does the student have to understand and organize effectively the content of what is to be written about, but also must display satisfactory mastery of such basics as word-usage, capitalization, punctuation, and spelling. Written communication is, in addition to reading, one of the cornerstones of a language arts education. Also, written communication, for many adults, is one of the foundations of a successful career and/or life.

Similarly, students who are masters of science-as-inquiry will generally be able to display understandings of both scientific concepts as well as the nature of science itself. Moreover, such students will possess the skills necessary to become independent inquirers about the natural world. And, finally, students who are proficient in science-as-inquiry will typically display dispositions to employ science as inquiry skills when they tackle science-related tasks. Putting it in the words of others, such students will "develop the habits of mind associated with science" (State Science Education Supervisors as cited by Brett Moulding in an October 3, 2003 personal communication).

The nature of science-as-inquiry is well explicated in NSES. On page 105 of that report, science-as-inquiry is described as moving beyond the "processes of science" such as observation, inference, and experimentation toward a vision that requires students to use such processes and scientific knowledge as the use of scientific reasoning and critical thinking to develop their understanding of science. Clearly, the authors of NSES regard science-as-inquiry as an integral, if not the most important, element of science education.

All right, we have two potent skills and the affective corollaries of those skills. It occurs to me that there are clear parallels between what we have learned from English teachers who employ "the writing process" as they teach kids to write, and the way that science teachers can teach their students to master science-as-inquiry. In a nutshell, here's what I see as the most important similarities.

What skilled English teachers do when they promote their students' writing prowess is provide an evaluative framework by which students can judge the caliber of their own writing (as well as the writing of others). These frameworks (or scoring guides) are referred to as "rubrics." If the rubric is properly formulated, it will contain about a half-dozen evaluative criteria (for example, "effective word-usage") that can be called on to appraise the caliber of an in-process or completed piece of writing. An evaluative criterion is simply a factor that's used in order to judge the caliber of a student's performance. Students who understand the nature of a composition rubric, therefore, possess a cognitively comprehended, evaluative framework for judging their writing. Students who truly understand these rubrics are then given as much guided and independent practice as possible in using those evaluative frameworks. For instance, effective English teachers not only ask students to generate their own compositions for teacher-grading, but also call on students to function as peer critics whereby they judge their classmates' writing according to the evaluative rubric being used. In addition, students can be required to apply their evaluative framework to "the classics," that is, pieces of writing regarded as superlative by our society. (One of the advantages of such critiquing activities is that students can learn how great writers sometimes depart from a conventional evaluative framework and yet turn out a wonderful piece of writing. Evaluative frameworks such as those provided in a rubric, in other words, are guidelines, not sacrosanct strictures.)

The more practice that emerging writers have in employing their evaluative frameworks, the more facile they become in using those frameworks. And that increasing skill soon translates into heightened confidence and a positive disposition toward writing.

When stripped down to its essentials, the promotion of students' writing prowess hinges on the provision of a defensible evaluative framework (in the form of a rubric) and ample guided/independent practice in writing.

And that's precisely what I think we should do with respect to promoting science students' mastery of a science-as-inquiry skill. We need to devise a powerful evaluative rubric that focuses on a modest number of evaluative criteria (for instance, defensibility of an investigation's design), get students to understand that rubric, and give them a substantial amount of guided and independent practice in applying the rubric. As in the case of the teaching of writing, science teachers need to have students work solo, in groups, and constantly employ the "science-as-inquiry rubric" to judge the science-related activities under consideration. (By applying the rubric widely, in particular to classic *atypical* investigations of the past, it can be demonstrated that there is no "one way to do science.")

As with the teaching of composition skills, the more adept that students become at using their emerging science-as-inquiry prowess, the more positively they will regard that approach to inquiry. Positive attitudes toward a subject are generally correlated with a student's mastery of that subject.

Thus, I think that developers of NCLB science tests need to get cracking on coming up with a sensible rubric to help guide students as they carry out their science-as-inquiry endeavors. The rubric dare not be too stuffed with content. Its evaluative criteria, however, though few in number, will surely embrace several sub-skills---but not too many! In that regard, it seems apparent that we must isolate a sufficient amount of science-related knowledge so that it will be possible to evaluate the accuracy of the student's use of that knowledge. I suggest that we could focus on three content arenas, namely, (1) physical sciences, (2) life sciences, and (3) earth and space sciences. In each of these three content arenas we could isolate a certain amount of arena-specific content that students would have to learn, and that we could then incorporate as one of the rubric's evaluative criteria. This would be akin to the problem facing the developers of composition rubrics when they list as one of their rubric's criteria "proper use of mechanics." Okay, mechanics might refer to spelling, punctuation, and capitalization. However, an instructionally defensible rubric will spell out (perhaps in a separate appendix) *which* spelling is fair game, *which* punctuation rules must be adhered to, and *which* capitalization conventions are to be observed. That's the sort of content delineation we would need to see so that science-as-inquiry does not become a general-purpose cognitive capability that has nothing really to do with science.

Science-as-inquiry According to NSES

I've looked repeatedly at the way that the NSES authors break out science-as-inquiry at the three grade ranges. Although I found troubling the fact that in the explanations about the meaning of a particular ability (that is, sub-skill) there were far too many instances in which subcomponents of one ability appeared as a subcomponent of other abilities, it does seem possible to isolate five basic abilities that show up in all three grade ranges. I'm going to list them below in just a moment. But if you look carefully at this content in NSES, you'll find that different language is often employed at different grade ranges. Yet, stripping aside such differences, I think there are five abilities that would constitute suitable foci for both instruction and assessment. And when one tries to build an instructionally supportive accountability test, one must try to isolate ingredients that not only can be properly measured, but also can be well taught. I have presented these five evaluative criteria below.

Under each of the five evaluative criteria that, in concert, underlie the student's skill in employing science-as-inquiry, I have listed several subordinate factors that might be employed to appraise the adequacy of a student's response with respect to a particular evaluative criterion. That response would be made to some sort of performance task calling for the use of a science-as-inquiry approach. Each evaluative criterion is numbered and each evaluative criterion's subordinate factors are bulleted.

A Potential Science-as-Inquiry Rubric's Evaluative Criteria

1. Identify questions that can be answered by scientific investigations

   - Clarity of questions
   - Whether questions are directed toward objects or phenomena that can be described, explained, or predicted

2. Design and conduct a scientific investigation

   - Quality of the investigation's design
   - Quality of the investigation's implementation

3. Use of appropriate ways to gather, analyze and interpret data

   - Suitability of tools and techniques
   - Accuracy of analyses
   - Defensibility of interpretations

4. Development of descriptions, examples, and predictions

   - Caliber of evidence-based support

5. Communication of procedures and conclusions

   - Clarity and accuracy of communications
   - Audience-appropriateness of communications

If developers of NCLB science tests were ever were to move in a direction such as this, there would be loads of clarifying and revising required. I tried to derive my ideas from NSES recommendations at the three relevant grade ranges, but there would definitely need to be some variations at the different grade ranges. For example, there is considerable attention given in NSES to alternative explanations and models as children grow older. Decisions would need to be made regarding whether that sort of sub-skill should be worked into the mix.

The skills we end up with, for they will be fewer in number, must be unarguably important. Indeed, if I had my way, the skills developers of NCLB science tests would end up with would be viewed by laypeople as "knock your socks off, high level science skills."

If developers of NCLB science tests were to use performance tasks as a key component of this approach to assessment, they'd have to be truly inventive. Not only would they need to make sure that any tasks were developmentally appropriate, but they'd have to figure out how to supply the students with as authentic a context as possible. In addition, students would be *given* the necessary information, that is, the information needed for skill-mastery demonstration or would have had to *already learn* such content. It seems clear that there would need to be a defined body of content (for example, in physical science) that would be required in order for students to carry out the task successfully and, thereby, display their ability to employ science-as-inquiry.

NCLB Tests That Help

Students' scores on NCLB science tests are not to be used in the "adequate yearly progress" requirements that now form the heart of a particular school's annual evaluation. However, it should be obvious that the skills and knowledge measured by NCLB tests will have a profound effect on the say that science is taught in this nation. If students in a school or district score well on these tests, applause will follow. If students' scores are low, the opposite result will take place. Teachers and school administrators will be heavily influenced by what's measured on any state-level science tests that satisfy the assessment requirements of NCLB.

That being the case, I hope desperately that the architects of these important new science assessments will try to build tests that are instructionally supportive. Tests that assess only a few truly important skills and/or bodies of knowledge, describe what they assess, and supply instructionally informative results are instructionally supportive. Such NCLB tests will help our students master what they need to learn in science. NCLB tests that are not instructionally supportive, however, definitely won't.

November, 2003