LeRoy, K., & Lee, O. (in press). What research says about science assessment with English language learners. Arlington, VA: National Science Teachers Association.

This work is supported by the National Science Foundation (NSF Grant #ESI–0353331). Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the authors and do not necessarily reflect the position, policy, or endorsement of the funding agencies.

What Research Says about Science Assessment with English Language Learners

As high-stakes testing and accountability in science looms on the horizon, science teachers are faced with the dilemma of identifying key science concepts and effective instructional practices to maximize achievement for all students. Education reform and specifically the *No Child Left Behind* [NCLB] Act of 2001 require that *all* students achieve high academic standards in core subject areas. Assessment of science achievement necessitates consideration of fairness to different student groups. Fairness in this context means "the likelihood of any assessment allowing students to show what they understand about the construct being tested" (Lawrenz, Huffman, & Welch, 2001, p. 280). How do we ensure that assessments are valid and equitable for all students? With English language learners (or ELL students), assessments should distinguish among academic achievement, English language proficiency, and general literacy. Although large-scale assessments may include various accommodation strategies, they are rarely administered in languages other than English. Even when assessments are administered in students' home language in addition to English, ensuring the comparability of assessment instruments between two languages is complicated.

This chapter addresses what research says about science assessment with ELL students. Specifically, we draw from our ongoing research and development efforts to promote science and literacy achievement of ELL students in a large urban school district (visit our project website at http://www.education.miami.edu/psell). First, we describe our intervention with a focus on inquiry-based science teaching and learning. Second, we describe how we design science and literacy assessment instruments for ELL students in our research. Third, we describe our efforts to align classroom assessments in science and literacy with high-stakes assessments for ELL students. Finally, based on our work and other research findings, we offer suggestions for what teachers can do to assess science and literacy achievement of ELL students.

Science Instructional Intervention with ELL Students

What science instructional practices will support increased science and literacy achievement of ELL students? Some empirical evidence has demonstrated success with inquiry-based science; however, there is a huge variation on what that looks like in today's science classrooms. With both content standards in place and inquiry at the cornerstone of what science curriculum should include, how can this be translated into quality teaching in today's classrooms that meets the needs of all students?

Through university and school district collaborative research, "Promoting Science among English Language Learners (P-SELL) in a High-Stakes Testing Policy Context," we implement an instructional intervention to promote ELL students' science and literacy achievement in the context of high-stakes testing and accountability. Our intervention involves teachers and students at grades 3 through 5 located in 15 elementary schools in a large urban district. These schools enroll greater-than-district proportions of ELL students and students from low socioeconomic status (SES) backgrounds, and have performed poorly according to the state's accountability plan. The research tests two conventional wisdoms: (1) Can ELL students learn academic subjects, such as science, while also developing English proficiency? and (2) Can ELL students, who learn to think and reason scientifically, also perform well on high-stakes assessments?

Research on science instruction with ELL students highlights hands-on, inquiry-based science in enabling these students to develop scientific understanding and acquire English language proficiency simultaneously (Amaral, Garrison, & Klentschy, 2002; Lee, Deaktor, Hart, Cuevas, & Enders, 2005; Rosebery, Warren, & Conant, 1992). First, hands-on activities are less dependent on formal mastery of the language of instruction, thus reducing the linguistic burden on ELL students. Second, hands-on activities through collaborative inquiry foster language acquisition in the context of authentic communication about science knowledge and practice. Third, inquiry-based science promotes student communication of their understanding in a variety of formats, including written, oral, gestural, and graphic. Fourth, by engaging in the multiple components of science inquiry, ELL students develop their grammar and vocabulary as well as their familiarity with scientific genres of speaking and writing (Lee, Buxton, Lewis, & LeRoy, 2006). Finally, language functions (e.g., describing, hypothesizing, explaining, predicting, and reflecting) can develop simultaneously with science inquiry and process skills (e.g., observing, describing, explaining, predicting, estimating, representing, inferring) (Casteel & Isom, 1994). By engaging in inquiry-based science, ELL students learn to think and reason as members of a science learning community.

Essential to our intervention are the project-developed curriculum units and instructional practices that promote science inquiry and English language development of ELL students. Our intervention emphasizes the use of the *Inquiry Framework* (see Figure 1) that supports science inquiry with diverse student groups. Although some advocates of more open-ended, student-centered inquiry would argue against a framework for organizing and planning inquiry, our practical experience as urban science educators with ELL students has demonstrated the importance of such a framework as an initial step for teachers and students. While making the inquiry process explicit, the framework also allows for flexibility to foster students' initiative and responsibility for their own learning. The icons in the framework serve as points of reference for assisting students in thinking about and organizing their own inquiry. The icons also encourage the use of graphic representations in communicating science, especially for ELL students and students with limited literacy development (Lee et al., 2005). In using the framework, it is important to recognize that the process is not to be followed as a lock-step procedure, but to be considered as a guide. The way students engage in inquiry will vary depending on their prior experience with science, their level of literacy development, and the kinds of questions that are promoted through interaction and discussion.

Our intervention also scaffolds student initiative and responsibility in conducting inquiry as teachers gradually reduce their level of guidance. The National Research Council (2000) presents a continuum of teacher direction to learner self-direction as it relates to essential features of science inquiry. The intervention is designed to move progressively along the teacher-explicit to student-initiated continuum to promote science inquiry. Initially, students may need a great deal of assistance to engage in inquiry. As they develop inquiry skills, they will need less and less assistance. Eventually, they can explore and do inquiry on their own. As students engage in multiple components of inquiry, they learn to engage in some areas more easily (e.g., implementing activities and reporting results), while they require more assistance and experience in other areas (e.g., questioning and applying findings). The *Science Inquiry Matrix* (see Figure 2) illustrates this continuum, as teachers gradually relinquish authority and encourage students to assume responsibility for inquiry.

Science inquiry should also involve such conventions as control of variables, use of multiple trials, and accuracy of measurement. Controlling variables will allow students to establish the cause and effect relationship between the variable being tested (i.e., independent variable) and the variable that is being measured (i.e., dependent variable). Multiple trials enhance reliability of the results, and this can be achieved by having groups of students or individual students do an experiment and then compare the results. Furthermore, students should strive for accuracy and precision when they use the tools of science for measurement (i.e., metric rulers, balances, graduated cylinders, thermometers). These conventions are integral to the science inquiry framework (see Figure 1), as teachers gradually withdraw assistance and students learn to take initiative and assume responsibility for conducting inquiry (see Figure 2).

How to Design Science and Literacy Assessment Instruments for ELL Students

In our research, we use multiple approaches of science assessment with ELL students that include (a) a project-developed pre-post science test directly aligned to the curriculum, (b) a writing prompt that addresses a key science concept covered within the curriculum, and (c) a reasoning interview protocol that engages students to conduct a performance task. Although these instruments are developed in English to maintain continuity between the language of instruction and the language of assessment, teachers are instructed to use students' home language if needed and students are allowed to use their home language.

Science Test

We developed a pre-post science test for each grade level. The test measures students' knowledge of key concepts and big ideas of the science topics in the curriculum that is taught during the school year. The test also measures students' understanding of science inquiry by asking them to construct graphs and tables using the data provided, offer explanations for the data, and draw conclusions. The test consists of project-developed items, public-release items from the state science assessment, public-release NAEP items, and public-release TIMSS items. Item formats include multiple-choice items and short and extended written response items, with many of the items having multiple components. For the project-developed items, we have developed a scoring rubric. For public-release items from the state science assessment, NAEP, and TIMSS, we use the available scoring rubrics. The maximum points for each specific item or item subcomponents range depending on the level of cognitive or conceptual difficulties.

Writing Prompt

We developed a pre-post expository writing prompt to assess both literacy (writing) and science content. The writing prompt on the water cycle for third grade students and the scoring rubric is presented in the Appendix. Students receive two scores on the writing prompt: (a) *form* measures students' use of conventions, organization, style, and voice; and (b) *content* measures students' ability to give scientific explanations and to use scientific vocabulary. The distinction between form and content allows teachers to assess science learning and English language proficiency separately.

Reasoning Task

We developed a reasoning interview protocol for each grade level. Reasoning interviews are conducted individually and videotaped. The topics include measurement with third grade students, force and motion with fourth grade students, and earth systems with fifth grade students. In addition to measuring students' conceptions of each science topic, reasoning interviews measure students' ability to apply measurement concepts using scientific tools, to conduct a scientific experiment on a force and motion task, and to engage in scientific discourse on an earth systems task. The three reasoning tasks represent critical elements of students' ability to engage in science inquiry (National Research Council, 2000). The performance-based reasoning interviews with individual students complement the larger-scale assessment of science test and the writing prompt being used with all students in our research.

Assessment Results

These three assessments, together, enable us to examine the first research question—Can ELL students learn academic subjects, such as science, while also developing English proficiency? Using hierarchical linear modeling (HLM) analysis, the results of our first-year intervention with third graders revealed that students in the treatment group displayed a statistically significant increase in science achievement test scores at the end of the school year (see Table 1; detailed results are reported in Lee, Maerten-Rivera, Penfield, LeRoy, & Secada, in press). Furthermore, the results of the HLM analysis showed that the students who were enrolled in English to Speakers of Other Languages (ESOL) programs made achievement gains comparable to the students who had exited from ESOL or never been in ESOL. Over the five-year period of our research, we will continue to examine the research question using longitudinal data.

How to Align Classroom Assessment in Science and Literacy with High-Stakes Assessments for ELL Students

The research on science assessments, especially in the context of high-stakes assessments with diverse student groups including ELL students, is currently limited. Science until very recently has not been part of high-stakes assessments in most states. Even when science is tested, it usually does not count toward accountability. As a result, research on science assessment *accommodations* with ELL students is even more limited. As science becomes a part of accountability measures under the NCLB Act, more research is expected.

A primary motivation for our research involves ongoing concerns about low science achievement of ELL students, especially given the national context of the impending high-stakes testing policy in science under the NCLB Act. When the third grade students in our research advance into fifth grade in 2006, they will be the first cohort of students for whom the statewide science assessment will factor into school accountability.

The State assessments are administered in reading, mathematics, writing, and science. Reading and mathematics are assessed in grades 3-10 and writing and science are assessed once at the elementary, middle, and high school levels. The State science assessment includes multiple choice and performance task items in the physical and chemical sciences, the earth sciences, the life and environmental sciences, and scientific thinking. Items on the State science assessment encourage student reasoning, planning, analyzing, and using scientific thinking. Students should be proficient in determining the logical steps or outcome in an experiment; comparing or contrasting structures or functions of different organisms or systems; applying and using concepts from a scientific model or theory; drawing conclusions; analyzing an experiment to identify a flaw and propose a method for improving it; and predicting outcomes as a result of a change within a system.

Our ongoing research is enabling us to examine the second research question—Can ELL students, who learn to think and reason scientifically, also perform well on high stakes assessments? Following the first year of our intervention with third graders, the treatment group students showed higher scores on the State mathematics test, particularly on the measurement strand emphasized in the intervention, than the comparison group students (detailed results are reported in Lee et al., in press). Of the eight point maximum score for the measurement strand, the mean of the treatment group was 5.00 (*SD* = 1.91) compared to the mean of 4.39 (*SD* = 2.02) in the comparison group. We will continue examining this question during the final three years of our research (2006 through 2009) when the state science assessment will be in place.

What Teachers Can Do to Assess Science and Literacy Achievement of ELL Students

An important aspect of classroom assessment includes the use of meaningful and relevant topics, tasks, and activities. Teachers can employ assessment practices for ELL students, which may serve to benefit all students in a classroom. First, teachers may use two separate scoring criteria for writing prompts or short and extended response tasks to assess science learning and English language proficiency separately (see Appendix). This assessment practice enables teachers to identify strengths and weaknesses of ELL students in each area (Lee et al., 2005).

Second, teachers may assess ELL students in their home languages as well as in English. Allowing students to communicate their science knowledge and abilities in their home languages promotes both general literacy and academic learning which, in turn, promotes English language proficiency. Achievement in these three areas develops simultaneously (Lee & Fradd, 1998).

Third, teachers should promote the use of multiple representational formats, keeping in mind that the goal is to move ELL students toward established literacy standards. Students who cannot write in either home language or English can express ideas in drawings, graphs, and tables, as well as in oral communication. When students realize that they are expected to produce meaningful representations of their knowledge in assessment settings, they engage in science learning activities and tasks in more meaningful ways.

Finally, teachers who are aware of linguistic and cultural influences on ELL students' responses support the assessment process. Whereas efforts have traditionally focused on eliminating these influences, an emerging approach advocates understanding how home language and culture can be incorporated to guide the entire assessment process (Solano-Flores & Trumbull, 2003).

Conclusions

Since instruction and assessment complement and reinforce each other, it is essential to provide high quality instruction for ELL students and to asses their achievement outcomes in a manner that can guide subsequent instruction. Science assessment with ELL students presents both promises and challenges for their science learning and English language development. Assessment in ELL students' home language in addition to English enables them to express their science knowledge and abilities; however, this does not ensure valid and equitable assessment if the language of instruction is in English. As science becomes a part of accountability measures across the states under the NCLB Act, accommodations for ELL students should be addressed. Unfortunately, research on science assessment accommodations with ELL students is insufficient to guide large-scale assessments.

In this chapter, we describe our own ongoing research and development on science instructional intervention and assessment with ELL students. Preliminary results indicate that (a) ELL students can learn science while developing English proficiency and that (b) ELL students can engage in science inquiry and reasoning while performing well on high-stakes assessments. The emerging literature in this field can offer insights for valid and equitable science assessment of ELL students, which, in turn, can be used to further enhance science instruction for all.

References

Amaral, O. M., Garrison, L., & Klentschy, M. (2002). Helping English learners increase achievement through inquiry-based science instruction. *Bilingual Research Journal, 26*, 213-239.

Casteel, C. P., & Isom, B. A. (1994). Reciprocal processes in science and literacy learning. *The Reading Teacher, 47,* 538-545.

Lawrenz, F., & Huffman, D., & Welch, W. (2001). The science achievement of various subgroups of alternative assessment formats. *Science Education, 85*(3), 279-290.

Lee, O., Buxton, C., Lewis, S., & LeRoy, K. (2006). Science inquiry and student diversity: Enhanced abilities and continuing difficulties after an instructional intervention. *Journal of Research in Science Teaching, 43*(7), 607-636.

Lee, O., Deaktor, R. A., Hart, J. E., Cuevas, P., & Enders, C. (2005). An instructional intervention's impact on the science and literacy achievement of culturally and linguistically diverse elementary students. *Journal of Research in Science Teaching, 42*(8), 857-887.

Lee, O., & Fradd, S. H. (1998). Science for all, including students from non-English language backgrounds. *Educational Researcher, 27*(3), 12-21.

Lee, O., Maerten-Rivera, J., Penfield, R., LeRoy, K., & Secada, W. (in press). Science achievement of English language learners in urban elementary schools: Results of a first-year professional development intervention. *Journal of Research in Science Teaching*

National Research Council. (2000). Inquiry and the national science education standards: A guide for teaching and learning. Washington, DC: National Academy Press.

No Child Left Behind Act of 2001. Public Law No. 107-110, 115 Stat. 1425. (2002).

Rosebery, A. S., Warren, B., & Conant, F. R. (1992). Appropriating scientific discourse: Findings from language minority classrooms. *Journal of the Learning Sciences, 21(1)*, 61-94.

Solano-Flores, G., & Trumbull, E. (2003). Examining language in context: The need for new research and practice paradigms in the testing of English-language learners. *Educational Researcher, 32*(2), 3-13.

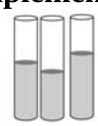
Inquiry Framework	
1.Questioning 	**State the problem** ▪ What do I want to find out? (written in the form of a question) **Make a hypothesis** ▪ What do I think will happen?
2. Planning 	**Make a plan by asking these questions (think, talk, write)** a. What materials will I need? b. What procedures or steps will I take to collect information? c. How will I observe and record results?
3. Implementing  	**Gather the materials** ▪ What materials do I need to implement my plan? **Follow the procedures** ▪ What steps do I need to take to implement my plan? **Observe and record the results** ▪ What happens after I implement my plan? ▪ What do I observe? ▪ How do I display my results? (using a graph, chart, table)
4. Concluding 	**Draw a conclusion** ▪ What did I find out? ▪ Was my hypothesis correct or incorrect?
5. Reporting 	**Share my results (informal)** ▪ What do I want to tell others about the activity? **Produce a report (formal)** ▪ Record what I did so others can learn. ▪ Consider different ways to express my information.

Figure 1. Inquiry framework

Inquiry levels	**Questioning**	**Planning**	**Implementing**	**Concluding**	**Reporting**
0	Teacher	Teacher	Teacher	Teacher	Teacher
1	Teacher	Teacher	*Students*	Teacher	*Students*
2	Teacher	Teacher	*Students*	*Students/ Teacher*	*Students*
3	Teacher	*Students/ Teacher*	*Students*	*Students*	*Students*
4	*Students/ Teacher*	*Students*	*Students*	*Students*	*Students*
5	*Students*	*Students*	*Students*	*Students*	*Students*

Figure 2. Science inquiry matrix

Table 1

Descriptive Statistics for Science Test Scores (total possible score = 24 points)

Variable	Test	Subgroup	*N*	*M*	*SD*
All Students	Pre		818	7.40	3.36
	Post		818	14.34	4.30
	Gain		818	6.95	4.18
ESOL	Pre	ESOL levels 1 to 4	118	6.55	3.40
		ESOL exited or non-ESOL	698	7.53	3.34
	Post	ESOL levels 1 to 4	118	12.39	4.62
		ESOL exited or non-ESOL	698	14.67	4.16
	Gain	ESOL levels 1 to 4	118	5.84	4.08
		ESOL exited or non-ESOL	698	7.14	4.17

Appendix

Writing Prompt: The Water Cycle

Pretend you are a drop of water. Before you begin writing, think about how water changes form in the water cycle. Explain to the reader how you are changed as you go through the water cycle.

Writing Scoring Rubric for Form

The writing scoring rubric for form considers the following components:

Convention
- Spelling
- Correct plurals and comparisons
- Capitalization and punctuation
- Subject/verb agreement and verb and noun forms

Organization
- Indentation for new paragraphs
- Idea development

Style/Voice
- Sentence structures to communicate ideas
- Coherence from sentence to sentence

4 – Complete/Comprehensive

- Spells all high frequency and most irregular words correctly with up to 3 errors
- Uses correct plurals and comparisons (e.g., good, better, best)
- No errors in capitalization and punctuation
- Subject/verb agreement and verb and noun forms are generally correct with up to 2 errors
- Uses accurate indentation for each new paragraph
- Ideas are presented logically
- Uses a variety of sentence structures to communicate ideas
- Writing is highly coherent

3 – Adequate

- Spells most high frequency (up to 3 errors) and many irregular words (with up to 5 errors) correctly
- Generally uses correct plurals and comparisons (e.g., good, better, best) (with up to 1 error)
- Few errors (up to 3) in capitalization and punctuation
- Occasional errors in subject/verb agreement, which do not impede communication (with up to 3 errors)
- Generally uses accurate indentation for the majority of each new paragraph
- Ideas are developed fairly logically

- Uses a variety of sentence structures to communicate ideas, although most (more than half of all sentences) are simple constructions
- Writing is fairly coherent

2 – Emerging/Expanding

- Some errors in spelling (up to 8 errors)
- Some errors in plurals and comparisons (e.g., good, better, best) (up to 3 errors)
- Some errors in capitalization and punctuation (up to 5)
- Errors in subject/verb agreement may somewhat impede communication
- Writing may not be organized into paragraphs
- Ideas are partially developed
- Sentence structures are limited to simple constructions with little attempt at variety
- Writing is partially coherent

1 – Minimal/Inaccurate

- Frequent errors in spelling (9 or more errors)
- Frequent errors in plurals and comparisons (e.g., good, better, best) (4 or more errors)
- Frequent errors in capitalization and punctuation (more than 5 errors)
- Lacks subject/verb agreement; verb and noun forms are incorrect and impeded communication
- No attempt to organize writing into paragraphs
- Ideas are not developed
- Sentence structure is fragmented/incomplete; lack of proper sentence structure may impede communication
- Writing lacks coherence

0 – No response/Unscorable

- The response is simply a rewording of the prompt
- The response is simply a copy of published work
- The student refused to write
- The response is illegible
- The response contains an insufficient amount of writing to determine if the student was addressing the prompt

Writing Scoring Rubric for Science <u>Content</u>

Elements of Writing:

Since the prompts require expository writing, the scoring rubric considers science vocabulary and explanation. Use of science vocabulary and a comprehensive explanation include the following components:

- scientific vocabulary
- three states of water (solid, liquid, gas, ice water, water vapor)

- three processes of change during the water cycle (evaporation, condensation, and precipitation)
- heating and cooling related to the water cycle
- concept of a cycle (continuous, ongoing, repetitive)
- sequence of events

Ratings:

4 – Complete/Comprehensive
- accurate use of science vocabulary with adequate explanations
- correctly names all three states of water as related to the water cycle
- accurately describes all three processes of change during the water cycle
- mentions process of both heating and cooling
- mentions concept of a cycle and the ongoing, repetitive nature of the process
- does not demonstrate sequencing errors

3 – Adequate
- accurate use of science vocabulary with adequate explanations
- correctly names all three states of water as related to the water cycle
- accurately describes all three processes of change during the water cycle
- may mention process of heating or cooling but does not include both elements
- mentions the concept of a cycle
- does not demonstrate sequencing errors

2 – Emerging/Expanding
- expanding accurate use of scientific vocabulary with some explanation
- correctly names two of three states of water as related to the water cycle
- accurately describes two of three processes of change during the water cycle
- does not mention process of heating or cooling
- does not mention the concept of a cycle
- may demonstrate sequencing errors

1 – Minimal
- inaccurate use of scientific vocabulary or minimal accurate use of scientific vocabulary without explanation
- accurately names only one of three states of water as related to the water cycle
- accurately describes one of three processes of change during the water cycle
- does not mention process of heating or cooling
- does not mention the concept of a cycle
- inaccurate description of sequence of events

0 – No response, unidentifiable, and/or irrelevant content