LEP Projects Report 6

Grade-Level Standards-Based Science Outcomes for English Language Learners and Language Minority Students: A Review of the Literature

NATIONAL
CENTER ON
EDUCATIONAL
OUTCOMES

In collaboration with:

Council of Chief State School Officers (CCSSO)
National Association of State Directors of Special Education (NASDSE)

LEP Projects Report 6

Grade-Level Standards-Based Science Outcomes for English Language Learners and Language Minority Students: A Review of the Literature

Kristin Kline Liu

April 2009

Liu, K. K. (2007). *Grade-level standards-based science outcomes for English language learners and language minority students: A review of the literature* (LEP Projects Report 6). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.

NATIONAL
CENTER ON
EDUCATIONAL
OUTCOMES

The Center is supported through a Cooperative Agreement (#H326G050007) with the Research to Practice Division, Office of Special Education Programs, U.S. Department of Education. Additional support for targeted projects, including those on English language learners, is provided by other federal and state agencies. The Center is affiliated with the Institute on Community Integration at the College of Education and Human Development, University of Minnesota. Opinions expressed herein do not necessarily reflect those of the U.S. Department of Education or Offices within it.

IDEAs that Work
U.S. Office of Special Education Programs

NCEO Core Staff

Martha L. Thurlow, Director
Deb A. Albus
Jason R. Altman
Manuel T. Barrera
Laurene L. Christensen
Christopher J. Johnstone
Jane L. Krentz
Sheryl S. Lazarus
Kristi K. Liu
Ross E. Moen
Michael L. Moore
Rachel F. Quenemoen
Christopher Rogers
Dorene L. Scott
Vitaliy Shyyan
Miong Vang
Yi-Chen Wu

National Center on Educational Outcomes
University of Minnesota • 207 Pattee Hall
150 Pillsbury Dr. SE • Minneapolis, MN 55455
Phone 612/626-1530 • Fax 612/624-0879
http://www.nceo.info

Introduction

Science is now entering the spotlight, after a time of standards-based education reform focused almost solely on increasing reading and math achievement for all students. There are many different sets of science standards that could be used to guide instructional decision making in a K-12 classroom. The publication of *A Nation at Risk* in the 1980s raised national concern about a possible crisis in the state of U.S. public education. In response, in 1993 the American Association for the Advancement of Science (AAS) published *Science for All Americans* (the predecessor to *Benchmarks for Science Literature*), followed a few years later by the National Research Council's publication of the *National Science Education Standards* (NRC, 1999). Scientists, science educators, and policymakers participated in creating these visions of what higher standards for science teaching and learning could look like. An analysis of these "national" standards showed that there are 147 K-12 science standards between the two documents, with a high degree of similarity in content (Marzano & Kendall, 1998; Saderholm & Tretter, 2008).

During that same time period, concerns over a widening achievement gap between subgroups of students spurred a call for greater educational equity (McClure, 2005). The 1994 reauthorization of Title I legislation, as well as the No Child Left Behind Act in 2001, mandated that every state create challenging content standards and assessments to measure student attainment of those standards in reading, math, and science. Assessments were required first in reading and math, and later, in 2007, for science.

Science for All Americans and the *National Science Education Standards* have received a great deal of attention, and have had a lasting impact on the ways teachers define good science teaching (Lauer, Snow, Martin-Glenn, Van Buhler, Stoutemyer, & Snow-Renner, 2005). Both sets of standards emphasized the importance of learning scientific inquiry processes. In doing so, these standards required educators to shift from a teacher-directed classroom where a teacher transmitted scientific facts to students, to a more student-directed classroom in which students were active participants constructing their own learning.

In spite of the importance of the *National Science Education Standards* and *Science for All Americans*, state science standards, as mandated by federal legislation, are the ones that directly impact daily life in schools and classrooms because of the accountability requirements attached to them. While state standards may bear some resemblance to national science standards, there may be areas of difference as well. Starting in 2007, all states are required to assess student attainment of science standards at least once in grades 3 through 5, once in grades 7 through 9, and once in grades 10 through 12. Patterns of differential science achievement between subgroups of students (e.g., English language learners, students with disabilities, students in various ethnic or racial categories, students from low socioeconomic backgrounds, migrant students, etc.) are cause for concern and must be addressed through changes in programs, services, and curricula as well as through instruction. According to Lee and Luykx (2007) this new accountability for science teaching and learning has an unprecedented power to change the field.

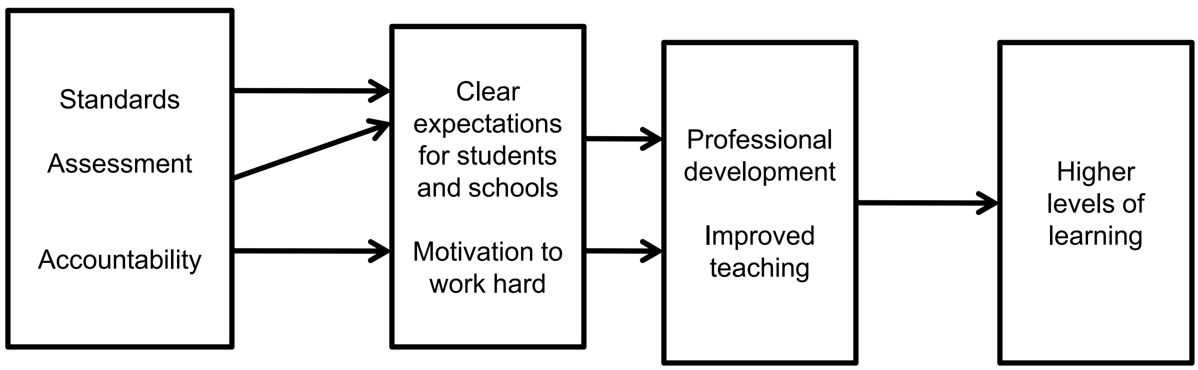
In many states and school districts, the first round of state science test scores have recently been made public and discussion of potential changes in science instruction are already underway. As we stand on the brink of a time of change and look ahead, English language learners (ELLs) are a group that warrants specific attention in the science classroom. Because science instruction is primarily conducted in English—the language that ELLs are learning—English proficiency is a major factor related to science outcomes for these students (Lee & Luykx, 2007). As Lee (2004) states, "A challenge facing teachers of students acquiring English as a new language (English language learners; ELLs) is enabling these students to learn academic content across subject areas, as the students simultaneously acquire English language and literacy" (p.65). Given the challenges that we know exist for educators instructing these students, the question arises, what do we know about their outcomes in standards-based science at this point in time? This review focuses on what peer-reviewed research literature tells us about standards-based science outcomes for ELLs, along with the broader group of students who speak a language other than English at home (also known as language minority students) and who may have been an ELL at one time. Earlier reviews addressed the issues of science for diverse students (Lee, 2005; Lee & Luykx, 2006, 2007) and standards-based science outcomes for native-English speaking students (Lauer, Snow, Martin-Glenn, Van Buhler, Stoutemyer, & Snow-Renner, 2005). No previous review has combined the two topics.

Theoretical Framework for this Review

Federal legislation that mandates content standards and accountability is based on a broad policy-level model of educational improvement. That model, the Expanded Theory of Action of Standards-based Reform: An Education Improvement System (NRC, 1999), is shown in Figure 1.

Figure 1 communicates the idea that content standards clearly define what all students should know and what schools should teach (NRC, 1999). Standards also act as benchmarks for assessing student attainment of the standards. Standards-based assessments give educators, students, and parents information about student progress toward that attainment. According to the NRC (1999), the most useful tests for this purpose are those that are sensitive to instruction. Teachers are given freedom to design instructional programs, but they must be accountable for student learning as measured by those standards-based assessments (NRC, 1999). Incorporated into the Theory of Action is the belief that teachers will use effective instructional practices if they have flexibility to do what they think best and if they are motivated to work hard. The assumption is that if teachers are motivated, they will change classroom teaching so that all students can learn, instead of assuming that some students will not learn. In addition to motivation, teach-

Figure 1. Expanded Model of the Theory of Action of Standards-Based Reform: An Education Improvement System (NRC, 1999)

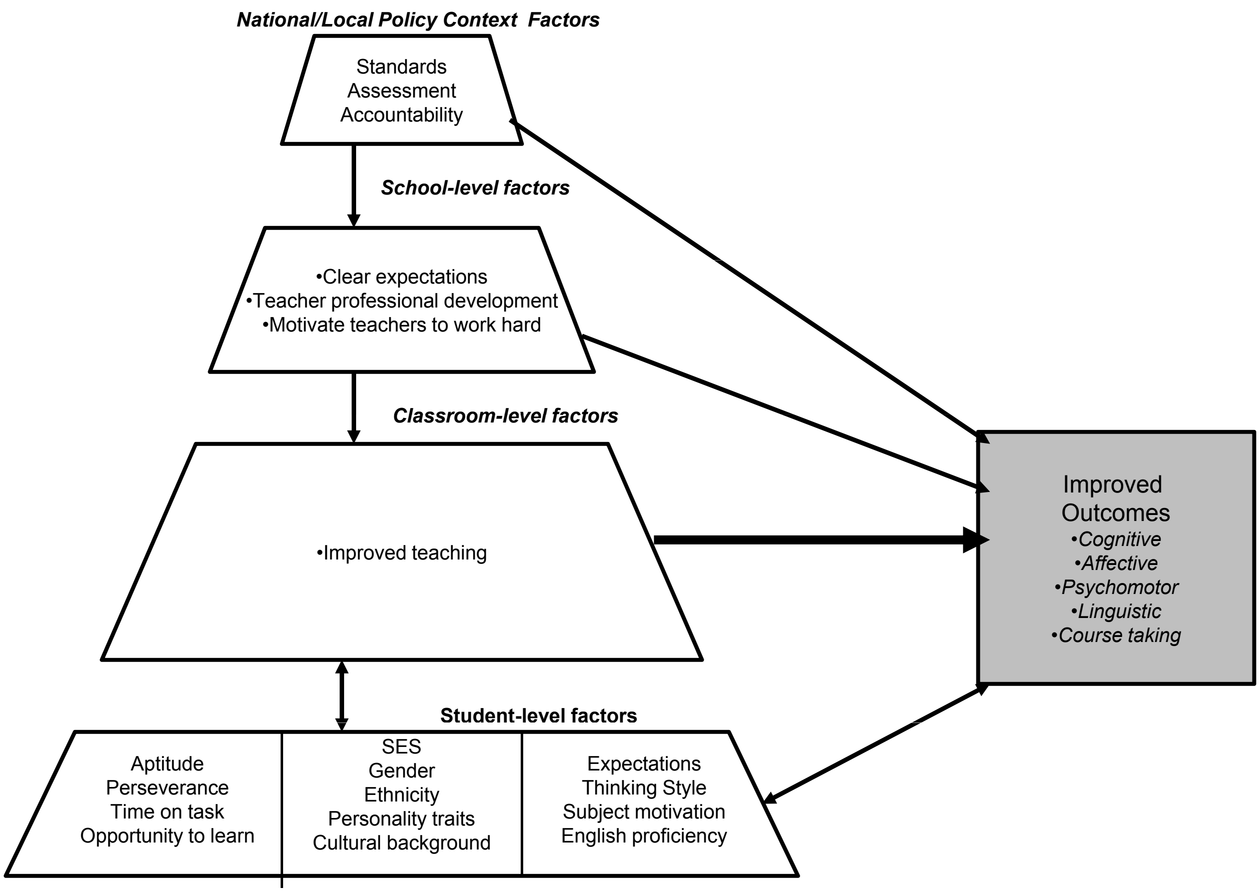


ers need professional development to support them in learning new practices. In this theory of action, when the different components of reform are all in play at the same time, higher levels of learning result (NRC, 1999).

The Expanded Model of the Theory of Action of Standards-based Reform is useful for guiding instructional policy at a broad level. However, it is less useful for guiding educational research on standards-based reform for particular groups of students because it does not reach down into the classroom. This theory of action also does not account for who linguistically and culturally diverse students are and how their characteristics influence instruction. For the purposes of this paper, an adapted version of Creemer's Comprehensive Model of Educational Effectiveness (Creemers & Kyriakides, 2008) incorporates elements of the Expanded Theory of Action of Standards-based Reform (NRC, 1999) into a multi-level model better suited for research purposes. Some additional factors have been added to the model to account for the learning of ELLs and language minority students who speak a language other than English (e.g., cultural background, English proficiency). Figure 2 shows the adapted model which is titled the Theory of Educational Effectiveness Under Standards-Based Reform.

The key attribute of this model is that multiple levels within the education system are represented and each one creates conditions that support the level beneath it (Creemers & Kyriakides, 2008). At the top of the diagram, national and state policy on standards, assessment, and accountability include both federal and state legislation, state content standards, and standards-based assessments for accountability purposes. At the school level, rules and policies about standards-based instruction in an individual building (e.g., grading policies, mandatory classroom routines or instruction practices, use of technology, etc.) combine with teacher professional development to create conditions that support teachers as well as motivating them to work hard.

Figure 2. Revised Theory of Educational Effectiveness under Standards-based Reform (Adapted from NRC 1999; Creemers & Kyriakides 2008)



Motivated teachers go in to the classroom where elements related to high quality instruction (e.g., curriculum use, teacher behavior, grouping procedures, etc.) come into focus. Teacher practice, as represented by the double-ended arrow between student factors and the classroom, is influenced by characteristics of the students in the class. Good teachers adapt instruction to meet the instructional needs of their students. At the same time, high quality instruction may influence characteristics of the students, such as their perseverance, their likes and dislikes, the time they spend focused on learning tasks, their content area language proficiency, and their overall opportunity to learn in the classroom.

The classroom-level factors, as shown by the bold arrow, are the most important contributors to improved student outcomes, but as in Creemers' original model (Creemers & Kyriakides, 2008), all of the levels in the system have a role to play. Student outcomes incorporate higher levels of learning, but in this model outcomes are defined broadly, as suggested by Lee and

Luykx (2007), to also include cognitive, affective, psychomotor, linguistic, and other types of outcomes such as science course taking patterns.

A word of caution is in order here about the differences between a school effectiveness model and a standards-based model. Creemers' original Comprehensive Model of School Effectiveness identified student factors as possible predictors of student outcomes (Creemers & Kyriakides, 2008, p. 43). In a standards-based system educators need to thoroughly understand the characteristics of the students they serve in order to design appropriate instruction, but the emphasis is on supporting the achievement of all students regardless of their characteristics and background. As former Secretary of Education Rod Paige once said in reference to the power of standards-based reform to increase outcomes for all students:

> We must let go of the myths and perceptions about who can learn and who can't… Anyone who's been in education as long as I have has heard all the excuses. 'Those kids' are too poor. 'Those kids' are too disadvantaged. We're doing the best we can with 'those kids.' What they're really saying is: We don't believe intellectual heft is in the DNA of poor and minority children…Teachers who believe that certain social groups are slower to learn and react by lowering the bar for performance rob those children of opportunities to grow intellectually and achieve their dreams. (Paige, 2003)

These words are applicable to the instruction of ELLs and language minority students as well. We cannot help them meet standards in science if we do not first believe they are capable of meeting them with appropriate instruction.

In Figure 2, as in Creemer's original model, the key elements that relate to higher levels of learning are the quality of instruction, time on task, and opportunity to learn (Creemers & Kyriakides, 2008). These items are bolded at the level in which they have been placed in the diagram, but every level of the system plays a role in developing conditions that sustain them.

The Revised Theory of Educational Effectiveness under Standards-based Reform (Figure 2) acts as a guide for this review of the standards-based science outcomes literature for language minority students and ELLs. It suggests that studies addressing any one or combination of the levels within the system (national/state policy, school, classroom, and student) are important to examine because they all, jointly, influence science outcomes for students. At the same time, the model emphasizes the central role of classroom instruction. Furthermore, the model suggests that, while cognitive outcome as measured by science test scores may be prioritized in standards-based reform, other kinds of related outcomes must be considered as well. Affective outcomes such as attitudes toward science and psychomotor outcomes such as science inquiry skills (e.g., using scientific tools to take measurements) are also deserving of attention.

Content of Standards

When discussing standards-based science outcomes for language minority students and ELLs it is important to consider what states actually require students to know and be able to do. Table 1 shows the broad categories addressed by 7th grade science standards in five states that were selected because they have sizeable populations of language minority students and ELLs.

Table 1. Science Standards from 5 Selected States

	State #1	State #2	State #3	State #4	State #5
Science Standards Focused on Content Knowledge	Structure and Function in Living Systems	Life Science	Relationship Between Structure and Function in Living Systems	Understand the Fundamental Concepts, Principles and Interconnections of the Life, Physical and Earth/Space Sciences	Life Science
	Cell Biology		Relationship Between Organisms and the Environment		
	Genetics				
	Evolution				
	Physical Principles in Living Systems				
		Physical Science	Equilibrium		Physical Science
			Force and Motion		
			Physical and Chemical Properties of Substances		
			Interactions Between Matter and Energy		
			Alterations in Earth Systems		Earth and Space Science

Table 1. Science Standards from 5 Selected States (continued)

	State #1	State #2	State #3	State #4	State #5
				Understand the Relationships Among Science, Technology and Society in Historical and Contemporary Contexts	Nature of Science
					Characteristics of Scientific Knowledge
					Roles of Theories, Laws, Hypotheses and Models
Science Standards Focused on Procedural Knowledge	Investigation and Experimentation	Scientific Thinking and Practice	Use Inquiry Methods	Understand the Processes of Scientific Inquiry and Technological Design to Investigate Questions, Conduct Experiments and Solve Problems	Nature of Science
			Use Tools and Methods to Conduct Inquiry		
			Use Critical Thinking and Scientific Problem Solving to Make Informed Decisions		
			Use Safe, Environmentally Appropriate, Ethical Practices		

From this table it is possible to see that in all five states there are two major categories of science standards: those that relate to knowledge of specific science content (declarative knowledge) and those that relate to knowledge of science processes (procedural knowledge). The majority of these states require all 7th graders, including language minority students and ELLs, to learn concepts in the areas of biology/life sciences, physical sciences (including chemistry and physics), and earth science. The states in the table appear to define these areas somewhat differently. For example, state 2 has a broad standard in the area of Life Science while State 3 has more detailed standards that address the relationship between structure and function in living systems or the relationship between organisms and the environment.

The five states in the table also require demonstration of scientific processes. Two states specifically mention the ability to perform scientific inquiry procedures as part of their standards, while others mention inquiry procedures in the more detailed benchmarks not shown here. Although definitions of inquiry may vary as well, inquiry often includes skills involved in identifying a problem to study, formulating hypotheses or research questions, designing and implementing a research study, and collecting, analyzing and communicating data (cf. Stoddart, Pinal, Latzke & Canaday, 2002). Inquiry procedures can be relatively unstructured and student-led or they can be structured in advance by the teacher (Cuevas, Lee, Hart & Deaktor, 2005; Lee, 2003; Lee, Buxton, Lewis & LeRoy, 2006).

In addition to the common content and process standards, a few of the five states shown in Table 1 list additional standards in either science content or processes. For example, State 5 requires students to know the characteristics of scientific knowledge and the roles of theories, laws, hypotheses, and models. In contrast, State 4 requires students to understand the historical relationships between science and technology.

While standards related to scientific language are not immediately apparent in the table, knowledge of scientific vocabulary and discourse patterns is embedded into the more detailed descriptions of what students need to know and do to achieve each standard. These detailed descriptions are often called benchmarks and several benchmarks may relate to one broader standard. For example, state 3 has a broad standard requiring 7th graders to understand the properties, structures, and processes of living things, as well as the interrelationship of living things and their environment. Under that standard state 3 includes the following benchmark:

> Know how to classify organisms: domain, kingdom, phylum, class, order, family, genus, species.

In this state, language minority students and ELLs, along with native English speaking students, would need to learn the vocabulary words "domain", "kingdom", "phylum", "class", "order", "family", "genus", and "species" in order to correctly classify organisms.

In state 3 the following standard not only requires students to learn particular vocabulary items, but also asks them to internalize a Western scientific way of thinking about catastrophic events and to communicate that Western explanation.

> Science concepts. The student knows that natural events and human activity can alter Earth systems. The student is expected to:

(A) Describe and predict the impact of different catastrophic events on the Earth;

(B) Analyze effects of regional erosional deposition and weathering; and

(C) Make inferences and draw conclusions about effects of human activity on Earth's renewable, non-renewable, and inexhaustible resources.

Some language minority students and ELLs may come from a culture where natural disasters are explained as acts of an angry god or punishments for human misbehavior rather than scientific phenomena (Lee, 2003). For such students, switching to a Western scientific view of these phenomena requires more than simply learning facts, it requires a shift in values and beliefs as well (Lee, 2003). In addition, this standard also requires students to understand and use scientific discourse patterns (e.g., using evidence to support predictions and conclusions, making inferences, etc.) as part of demonstrating the knowledge of content.

Looking at samples of science standards provides important context for reviewing studies of standards-based science outcomes. These selected state standards illustrate the kinds of areas in which language minority students and ELLs need to demonstrate academic achievement and provides examples of some of the specific content and skills they may need to learn.

Procedure for Obtaining and Reviewing Literature

A multi-step search procedure identified literature on science outcomes or language minority students and ELLs. First, a library meta-search engine that could search multiple databases at the same time was used to look for articles in databases such as ERIC, Educational Full Text, PsycInfo, and Linguistics and Language. Search terms included combinations such as science and English language learners, science and language minority students, science and bilingual students, science and senior high, science and junior high, science achievement and student outcomes, etc. Second, reference lists of identified articles were scanned for additional resources. Finally, an Internet search using combinations of search terms similar to those used with the meta-search engine was used to verify that the relevant articles had been located.

The following criteria were used to determine which articles to review:

1. Peer reviewed journal articles—The intent of this paper is to review the status of science outcomes as described in peer reviewed journal articles.

2. Publication dates between 1994, the year of federal legislation mandating state standards, and 2008.

3. Research conducted within the United States in kindergarten through 12th grade (K-12) settings.

4. Research participants were either K-12 students or their teachers.

5. The article had to contain student science outcomes or science outcomes as reported by teachers. The scope of outcomes for this review was intentionally broad and included grades, test scores, attitudes toward science, course taking patterns, and other affective, linguistic, or psychomotor outcomes.

6. The science outcomes had to pertain directly to ELLs or language minority students. Studies were not included in this review if the ethnicity of students was the only identifying factor (e.g., Hispanic, Asian).

From this search process, 22 articles were identified that met all of the study criteria (see Appendix A for a complete listing). They were all reviewed, regardless of whether outcomes were standards-based or not, in the belief that the total group of studies would have something to say about the priorities and interests of researchers in the field. A study coding sheet was used to highlight the research design of each study along with the theoretical and conceptual frameworks underlying the research and the links between the research and standards-based instruction or outcomes.

Study Findings

The 22 studies included in this review represent diverse research perspectives, and a multitude of research designs (see Appendix A for a complete listing of studies, research methods, measures, study samples, and findings). These studies described a variety of outcomes, and represented varying degrees of connection to state standards. Some studies followed rigorous research procedures while others did not. Six studies that included ELLs or language minority students, along with other students, in the study sample did not provide disaggregated data for ELLs or language minority students, limiting the conclusions that could be made based on the data (Bouchey, 2004; Bouchey & Harter, 2005; Hadi-Tabassum, 1999; Lau & Roeser, 2002; Maheady, Michielli-Pendl, Harter & Mallette, 2006; Varelas, Pappas, Kane, Arsenault, Hankes & Cowan, 2007). As a result of this tremendous variability, it is difficult to make strong statements about the conclusions that can be drawn from this small but growing body of literature.

Links to State Standards

Appendix B shows each of the 22 studies included in this review, the outcomes associated with each study, and whether those outcomes had a connection to state standards. A number of studies were linked to either national science tests such as the National Assessment of Educational Progress (Lau & Roeser, 2002; Lee et al., 2008b; Lee et al., 2008c; Lee et al., 2005; Luykx et al., 2007) or to the national science standards published by different groups (Lynch et al., 2002). Because state and national science standards may be very different, this review begins with a look specifically at the peer-reviewed literature pertaining to state standards. A total of nine studies addressed outcomes related to state standards. These were:

- Amaral et al. (2002)
- Lee et al. (2008b)
- Lee et al. (2008c)
- Lee et al. (2005)
- Luykx et al. (2007)
- Lynch et al. (2002)
- Maheady et al. (2006)
- Rivera & Stansfield (2003-2004)
- Torres & Zeidler (2002)

Cognitive Outcomes

Standards-based measures of outcomes were usually associated with the cognitive domain. These outcomes typically were in the form of scores on state tests (Amaral et al., 2002; Lee et al., 2008b; Lee et al., 2008c; Rivera & Stansfield, 2003-2004; Torres & Zeidler, 2002) or researcher-developed tests based on curricula aligned with state standards (Lee, Deaktor, et al., 2008b, Lee et al., 2005; Maheady et al., 2006).

Two studies linking curricula with state test scores (Amaral et al., 2002; Lee et al., 2008c) indicated that test scores went up after implementation of an inquiry-focused curriculum based on state standards. However, only one of these studies (Amaral et al., 2002) reported state science test scores. Lee et al. (2008c) reported increases in math test scores because there was no state science test at the time of their study and the science curriculum that was implemented for the study incorporated math skills. While documented increases in state test scores are positive signs of improved classroom practices and student learning, neither study made clear how far ELLs and language minority students still were from passing state tests because the studies only focused on the amount of gain students made.

Of the three studies that reported scores on researcher-developed assessments aligned with state standards (Lee et al., 2008b, Lee et al., 2005; Maheady et al., 2006), Maheady et al. (2006) did not break down test results to show specifically the performance of the ELLs in the sample. The two related studies by Lee and colleagues (Lee et al., 2008b; Lee et al., 2005) showed a significant increase in standards-related test scores for ELLs and language minority students after implementation of the same inquiry-based curriculum across multiple years. In addition to explicit teaching of inquiry skills, the curriculum included a scientific language component that the researchers believed to be successful. In spite of the increase in test scores, ELLs and language minority students still scored well below other groups of students in the study. It was unclear whether their level of performance may, perhaps, have been well below what standards required for that grade level. The authors did not address the issue. The gaps in performance levels between ELLs or language minority students and other students tended to narrow a bit for students who had participated in the standards-based curriculum for two years, but gaps still remained.

Rivera and Stansfield (2003-2004) studied the effect of simplification of the language of state science test items on the performance of ELLs and non ELLs in one state. There were so few ELLs enrolled in K-12 schools in the state, and so few participating in the state science test, that no definitive conclusions were made about the effect of linguistic simplification for these students.

Affective and Other Outcomes

Studies that only measured affective outcomes such as perceptions of science-related creativity or math and science ability, and perceptions of adult support for science learning, tended not to link results to state science standards. Maheady et al. (2006) was the only exception. These researchers measured student attitudes toward a particular format of cooperative learning called Numbered Heads Together (NHT) with, and without, a behavioral incentive package. While the intervention was implemented using a standards-based science curriculum, the focus of the intervention was on the effectiveness of NHT. Every student, including ELLs, performed better on standards-based classroom assessments with some form of NHT, but the addition of the behavioral incentive had the greatest effect on scores. It also positively affected students' attitudes toward science instruction.

Two studies that focused on other forms of outcomes (Cuevas et al., 2005; Lee et al., 2006), namely the ability to conduct science inquiry tasks, were related to the same grant project that implemented a standards-aligned curriculum to directly teach inquiry processes. Over time, the students took a greater responsibility for performing inquiry tasks, after they had gained skills. Both studies (Cuevas et al., 2005; Lee et al., 2006) found that the inquiry skills of ELLs did increase after direct instruction. Lee et al. (2006) discovered that ELLs had some of the largest gains in their ability to perform inquiry tasks. However, students became more effective only on certain parts of the inquiry tasks, such as their abilities to support theories with data and controlling specific variables (e.g., amount of water used, the length of time and the location of an experiment). They continued to struggle with controlling other variables (e.g., the size and shape of containers, heat, location and length of time) (Lee et al., 2006). Lee et al. (2006) attributed some of the difficulty with inquiry-based learning to students' cultural backgrounds and values.

> Although conducting scientific inquiry is a challenge for most students, it presents additional challenges for students from societies that may not encourage them to engage in some aspects of inquiry practices, such as asking empirical questions about natural phenomena, designing and implementing systematic investigations, and finding answers on their own…In many societies, cultural norms prioritize respect for teachers and other adults as authoritative sources of knowledge. In other words, validity of knowledge is often based on the validity of its source, rather than the validity of knowledge claims. Children who are taught to respect the wisdom and authority of their elders may not be encouraged to question received knowledge in ways that are compatible with Western scientific practices (Lee et al., 2006, p. 611)

The researchers believed that developmental readiness to perform inquiry tasks was also a factor that caused 4th graders to perform better than 3rd graders in the study. Lee et al. (2006) provided an important implication for teaching inquiry science to ELLs. Furthermore, as student understanding of and skill with inquiry procedures grew, they often struggled to express that understanding with their developing English skills.

The picture that emerged from these nine studies was that a handful of researchers examined ways to increase the grade-level standards-based science achievement of ELLs and language minority students. However, these students may still have performed below the level that was required to pass standards-based assessments and to graduate within a standards-based system. The studies did not provide much information on which to judge how far ELLs and language minority students have to go, and in what areas or aspects of science, to be successful. With limited and incomplete information it is difficult to use the research literature as a guide to plan instructional improvements.

Given that there is so little research with clear data on state standards-based science outcomes specifically for ELLs and language minority students, perhaps the most important issues that arise from this review relate to the state of the field and recommendations for the future. These issues and recommendations are based on all 22 studies, not only those related to state standards.

Relationship of Studies to Model

Referring back to the Revised Theory of Educational Effectiveness under Standards-Based Reform in Figure 2, how did the studies reviewed match up to the elements of the figure? As Table 2 shows, the majority of reviewed studies addressed either student factors (n=13) such as attitudes, cultural background and language proficiency levels, or classroom factors (n=11), particularly curriculum implementation and associated teacher training. Four of the studies (Duran et al., 1998; Hampton et al., 2001; Lee et al., 2005; Lynch et al., 2002) examined factors at both the student and the classroom levels. Amaral et al. (2002) reported the effects of a science curriculum on students in an entire school district, and, for the purposes of this review, is included with the school level. Only one study (Rivera & Stansfield, 2003-2004) addressed national/local policy context factors in the form of a state test accommodation not commonly allowed in state testing policies.

Table 2. Factors Addressed by Studies Reviewed

	Factors			
Study	Student	Classroom	School/District	National/Local
Amaral et al. (2002)			X	
Beghetto (2007)	X			
Bouchey (2004)	X			
Bouchey & Harter (2005)	X			
Cuevas et al. (2005)		X		
Duran et al. (1998)	X	X		
Hadi-Tabassum (1999)	X			
Hampton et al. (2001)	X	X		
Lau & Roeser (2002)	X			
Lee et al. (2008a)		X		
Lee et al. (2008b)		X		
Lee et al. (2008c)		X		
Lee et al. (2006)		X		
Lee et al. (2005)	X	X		
Luykx et al. (2007)	X			
Lynch et al. (2002)	X	X		
Maheady et al. (2006)		X		
Medina & Mishra (1994)	X			
Rivera & Stansfield (2003-2004)				X
Shaw (1997)	X			
Torres & Zeidler (2002)	X			
Varelas et al. (2007)		X		

Because the intent of standards-based reform is to assist all students in achieving grade-level content by improving the quality of classroom conditions that support learning, an emphasis on classroom factors in the research literature seems important and appropriate. By focusing on the classroom, researchers have placed the responsibility for increasing student outcomes primarily on educators and schools, not on the students. Further studies on classroom factors are important. In addition, the 22 studies reviewed here also emphasize the role of student factors, such as attitude and motivation, in learning science. These studies illuminate the ways in which students are unique. They have different attitudes toward science, differing levels of support for science learning, and differing motivations to succeed in the science classroom. Differences are important to study so long as the emphasis remains on stimulating teachers to teach standards-based content differently in order to better meet student needs.

Research Perspectives Relate to Outcomes Studied

In a previous review of studies on science instruction for diverse students, Lee and Luykx (2007) reported that few studies provided concrete, quantitative outcomes (p. 177). One possible reason for this finding may be that many studies in the field come out of disciplinary traditions and theoretical or conceptual frameworks with non-cognitive types of emphases. Table 3 highlights that half of the total group of studies reviewed (n=11) fall into this category. The few studies in Table 3 that included cognitive outcomes, such as Bouchey (2004), also included other types of outcomes as well. Footnotes to the table describe terms used to refer to theoretical or conceptual frameworks, as used by article researchers.

Two examples from the table provide illustrations of the way in which aligning research traditions with theoretical or conceptual frameworks, and thus with research questions, can lead to emphases on varying outcomes. First, Duran, Dugan, and Weffer (1998) chose to study science and language minority students from the perspective of social constructivism, which is associated with the tradition of social psychology. Duran et al. (1998) define social constructivism in the following way:

> Our research was conducted within the social constructivist framework of Vygotsky (1978, 1987), as extended by Wertsch (1991), which offers a theoretical perspective about the development of higher order mental functions and the interdependence of language and thought in learning….Wertsch (1991) uses the idea of 'a social language as a meditational means' to provide a dialogic account for the social nature of child learning…Learning occurs as children internalize the dialogues of their social interaction by gradually using these for their intention. (pp. 313-314)

Table 3. Research Perspectives of Studies Reviewed

Study	Disciplinary Roots	Theoretical or Conceptual Framework	Focus/Research Questions or Hypotheses	Student Outcomes
Beghetto (2007)	Psychology	Achievement Goal Theory [1]	Perceived Science Competence 1. After controlling for differences in age, gender, and ethnicity, are students' goal orientations related to their perceived science competence beliefs? 2. Is there an association between students' self-perceptions of their ability to generate creative ideas and their perceived competence in science? 3. Does a link exist between students' perceived science competence and their perceptions of teacher support and press?	Affective
Bouchey, (2004)	Social Psychology	Symbolic interactionism [2]	Interplay between others' influence and students' psychological factors in predicting achievement 1. Parents', teachers' and classmates' beliefs regarding the importance of math and science and the student's ability will predict students' reflected appraisals of others' beliefs 2. Students' reflected appraisals would predict their own self-perceptions, scholastic behavior, and math and science performance 3. Support from others predicts students' reflected appraisals	Cognitive Affective
Bouchey & Harter (2005)	Social Psychology	Symbolic interactionism Expectancy value theory[3]	Test a meditational model of relations among (a) students' perceptions of others' beliefs and behavior regarding schoolwork, (b) students' own academic self-perceptions and behavior, and (c) academic performance.	Cognitive Affective
Cuevas, Lee, Hart, & Deaktor (2005)	n/a	*Instructional Congruence Teacher-explicit to student-exploratory continuum	Impact of an inquiry-based instructional intervention 1. What is the impact of the instructional intervention on students' ability to conduct science inquiry overall and to use inquiry skills of questioning, planning, implementing, concluding and reporting? 2. What is the impact of the instructional intervention on narrowing gaps in the ability to conduct inquiry among demographic subgroups of students with respect to grade, achievement, gender, ethnicity, SES, home language, and English proficiency?	Ability to conduct inquiry-based science
Duran, Dugan, & Weffer (1998)	Social psychology	Social constructivism (as extended by Wertsch)[4]; social semiotics	Describe how students construct biology concept meanings based on extant language skills and engage students in constructing meaning through meditational means.	Learning to talk science

*Researcher inferred

[1]**Achievement Goal Theory**: Perceived competence relates to a students' motivation to learn.
[2]**Symbolic interactionism**: Self-concept is constructed based on internalizing others' beliefs about us.
[3]**Expectancy Value Theory**: Children's perceptions of their socializers' beliefs, expectations and attitudes predict children's' self-concept.
[4]**Social constructivism**: Language and thought are interdependent in learning; learning occurs as children internalize social interactions and use the interactions as the basis for expressing their own ideas.

Study	Disciplinary Roots	Theoretical or Conceptual Framework	Focus/Research Questions or Hypotheses	Student Outcomes
Lau & Roeser (2002)	Psychology	Pathways to achievement outcomes (Snow)[5] Social-cognitive theories of motivation	How cognitive and motivational factors associated with performance and commitment pathways to achievement contribute to the prediction of achievement outcomes in science. 1. Both sets of factors would add incremental predictive validity to outcomes 2. Cognitive abilities would be most closely associated with science performance measured by grades and test scores 3. Motivational variables would be most closely associated with students' future science activities Cognitive and motivational factors will have both direct and indirect effects on outcomes	Cognitive Affective
Lee et al. (2008a)	Cultural Anthropology	*Instructional Congruence Teacher-explicit to student-exploratory continuum	Teachers' perceptions of the teaching and learning impact of an inquiry-based curriculum plus teacher training intervention	Teacher perceptions of student learning
Lee et al. (2006)	Cultural anthropology	Design research (cf. Brown, 1992, 1994; Lehrer et al., 2000; Metz, 1995, 1997; Rosebery, Warren & Conant, 1992) *Instructional Congruence Teacher-explicit to student-exploratory continuum	Preliminary investigation into the impact of an instructional intervention on science inquiry abilities among linguistically and culturally diverse elementary students.	Ability to conduct inquiry
Lynch, Kuipers, Pyke, & Szesze (2005)	Psychology Social Psychology	Conceptual Change Theory[6] Culturally-defined Activity Systems (Sociocultural theory)[7]	Results of a planning grant studying the effects of a highly rated science curriculum unit on a diverse population.	Affective Cognitive
Maheady, Michielli-Pendl, Harter, & Mallette (2006)	Psychology	*Applied Behavior Analysis	Examined the effects of Numbered Heads Together with and without a behavioral incentive package on 6th graders daily quiz scores and pretest-posttest performance in chemistry	Cognitive Affective
Varelas, Pappas, Kane, Arsenault, Hankes, & Cowan (2007)	Psychology	Sociocultural perspective of teaching and learning[8]	Focus on the concept of matter and explore how young children in urban schools bridge their spontaneous concepts and everyday experiences with scientific concepts introduced to them by children's literature information books and their teachers	Meaning making in collaborative interactions

[5]**Pathway to Achievement Outcomes**: Cognitive abilities (performance) and affective processes (commitment) both contribute to achievement outcomes.
[6]**Conceptual Change Theory**: Each learner has the opportunity to interact with curriculum materials to develop more advanced conceptual knowledge.
[7]**Culturally-Defined Activity Systems (Sociocultural theory)**: The classroom is a dynamic environment where students and teachers participate in many overlapping communities of practice.
[8]**Sociocultural perspective of teaching and learning**: Emphasizes the meaning making that takes place in classroom communities; material artifacts as tools.

Because this study focused on social speech as a mediator of individual student learning, as expressed through language, it is logical that the researchers collected interview and observational data on outcomes related to student production of meaning in classroom interactions.

A second example of the way disciplinary traditions and theoretical or conceptual frameworks align with science outcomes can be found in Beghetto's (2007) study. The researchers examined science through the perspective of Achievement Goal Theory, which comes out of the field of psychology. As Beghetto (2007) explained:

> This line of research suggests that although cognitive ability is important to success in science—ability alone is not enough. Indeed, students who otherwise have the ability to be successful in science, yet believe they are not capable of success, likely sell themselves and the field short. Students' perceived science competence is related to students' motivation to learn, future aspirations, and ultimately their achievement in science. (p. 800)

As a result of this focus, Beghetto collected data on students' perceived efficacy, perceived science competence, achievement goals, and perceived teacher support. Cognitive outcomes were not included in Beghetto's study. Again, in the context of this particular disciplinary tradition and theoretical framework, outcomes relating to student attitudes and motivation make sense.

The remaining studies reviewed focused solely on cognitive outcomes such as test scores and grades. Six of these studies, several from the same research project, incorporated test scores to measure the effectiveness of implementing new curricula (Amaral et al., 2002; Cuevas et al., 2005; Lee, Deaktor et al., 2008; Lee, Maerten-Rivera et al., 2008; Lee et al., 2006; Lee & Luykx, 2006). Studies like those of Duran et al. (1998) and Beghetto (2007) add richness to our understanding of the context of science outcomes. However, the model of educational effectiveness under standards-based reform (Figure 2) does emphasize standards-based cognitive outcomes as an end goal of instruction. Other types of outcomes may be important contributors to cognitive outcomes (e.g., levels of academic English proficiency, attitudes toward science), but these are not the end goal of instruction.

Conclusion and Recommendations

The results of this research review indicate that there is a small but growing research base that addresses science outcomes for language minority students and ELLs. The purpose of this review was to focus more specifically on what is known about outcomes related to state science standards for ELLs and language minority students. Of the 22 studies that addressed science outcomes for these students, many were related to either national science assessments or to national science standards. A smaller proportion was related to state standards. The small number of studies available combined with the diversity of topics studied and the multitude of viewpoints makes it difficult to distill findings from this body of literature. For example, theoretical and conceptual frameworks researchers chose for the larger group of studies influenced the way outcomes were conceptualized and measured. Some studies conceptualized outcomes in purely cognitive terms such as scores on tests, while others addressed affective or other types of outcomes.

There were only nine studies linking science outcomes to state standards in some way. Four of these nine studies were based on an implementation of one inquiry-based science curriculum used in the same grant project (Lee et al., 2008b; Lee et al., 2008c; Lee et al., 2005; Luykx et al., 2007). It is not possible to state definitive research findings with such a small number of studies, but two ideas contained in those eight studies stand out for further consideration.

First, scientific inquiry skills, which are often part of state science standards, are important to explicitly teach to language minority students and ELLs, but may be challenging for students to fully learn. Because of differing cultural backgrounds and different levels of English proficiency, educators should not assume that linguistically and culturally diverse students will implicitly have the knowledge or skills to perform inquiry tasks without being taught how to do so. A few studies in this review found that ELLs and language minority students did, in fact, perform significantly higher on inquiry-based tasks and test items after receiving instruction on the different inquiry elements plus scientific language. However, ELLs in particular still struggled with certain inquiry concepts and skills after the interventions.

Second, curriculum-related interventions had differential effects for language minority students and ELLs. While ELLs displayed a great deal of growth in science learning after the implementation of these interventions, they still scored below their language minority peers on content-related assessments and inquiry tasks. Researchers observed that a sizeable gap in performance levels between ELLs and other students was still evident. One major challenge in summarizing science outcomes for linguistically and culturally diverse students is that the available literature did not state how far below grade-level standards the performance of ELLs and language minority students was.

Both of these key points in the literature highlight the challenges inherent in teaching academic English language skills specific to science simultaneously with both science content knowledge and science processes in the mainstream classroom. ELLs and language minority students who have recently exited English as a Second Language services have to learn more knowledge and skills in order to achieve at grade level in the science classroom than do their native English speaking peers. The work of Lee and colleagues (Cuevas et al., 2005; Lee et al., 2008a; Lee et al., 2008b, Lee et al., 2008c; Lee et al., 2006; Lee et al., 2005) highlighted the need to incorporate scientific language instruction into the teaching of scientific inquiry skills while also focusing on content acquisition. It is vitally important to conduct more studies of this kind relating content, language, and skill integration in the science classroom.

Considering the extremely small number of studies relating specifically to standards-based science outcomes for ELLs and language minority students, the strongest statements that can be made about the 22 studies reviewed relate to current emphases in the field of science education for linguistically and culturally diverse students and possible research directions.

1. **The types of standards-related cognitive outcomes measures used and the way results were reported did not allow for clear identification of how well or how poorly language minority students performed in science.** Standards-related outcomes were typically associated with cognitive measures such as tests or grades. Test scores were not necessarily from (state) standards-based tests and may not have been associated with science tests at all. A few studies included measures of math and literacy as possible outcomes. Researcher-developed assessments that related to standards-aligned curriculum were common, but the studies frequently did not report linkages to state standards in detail. In several studies, gains in science learning as measured by tests were reported, but whether students met state standards was not reported. In still other studies data were not necessarily disaggregated to show the performance of ELLs and language minority students in the study sample. The field could benefit greatly from more studies incorporating cognitive measures of standards-based science learning, with clearly disaggregated data for linguistically and culturally diverse students, and specific indications of how far these students have to go to meet grade-level standards. Specific knowledge in this area could help educators and policymakers plan effective programming and choose the best standards-based and instructional practices for students.

2. **The influence of standards-based teaching and learning on study participants was not explicitly acknowledged as a contextual factor in research studies**. The majority of the 22 studies reviewed involved students who received (state) standards-based instruction in their daily lives at school or teachers who taught standards-aligned science curricula. Yet this influence on the research was often not explicitly acknowledged or considered, particularly in studies of student factors associated with science outcomes.

Greater contextualization of studies and more direct links to standards-based instruction are needed. Findings from studies that report affective or other types of outcomes, as well as those that report cognitive outcomes, could be strengthened by this more nuanced approach.

3. **Most of the studies reviewed focused on either student-level factors associated with science outcomes (see Figure 2) or classroom-level factors. Few, if any, studies examined school-level factors or state and local policy factors and their relationship to science outcomes for ELLs and language minority students.** In addition, only a handful of studies linked together factors from across levels (see Figure 2) to explain how multiple levels within the education system interact to play a role in student outcomes. Teachers make instructional decisions within the context of decisions that have already been made at other levels (Creemers & Kyriakides, 2008). Teachers also may make different instructional decisions depending on the unique characteristics of the students in their classrooms at any given time. Studies linking the various levels represented in Figure 2 would help educators, policymakers, and researchers better understand the complexities inherent in teaching linguistically and culturally diverse students.

At this point in time, the research literature simply does not provide enough information about state standards-based science outcomes for ELLs and language minority students to guide educational improvement efforts. This lack of research may be, in part, because state accountability for standards-based science outcomes, in the form of mandatory statewide achievement tests, is relatively new. Policymakers, educators, and researchers are just beginning to turn their attention to the instructional implications of standards-based science assessment scores for all students and largely have not yet begun to focus on outcomes for ELLs and language minority students. Nevertheless, the available research suggests some promising directions for future research efforts.

References

Amaral, O., Garrison, L., & Klentschy, M. (2002). Helping English language learners increase achievement through inquiry-based science instruction. *Bilingual Research Journal, 26*(2), 213-239.

Beghetto, R. (2007). Factors associated with middle and secondary students' perceived science competence. *Journal of Research in Science Teaching, 44*(6), 800-814.

Bouchey, H. (2004). Parents, teachers, and peers: Discrepant or complementary achievement socializers. *New Directions for Child and Adolescent Development, 106*, 35-53.

Bouchey, H., & Harter, S. (2005). Reflected appraisals, academic self-perceptions, and math/science performance during early adolescence. *Journal of Educational Psychology, 97*(4), 673-686.

Creemers, B. P. M., & Kyriakides, L. (2008). *The dynamics of educational effectiveness: A contribution to policy, practice and theory in contemporary schools.* New York, NY: Routledge.

Cuevas, P., Lee, O., Hart, J., & Deaktor, R. (2005). Improving science inquiry with elementary students of diverse backgrounds. *Journal of Research in Science Teaching, 42*(3), 337-357.

Duran, B., Dugan, T., & Weffer, R. (1998). Language minority students in high school: The role of language in learning biology concepts. *Science Education, 82*(3), 311.

Hadi-Tabassum, S. (1999). Assessing students' attitudes and achievements in a multilingual, multicultural science classroom. *Multicultural Education, 7*(2), 15-20.

Hampton, E., & Rodriguez, R. (2001). Inquiry science in bilingual classrooms. *Bilingual Research Journal, 25*(4), 417-434.

Lau, S., & Roeser, R. (2002). Cognitive abilities and motivational processes in high school students' situational engagement and achievement in science. *Educational Assessment, 8*(2), 139-162.

Lauer, P., Snow, D., Martin-Glenn, M., Van Buhler, R., Stoutemyer, K., & Snow-Renner, R. (2005). *The influence of standards on K-12 teaching and student learning: A research synthesis.* Mid-Continent Research for Education and Learning (MCREL).

Lee, O. (2003). Equity for linguistically and culturally diverse students in Science education: A research agenda. *Teachers College Record*, *105*(3), 465-489.

Lee, O. (2004). Teacher change in beliefs and practices in science and literacy instruction with ELLs. *Journal of Research in Science Teaching, 41*(1), 65-93.

Lee, O. (2005). Science education with English language learners: synthesis and research agenda. *Review of Educational Research, 75*(4), 491-530.

Lee, O., Adamson, K., Maerten-Rivera, J., Lewis, S., Thornton, C., & LeRoy, K. (2008). Teachers' perspective on a professional development intervention to improve science instruction among ELLs. *Journal of Science Teacher Education, 19*, 41-67.

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., Enders, C., & Lambert, J. (2008). Impact of a multiyear professional development intervention on science achievement of culturally and linguistically diverse elementary students. *Journal of Research in Science Teaching, 45*(6), 726-747.

Lee, O., Deaktor, R., Hart, J., 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., & Luykx, A. (2006). *Science education and student diversity: Synthesis and research agenda.* New York: Cambridge University Press.

Lee, O., & Luykx, A. (2007). Science education and student diversity: Race, ethnicity, language, culture and socioeconomic status. In S. Abell & N. Lederman (Eds.), *Handbook of research in science education* (pp. 171-197). Mahwah, NJ: Lawrence Earlbaum.

Lee, O., Maerten-Rivera, J., Penfield, R., LeRoy, K., & Secada, W. (2008). Science achievement of English language learners in urban elementary schools: Results of a first-year professional development intervention. *Journal of Research in Science Teaching, 45*(1), 31-52.

Luykx, A., Lee, O., Mahotiere, M., Lester, B., Hart, J., & Deaktor, R. (2007). Cultural and home language influences on children's responses to science assessments. *Teachers College Record, 109*(4), 897-926.

Lynch, S., Kuipers, J., Pyke, C., & Szesze, M. (2005). Examining the effects of a highly rated science curriculum unit on diverse students: Results from a planning grant. *Journal of Research in Science Teaching, 42*(8), 912-946.

Maheady, L., Michielli-Pendl, J., Harper, G., & Mallette, B. (2006). The effects of numbered heads together with and without an incentive package on the science test performance of a diverse group of sixth graders. *Journal of Behavioral Education, 15*(1), 25-39.

Marzano, R., & Kendall, J. (1998). *Awash in a sea of standards.* Denver, CO: Mid-Continent Research for Education and Learning (MCREL).

McClure, P. (2005). Where standards come from. *Theory into Practice, 44*(1), 4-10.

Medina, M., & Mishra, S. (1994). Relationships among Spanish reading achievement and selected content areas for fluent- and limited-Spanish-proficient Mexican Americans. *Bilingual Review, 19*(2), 134-142.

National Research Council (1999). *Testing, teaching, and learning: A guide for states and school districts.* Committee on Title I Testing and Assessment, Richard F. Elmore and Robert Rothman, editors. Board on Testing and Assessment, Commission on Behavioral and Social Sciences and Education. Washington, D.C.: National Academy Press.

Paige, R. (2003, March). *Remarks by Secretary Paige to the Commonwealth Club of California.* Speech presented at the Commonwealth Club of California.

Rivera, C., & Stansfield, C. (2004). The effect of linguistic simplification of science test items on score comparability. *Educational Assessment*, *9*(3&4), 79-105.

Saderholm, J., & Tretter, T. (2008). Identification of the most critical content knowledge base for middle school science teachers. *Journal of Science Teacher Education, 19*, 269-283.

Shaw, J. (1997). Threats to the validity of science performance assessments for English language learners. *Journal of Research in Science Teaching, 34*(7), 721-743.

Stoddart, T., Pinal, A., Latzke, M., & Canaday, D. (2002). Integrating inquiry science and language development for English language learners. *Journal of Research in Science Teaching, 39*(8), 664-687.

Torres, H., & Zeidler, D. (2002). The effects of English language proficiency and scientific reasoning skills on the acquisition of science content knowledge by Hispanic English language learners and native English language speaking students. *Electronic Journal of Science Education, 6*(3). Available: http://ejse.southwestern.edu/original%20site/manuscripts/v6n3/articles/art04_torres/torres.pdf

Varelas, M., Pappas, C., Kane, J., Arsenault, A., Hankes, J., & Cowan, B. (2007). Urban primary-grade children think and talk science: Curricular and instructional practices that nurture participation and argumentation. *Science Education,* 65-95.

Appendix A

Studies Reviewed

Study	Focus and Connection to Standards	Methods, Measures and Sample	Findings
Amaral, Garrison, & Klentschy (2002)	Project evaluation 1. What are participants' rates of achievement in science content? 2. How does the growth of ELLs compare to growth by different categories of language minority students? 3. How does participation in the Local Systemic Reform Initiative affect student achievement? *Optional science portion of test used for state standards*	Method: Quantitative non-experimental Measures: SAT9 Science test Sample: 615 grade 4; 635 grade 6 students in bilingual, structured immersion and English only classrooms from one district	SAT9 mean science scores went up consistently with longer participation in science program. District writing scores went up (more consistently for grade 6) with longer program participation; possibly due to science journaling. SAT9 math and reading scores increased with time in the science program. Math increases were less consistent and greater for grade 6. Science, writing, reading, and mathematics achievement of English language learners went up with longer science program participation.
Beghetto (2007)	Perceived Science Competence 1. After controlling for differences in age, gender, and ethnicity, are students' goal orientations related to their perceived science competence beliefs? 2. Is there an association between students' self-perceptions of their ability to generate creative ideas and their perceived competence in science? 3. Does a link exist between students' perceived science competence and their perceptions of teacher support and press? *No explicit link to standards*	Method: Quantitative non-experimental Measures: Survey of school experiences (student characteristics, perceived science competence, achievement goals, creative self-efficacy, perceived teacher support) Sample: 1,289 students in grades 6-12 (63% language minority)	Perceived science competence was related to four factors (a) age, gender, ethnicity; (b) mastery and performance-approach goals; (c) self-perceptions of creativity; and (d) perceptions of teacher support and degree of academic challenge. The combination of perceived teacher support and degree of academic challenge is related to positive self-perceptions of competence. Self-perceptions of creativity were found to be the strongest correlates with perceived science competence. Perceived teacher support alone was not enough for developing positive perceptions of academic competence. No data specific to language minority students

Study	Focus and Connection to Standards	Methods, Measures and Sample	Findings
Bouchey (2004)	Interplay between others' influence and students' psychological factors in predicting achievement 1. Parents', teachers' and classmates' beliefs regarding the importance of math and science and the student's ability will predict students' reflected appraisals of others' beliefs 2. Students' reflected appraisals would predict their own self-perceptions, scholastic behavior, and math and science performance 3. Support from others predicts students' reflected appraisals *No explicit link to standards*	Method: Quantitative non-experimental Measures: Socializers' (mothers, teachers, classmates) beliefs and behavior, students' reflected appraisals of importance of doing well in math and science and academic ability, students' perceptions of others' support for math and science, students' perceptions of importance of math and science and their own ability, students' time and energy given to math and science, current performance, prior achievement Sample: 378 middle school students (11% ELLs), 21 teachers, 150 mothers	Students' reflected appraisals of what their classmates, teachers, and parents think isn't an important mediator between the actual perceptions of these socializers and students' self-perceptions. Adults and classmates may not accurately report student's competence in science. Students may have also had inaccurate perceptions of others' perceptions. Over time parents and children, as well as students and classmates, may hold outdated perceptions to which they pay less attention. *No data specific to English language learners*
Bouchey & Harter (2005)	Test a meditational model of relations among (a) students' perceptions of others' beliefs and behavior regarding schoolwork, (b) students' own academic self-perceptions and behavior, and (c) academic performance. *No explicit links to standards*	Method: Quantitative non-experimental Measures: Perceived importance of math/ science to others, support for schoolwork, importance of schoolwork, beliefs about the student's competence, scholastic behavior, perceived competence, current performance and prior achievement. Sample: 378 middle school students (41 ELLs)	Students' self-perceived importance, competence, scholastic behavior and performance in science were predicted by reflected appraisals of adults' beliefs about the importance of science and adults' beliefs about student competence in math and science along with students' perceived support for science. Controlling for maternal education, Latino students reported lower mean levels of perceived competence than did European American students.

Study	Focus and Connection to Standards	Methods, Measures and Sample	Findings
Cuevas, Lee, Hart & Deaktor (2005)	Impact of an inquiry-based instructional intervention 1. What is the impact of the instructional intervention on students' ability to conduct science inquiry overall and to use inquiry skills of questioning, planning, implementing, concluding and reporting? 2. What is the impact of the instructional intervention on narrowing gaps in the ability to conduct inquiry among demographic subgroups of students with respect to grade, achievement, gender, ethnicity, SES, home language, and English proficiency? *No explicit links to standards*	Method: Mixed method (quasi-experimental) Measures: Pre-post inquiry-based instructional intervention. Students completed an elicitation protocol asking them to design a scientific investigation to solve a problem about the effect of surface area on the rate of evaporation Sample: 25 students in grades 3 & 4 from six linguistically and culturally diverse elementary schools (13 language minority students exited from ESL; others unclear)	The intervention enhanced the inquiry ability of all students regardless of subgroup, particularly former ELLs, low achieving students, and low socioeconomic status students.
Duran, Dugan, & Weffer (1998)	Describe how students construct biology concept meanings based on extant language skills and engage students in constructing meaning through meditational means. *No explicit links to standards*	Method: Qualitative Measures: Participant observation, questionnaires, interviews, group discussions, standardized tests at the end of the year Sample: 14 10th graders (all language minority students) enrolled in a 32-week academic preparatory science course at a college	Students initially watched the teacher practice science and waited for the teacher's interpretation of events. They did not construct their own meanings of science as long as the teacher provided meanings. The teacher showed students how to use resources in the text to construct meaning and emphasized relational patterns for connecting ideas. Students then used these patterns to organize other information in the course. Diagrams of concepts were also used as tools to promote collaborative meaning-making between teachers and students and between groups of students.
Hadi-Tabassum (1999)	A qualitative and quantitative look at the instructional curriculum and teaching of one two-way immersion 8th grade science classroom that describes implications for reform efforts. *No explicit links to standards*	Method: Mixed method (quasi-experimental) Measures: Attitude survey Sample: 25 at-risk 8th graders (12 ELL)	There was a 10 point increase in the range of student attitudinal scores about science from before the intervention to after the intervention.

Study	Focus and Connection to Standards	Methods, Measures and Sample	Findings
Hampton & Rodriguez (2001)	Value of implementing an inquiry science curriculum with second language learners. *No explicit links to standards*	Method: Qualitative Measures: Student attitude surveys, three focus groups with selected (n=20) interns about their experience after implementing two lessons in larger intervention, writing samples from all interns (n=100) providing feedback about the effectiveness of implementing the intervention, Likert scale and open-ended comments from classroom teachers. Sample: 80 third graders and 107 fifth graders (largely ELLs); 100 university interns who implemented a science intervention in elementary classrooms and the classroom teachers	Data indicates all participants had a strong positive feeling of the value of implementing an inquiry-based curriculum. The curriculum increased the first and second language skills of participating students along with their science content knowledge and skills.
Lau & Roeser (2002)	How cognitive and motivational factors associated with performance and commitment pathways to achievement contribute to the prediction of achievement outcomes in science. 1. Both sets of factors would add incremental predictive validity to outcomes 2. Cognitive abilities would be most closely associated with science performance measured by grades and test scores 3. Motivational variables would be most closely associated with students' future science activities 4. Cognitive and motivational factors will have both direct and indirect effects on outcomes *Test based on large-scale test items (e.g., NAEP, SAT, NELS: 88, etc.)*	Method: Quantitative non-experimental Measures: Survey of motivation and background characteristics, assessments of verbal, quantitative and spatial ability, science achievement tests, grades Sample: 491 10th and 11th graders enrolled in science classes in one high school in northern California (131 language minority students; 2 reported limited English skills)	(a) Students cognitive abilities were the strongest predictors of standardized test scores in science; (b) motivation enhanced the predictive ability of science test scores and grades beyond that of just ability; (c) motivation was the strongest predictor of engagement in science and the potential to choose science-related college majors and jobs.

Study	Focus and Connection to Standards	Methods, Measures and Sample	Findings
Lee, et al. (2008a)	Teachers' perceptions of the teaching and learning impact of an inquiry-based curriculum plus teacher training intervention. *Test based on science curriculum units aligned with state standards; Test based on NAEP/ TIMSS items*	Method: Descriptive Measures: 3-part questionnaire: (1) Effectiveness of each component of inquiry-based intervention for science teaching; (2) Effectiveness of instructional materials and teacher workshops; (3) Effectiveness of intervention overall. Sample: 44 elementary teachers of Haitian Creole and Spanish speaking ELLs	Teachers believed the intervention (curriculum materials and teacher training workshops) was effective in promoting student learning of science, language and math. Receiving science supplies was one of the most important strengths.
Lee, et al. (2008b)	Preliminary investigation into the impact of an instructional intervention on science inquiry abilities among linguistically and culturally diverse elementary students. *Test based on science curriculum units aligned with standards; Test based on NAEP/ TIMSS items*	Method: Quasi-experimental Measures: Researcher-developed science tests based on instructional units, test with NAEP/TIMSS items that corresponded to instruction Sample: Year 1=1273 3rd-5th graders; Year 2=1,620 4th-5th graders; Year 3=431 5th graders at six elementary schools whose teachers participated in teacher training on new inquiry-based curriculum (ELLs were 24% of sample)	Intervention was generally effective in promoting equity and achievement with diverse students. Significant pre-post test increases with large effects in each of 3 years at all grade levels on every measure. 3rd graders showed the most significant gains and largest effect sizes. Achievement gaps between subgroups narrowed in some cases and stayed consistent in others. Item-level comparisons with NAEP/TIMSS tests showed overall positive performance by study participants at end of each school year.

Study	Focus and Connection to Standards	Methods, Measures and Sample	Findings
Lee, et al. (2008c)	Impact of a 3-year implementation of a professional development intervention (and associated curriculum units) on science achievement of culturally and linguistically diverse elementary students. 1. How did participating students perform in science overall? 2. How did performance differ among demographic subgroups in terms of ethnicity, home language, English language proficiency, socioeconomic status, special education status, and gender? 3. How did students' performance on NAEP/TIMSS items compare with the performance of national and International samples of students? *State math test; Test based on NAEP/ TIMSS items*	Method: Quasi- Experimental (with control) Measures: Researcher-developed science test based on content of inquiry-curriculum; Statewide math test scores Sample: 42 3rd grade teachers, 818 3rd graders (15 ELLs, 38 recently exited language minority students; unclear how many others are language minority)	Treatment students showed statistically significant increase in science achievement. No statistically significant difference in gains between ELLS and former ELLs or non-ELLs. No significant difference in achievement gains for students retained in grade due to low state test scores and students not retained. Treatment group had a higher state math test scores, especially on the measurement strand associated with the science intervention, than comparison group.
Lee, et al. (2006)	Examine the impact of the first year of full implementation of the intervention on students' science and literacy achievement during the school year. *Used inquiry definition found in NRC standards; Inquiry-based curriculum aligned with state standards*	Method: Mixed method (Quasi-experimental) Measures: Pre-post intervention semi structured inquiry task on evaporation. Sample: 28 3rd and 4th graders from seven classrooms in six elementary schools where teachers had participated in inquiry-based science curriculum training (13 Language Minority students who were former ELLs). Teachers were purposively selected for effectiveness with curriculum. 1 high achieving boy and girl and one low-achieving boy and girl were chosen from each teacher's classroom.	Students demonstrated greater abilities with some parts of the inquiry task after the intervention but still had difficulties with other parts of the task. 4th graders had a better understanding and showed higher gains relating to particular aspects of the inquiry task (control of variables, use of measurement data, and tools to support theories) compared to third graders. Students from all demographic backgrounds showed large gains, but the largest gains were shown by "nonmainstream" and less privileged students, including former ELLs.

Study	Focus and Connection to Standards	Methods, Measures and Sample	Findings
Lee, et al. (2005)	Examine third grade students' science achievement after the first year of implementation of a professional development intervention (and associated curriculum units) 1. Did treatment students display pre/post gains in year 1? 2. Did science achievement gaps between levels diminish pre/post? 3. Did treatment and comparison groups perform similarly on the state math test? *Test based on science curriculum units aligned with state standards; Test based on NAEP/ TIMSS items*	Method: Quasi-experimental Measures: Researcher-developed pre-post intervention science tests; one NAEP/TIMSS test given at the start and end of school year, one researcher-developed writing prompt Sample:1,523 3rd and 4th graders from six schools whose teacher participated in inquiry-based science curriculum training (25% ELL)	Statistically significant increases on all measures of science and literacy at both grades. Generally stronger effects for science. Achievement gaps between subgroups of students narrowed on some measures at 3rd grade and narrowed for 4th graders. Growth rates for ELLs were similar to growth rates or non-ELLs but ELLs scored lower overall than other subgroups.
Luykx, Lee, Mahotiere, Lester, Hart, & Deaktor (2007)	Examine cultural and linguistic interference in the open-ended responses of 3rd and 4th grade students on paper-and-pencil science tests. *Test based on NAEP performance items*	Method: Qualitative (discourse analysis of errors) Measures: Project-developed science assessment administered pre-post intervention; intervention not part of study. Sample: 1500 3rd and 4th graders (includes many ELLs; number unspecified)	Analysis showed that misinterpretation of science test items was caused by phonological/orthographic and semantic interference from students' native languages. Differences in cultural beliefs and practices and "languacultural" features of the scientific language used in the items also caused misinterpretations.

Study	Focus and Connection to Standards	Methods, Measures and Sample	Findings
Lynch, Kuipers, Pyke, & Szesze (2005)	Results of a planning grant studying the effects of a highly rated science curriculum unit on a diverse population. *Science test based on curriculum aligned with AAAS standards (Conservation of Matter Assessment)*	Method: Mixed method (Quasi-experimental with control plus ethnographic study of classroom) Measures: Conservation of Matter Assessment related to Chemistry that Applies curriculum, motivation and engagement questionnaire Sample: 1500 8th graders in 5 diverse middle schools with oversampling from most diverse schools (139 ELLs—61 comparison, 78 treatment; 331 language minority students—140 comparison, 191 treatment)	Students in both conditions generally reported being engaged, being interested in learning and being focused on their own performance. Subgroups of students in treatment group (using Chemistry that Applies curriculum) outperformed comparison group students in same subgroup except for ELLs. There was a statistically significant interaction between ELL level and curriculum condition (treatment vs. comparison). ELLs either did not learn well with Chemistry that Applies curriculum or could not demonstrate their learning on the test. Language minority students made academic achievement gains with treatment condition. For whole group, Chemistry that Applies curriculum improved understanding of content but did not close previously existing achievement gaps between subgroups.
Maheady, Michielli-Pendl, Harter, & Mallette (2006)	Examined the effects of Numbered Heads Together with and without a behavioral incentive package on 6th graders daily quiz scores and pretest-posttest performance in chemistry. *Daily quizzes on state standards-based science instruction; Summative science tests on state standards-based science instruction*	Method: Single case experimental Measures: Pupil satisfaction measure, Daily quizzes, chemistry pre-tests and post-tests Sample: 23 6th graders (8 ELLs) in general education science class	Every student performed better on quizzes and tests when Numbered Heads Together or Numbered Heads Together plus a behavioral incentive were used. However, the addition of a behavioral incentive to Numbered Heads Together improved student performance on daily chemistry quizzes and student satisfaction with instruction over Numbered Heads Together alone or the comparison condition (whole group instruction).

Study	Focus and Connection to Standards	Methods, Measures and Sample	Findings
Medina & Mishra (1994)	Examined the relationship between Spanish reading achievement and native-language academic performance in math, social studies, and science for the total sample of ELLs and fluent Spanish proficient students in a maintenance bilingual program. *No explicit links to standards*	Method: Quantitative non-experimental Measures: La Prueba Spanish reading, math, social studies, and science test scores Sample: 980 7th and 8th graders in a maintenance bilingual program (462 fluent Spanish proficient students, 518 Spanish-speaking ELLs)	Positive, statistically significant correlations were shown for La Prueba subtest scores for all students. There was a significant association between native language proficiency (Spanish) of ELLs and their Spanish academic performance. Moderately high, positive, and significant relationships found between Spanish social studies and science, and between Spanish reading and science support the theory that initial development of native-language literacy provides support for learning content in the same language.
Shaw, J. (1997)	1. How do ELLs respond to performance assessment in science? 2. Are acceptable levels of inter-rater agreement achievable when scoring ELLs' responses to a science performance assessment? 3. Can science performance assessments written in English validly measure scientific literacy among ELLs despite their inevitable dependence on some extent of English language proficiency? 4. How do teachers of ELLs respond to performance assessment in science? *No explicit links to standards*	Method: Qualitative Measures: Field observations before, during, and after implementation of a performance assessment (Rate of Cooling Assessment), student, teacher and administrator interviews, student transcript analysis Sample: 96 high school students (all ELLs)	Both teachers and students had an overall favorable response to the performance assessment, but students' English skills affected their performance on some items. Specific teacher guidance about how to understand the demands of the assessment also significantly affected student scores.

Study	Focus and Connection to Standards	Methods, Measures and Sample	Findings
Torres & Zeidler (2002)	1. What are the effects of English language proficiency and levels of scientific reasoning skills on the acquisition of science content knowledge of Hispanic ELLs participating in grade 10 science classes? 2. Do English language proficiency and scientific reasoning skills interact to influence the acquisition of science content knowledge by Hispanic ELLs and Native English language speaking students participating in grade 10 science? *State science test*	Method: Quantitative non-experimental (Ex post facto) Measures: TOEFL English assessment, test of scientific reasoning skills, state science test Sample: 158 10th graders taking science (134 ELLs)	Significant 2-way interaction between English language proficiency and scientific reasoning skills with regard to performance on the standardized science test. This interaction implies that the combination of high levels of English proficiency and high levels of reasoning skills increase students' ability to learn content.
Varelas, Pappas, Kane, Arsenault, Hankes, & Cowan (2007)	Focus on the concept of matter and explore how young children in urban schools bridge their spontaneous concepts and everyday experiences with scientific concepts introduced to them by children's literature information books and their teachers. *No explicit links to standards*	Method: Qualitative Measures: Field notes, transcripts of videotaped class discourse, student artifacts (e.g., written work, pictures of class artifacts). Sample: 6 classes (2 each of grades 1, 2, and 3) and their teachers (ELLs and language minority students included; unclear number)	Material artifacts used in a sorting activity in science class (e.g., baggie with air, baggie with shaving cream) become devices that promoted children's engagement with science and shaped classroom, communication, discussion, and thinking. Children used four ways of reasoning about states of matter (i.e., macroscopic properties, prototypes, everyday functions, and process of elimination). Children's meaning making was interconnected with their classroom roles, how they worked with each other and how others responded to their ideas.

Appendix B

Study Outcomes and their Relationship to State Standards

	Cognitive Outcomes		Affective Outcomes		Other Outcomes	
Study	Related to State Standards	Other	Related to State Standards	Other	Related to State Standards	Other
Amaral et al. (2002)	Optional science portion of standardized test used for accountability (SAT9)					
Beghetto (2007)				Creative self-efficacy Perceived teacher support Perceived science competence Achievement goals		
Bouchey (2004)		Grades Standardized science test scores (ITBS)		Math importance Time and energy devoted to math and science Perceived math ability Others' perceptions of science ability Reflected appraisals of science ability Reflected appraisals of performance Reflected appraisals of support		
Bouchey & Harter (2005)		Grades Standardized science test scores (ITBS)		Perceived importance of math/ science to others Support for schoolwork Beliefs about the student's competence Importance of schoolwork Scholastic behavior Perceived competence		
Cuevas et al. (2005)					Ability to conduct science inquiry	

	Cognitive Outcomes		Affective Outcomes		Other Outcomes	
Study	Related to State Standards	Other	Related to State Standards	Other	Related to State Standards	Other
Duran et al. (1998)						Learning to talk science Recognition of meaningful concepts, tools and signs
Hadi-Tabassum (1999)				Science attitudes		
Hampton et al. (2001)				Science attitudes		Teacher perceptions of student learning
Lau & Roeser (2002)		Test based on large-scale test items (e.g., NAEP, SAT, NELS: 88, etc.)		Perceived ability to master science content and perform well on tests Task values related to science engagement Intended science choices in college		
Lee et al. (2008a)						Teacher perceptions of student learning
Lee et al. (2008b)	Test based on science curriculum units aligned with standards	Test based on NAEP/ TIMSS items				
Lee et al. (2008c)	State math test	Test based on NAEP/ TIMSS items				
Lee et al. (2006)					Ability to conduct student inquiry	
Lee et al. (2005)	Test based on science curriculum units aligned with standards	Test based on NAEP/ TIMSS items				
Luykx et al. (2007)		Test based on NAEP/ TIMSS items				

	Cognitive Outcomes		Affective Outcomes		Other Outcomes	
Study	Related to State Standards	Other	Related to State Standards	Other	Related to State Standards	Other
Lynch et al. (2002)		Science test based on curriculum aligned with AAAS standards (Conservation of Matter Assessment)		Mastery goal orientation Performance goal orientation		
Maheady et al. (2006)	Summative science tests on standards-based instruction Daily quizzes on standards-based instruction			Student satisfaction with intervention options		
Medina & Mishra (1994)		Spanish reading, math, science, and social studies tests (La Prueba)				
Rivera & Stansfield (2004)	Mean scores on field test portions of state science test					
Shaw (1997)		Science performance assessment (Rate of Cooling or ROC)				
Torres & Zeidler (2002)	State science test	Test of scientific reasoning skills in English or Spanish Language test (TOEFL)				
Varelas et al. (2007)						Meaning making in collaborative science

COLLEGE OF EDUCATION
+ HUMAN DEVELOPMENT

UNIVERSITY OF MINNESOTA

NCEO is an affiliated center of the Institute on Community Integration