REVIEW OF RESEARCH ON THE NEBRASKA LEP PLAN COMPONENTS

March 31st, 2011

Prepared by

Xin Wang, Ph.D.
Susie Bachler, M.A.
REL Central at McREL
for
The Nebraska Department of Education

Table of Contents

1

Executive Summary

In Nebraska, pursuant to Nebraska statutes 79-1014, districts must submit a Limited English Proficiency (LEP) Plan by mid October of each year as specified by statute. The LEP plan is meant to specifically address the needs of Limited English Proficient (LEP) students and will be used in calculation of state funding for the following fiscal school year. District LEP plans are approved or disapproved by the Nebraska Department of Education and are used to determine resource allocation and technical assistance from the state. Districts submitting LEP plans are required to explain how they will address four LEP components identified within the legislation:

1) identification of students with limited English proficiency;
2) instructional approaches;
3) assessment of ELL students' progress toward mastering the English language;
4) an evaluation to determine the effectiveness of the elements of the district's LEP plan.

The Nebraska Department of Education (NDE) requested that REL Central review the research to determine to what extent it supports the components of their required district LEP plan. Specifically, the NDE requested information on research that examines the influence of the components on student performance. NDE also requested that REL Central examine six specific questions (see "Six Questions Guiding the Review") in its review of the literature. The information from this report will also be used by NDE staff to provide technical assistance to districts and support schools in making improvements.

A search of the research literature identified numerous articles relating to these four components. Summaries of findings for each component are presented below.

Identification of Students with Limited English Proficiency

Research has found that identifying LEP students usually involves two steps: A Home Language Survey (HLS) and an English Language Proficiency (ELP) test (Abedi, 2008a; Goldenberg & Rutherford-Quach, 2010; Wolf et al., 2008). (The limitations of these two identification tools are discussed in the detailed review later in this report). Research suggests that the HLS is used in combination with ELP tests as criteria for the classification and reclassification of English Language Learner[1] (ELL) students (Abedi, 2007). According to Wolf et al. (2008), states or districts should weigh the reviews of current ELP test and select the test(s) that is/are best suited to the needs of their students, teachers, and programs. States and districts also need to examine the consistency and validity of the assessment systems in use to ensure that a student classified as ELL at one school will be similarly classified at another (Wolf et al., 2008).

Instructional Approaches

REL Central's review of the literature reveals a variety of ELL instructional methods in content areas including literacy, social science, math, and science. The effectiveness of these instructional strategies have been examined by several studies (e.g., Chang, Singh & Filer, 2009; Brown, 2007; Honnert & Bozan, 2005; McLaughlin et al., 2000; Saunders & Goldenberg, 1999; What Works Clearinghouse, 2006). Research suggests that teachers should combine interactive

---

[1] Research on students who are not proficient in English typically refers to students either as Limited English Proficient (LEP) or English Language Learners (ELL). These terms are interchangeable; for each study reviewed, we have retained the term used in that study.

and direct instruction approaches in an ELL classroom to provide instruction in specific literacy skills within interactive contexts. When teaching content areas other than literacy, studies suggest that teachers should provide visual aids, use students' native language, encourage student collaboration, and use other instructional activities to make the content knowledge more accessible to ELL students.

Five frequently used LEP program models are discussed in the current review. We report on the program models' underlying goals, the degree to which students' first language is used and maintained, resources and local conditions required, and the manner in which they are implemented at different grade levels as discussed in Genesee (1999). Despite empirical evidence in favor of long-term bilingual instruction (Collier & Thomas, 2004; Rolstad, Mahoney & Glass, 2005), these authors suggest that the choice of an LEP program model should be based on the needs of the students involved and the resources available. States and districts need to consider multiple factors, including the proportion of LEP students in the classroom or school, to determine the type of program most likely to be appropriate for their specific situation.

Assessment of ELL Students' Progress toward Mastering the English Language

In this report we review the research on two types of assessments that have been used to measure ELL students' language progress: standardized achievement tests and English Language Proficiency (ELP) tests, and discuss the limitations of the existing measures (Abedi, 2008a; Bailey & Butler, 2003; Del Vecchio & Guerrero, 1995; Garcia, McKoon, & August, 2006; Mahoney & MacSwan, 2005; Wolf et al. 2008). Research suggests that standardized achievement tests should not be used as the only criterion to assess ELL students for classification purposes due to the linguistic complexity of test items and the lack of comparability in determining levels of English proficiency. In order to provide valid and reliable measures of ELL proficiency, states and districts should select ELP assessments that incorporate academic language and align the content of ELP assessments with the academic language requirements of the state academic assessment and ELP standards (Bailey & Butler, 2004). Research also suggests that ELP assessment results should be used with caution or combined with other criteria for ELL redesignation.

An Evaluation to Determine the Effectiveness of the Elements of the District's LEP Plan

Despite the importance of evaluating districts' language instructional programs for compliance with Title III requirements, we found that few studies have examined evaluation approaches for ELL programs implemented in states and districts. Therefore, we chose to review ELL program evaluation conducted by various states and districts and discuss how program evaluation can be conducted in ways that meaningfully affect the achievement of ELL students (Gonzalez, 2001; Iribarren, 2009). Research found that evaluations of ELL programs often gather information from various stakeholders, include components that address Title III requirements, and collect evaluation information from multiple data sources. According to Llosa and Slayton (2009), the design of an effective evaluation of ELL instructional program should incorporate the following four features: 1) an investigation of the context; 2) the use of multiple types and sources of data; 3) the use of appropriate analytic tools; and 4) the use of extensive qualitative data.

Six Questions Guiding the Review

1. *What is the literature concerning the selection of a program/service for districts with low numbers of limited English proficient students and/or limited resources? For example, does the literature indicate that there is a program/service preference for schools or districts with very small numbers of students or limited resources?*

   Our review found that schools adapt their instructional models to the size and the needs of their ELL population. For example, pull-out ESL models are adaptable to changing populations or schools that have new ELLs at different grade levels; therefore, this model is often implemented in low-incidence schools or to serve students who do not share the same first language (Valdez, Svedkauskaite, & McNabb, 2002). Han and Bridglall (2009) also reported that, given a small number of ELLs with diverse native languages, schools tended to favor pull-out ESL programs.

   Previous work finds that it is important to provide teacher education programs for ESL/bilingual teachers and also to provide non-ELL teachers with training in working with LEP students (Menken & Antunez, 2001). This training is particularly urgent in schools with smaller proportion of ELL students, as the teachers in these schools may be less racially diverse and receive less professional development related to teaching ELL students.

2. *What are the strengths and weaknesses of the researcher-identified program/service models?*

   REL Central reviews the five most-frequently used ELL program models in the United States. The five models are assigned to two major categories: bilingual models and English-as-a-second-language (ESL) programs (Zelasko & Antunez, 2000). The strengths and weaknesses of these program models are discussed on pages 7 to 10.

3. *What is the literature that describes evaluation approaches for identified program/service models? What is the linkage between the instructional programs and the approach implemented to evaluate the effectiveness of the identified program/service?*

   Despite the importance of evaluating districts' language instructional programs for compliance with Title III requirements, few studies have examined evaluation approaches and methodologies for this purpose. In this report we review ELL program evaluations conducted by various states and districts and discuss how program evaluation can be conducted in ways that meaningfully affect the achievement of ELL students.

4. *What elements should be included in a good evaluation plan for an LEP program/service?*

   ELL program evaluations typically share three common features (pp. 19-20)

1) evaluation procedures of language assistance programs often involve multiple stakeholders to gather information through written reviews, surveys and/or interviews;
2) many evaluation plans include components that address the requirements of Title III;
3) in the evaluation plans reviewed, evaluation information is often obtained from multiple data sources.

Llosa and Slayton (2009) found that an evaluation of ELL program models should include an investigation of the context, the use of multiple types and sources of data, the use of appropriate analytic tools, and the use of extensive qualitative data.

5. *What guidance does the literature provide regarding how to use information about former LEP students as part of the evaluation of the district's current LEP program/services?*

Two CRESST reports argue that states/districts should examine their reclassification rates and evaluate the effects of their reclassification policies. The researchers recommend that states, districts, and schools monitor newly redesignated ELL students and maintain a tracking system to do so. Data from this practice could be used to evaluate potential consequences for reclassified and long-term ELL students, hence providing evidence to support or modify the states' redesignation criteria.

6. *How should a district use local and statewide results of a state test as part of the evaluation of the program/services?*

In the evaluation plans reviewed, LEP program decision making in schools and districts is data driven. Therefore, the outcomes of state assessments should be included in the evaluation plan as one way to evaluate the effectiveness of an ELL program (Llosa & Slayton, 2009). For example, the Idaho LEP Program Evaluation Guidance (2006) included data from ELP assessment and academic achievement. Gonzalez (2001) assessed program effectiveness by comparing outcome indicators (e.g. achievement outcomes) for LEP students being served and for LEP students whose parents refused program services (pp. 19-20).

This report provides a general overview of areas for Nebraska to consider when reviewing the LEP component plans in Nebraska districts. The literature reviewed represents only a fraction of available research. Therefore, readers are encouraged to directly read the cited articles and explore additional sources for further information. It is also important to note that the quality and quantity of the research vary considerably across the components. For example, for the component of LEP program evaluation, limited information and research is available. This limitation does not necessarily indicate a lower effect, but rather the need for more research in this area. Sample LEP program guidelines and evaluation plans from states and districts are provided in Appendix A and Appendix B.

Introduction

The Nebraska Department of Education (NDE) requested information from REL Central's reference desk on the research basis of the components of their required district LEP plan. As part of calculating state funding for a district, Nebraska requires districts to submit an LEP Plan, which details how they will work to educate students with limited English proficiency. The instructions for the plans specify four components that school districts must address when developing their plans: identification of Limited English Proficient (LEP) students, instructional approaches, assessment of student progress toward English proficiency, and evaluation of the effectiveness of districts' LEP programs (Schoonover, 2010).

This report provides a review of the research on these four components of the LEP plans required by the state statue.

Literature Search Methods

REL Central conducted multiple searches to identify relevant, rigorous research for each of the four components. The resources searched included: Google Scholar, bibliographic databases (e.g. ERIC, JSTOR, Academic Search Premier, Educator's Reference Complete, and Proquest), IES-supported organization websites (e.g. the Regional Educational Laboratories website, and the What Works Clearinghouse), websites of nationally-known research and policy organizations, and reference lists of documents identified through searching. The initial searches used the terms "Limited English Proficient" and "English Language Learners."

REL Central submitted a preliminary study outline and a list of articles found, as well as an annotated bibliography, to NDE for draft review. After NDE's review of these initial materials, additional searches using more refined search terms were conducted to ensure that all relevant sources were obtained.

This review examines the degree to which the components in the Nebraska LEP Plan are supported by research related to improving student achievement.

Studies that have been included in the review also address the following questions, as requested by the state.

1. What is the literature concerning the selection of a program/service for districts with low numbers of limited English proficient students and/or limited resources? For example does the literature indicate that there is a program/service preference for schools or districts with very small numbers of students or limited resources?

2. What are the strengths and weaknesses of the researcher-identified program/service models?

3. What is the literature that describes evaluation approaches for identified program/service models? What is the linkage between the instructional programs and the approach implemented to evaluate the effectiveness of the identified program/service?

4. What elements should be included in a good evaluation plan for an LEP program/service?

5. What guidance does the literature provide regarding how to use information about former LEP students as part of the evaluation of the district's current LEP program/services?

6. How should a district use local and statewide results of a state test as part of the evaluation of the program/services?

Initial Review of Materials Found

The initial literature searches identified articles, books and other documents such as conference proceedings, journal articles, and newsletters. REL Central researchers reviewed the abstracts of these materials to determine the relevance to the component and confirm that the material was a research piece. After reading each article or document, we categorized the articles into the four component areas. Although the majority of the documents are from 2000 through 2010, earlier documents were included in some instances if they were works from seminal authors and key organizations, or were cited in other studies that were included in the review.

The articles included in this review varied in type and rigor and included meta-analyses of multiple studies, literature reviews, and single studies that used a variety of methods (i.e. quasi-experimental design, descriptive research). The method used, as well as how it was used, determines the nature of the conclusions that can be drawn.

Findings: the Four Components

Each of the component sections that follow begins with a general introduction as background for that component. The introductory material is followed by topical sections describing the relevant research.

Identification of students with limited English proficiency

Studies in this area focus on the criteria and processes used to identify ELL students. Two reports from the National Center for Research on Evaluation, Standards, and Student Testing (CRESST), provided a comprehensive review of the procedures of ELL identification and classification and offered recommendations for state policy and practice to assess ELL students. The other studies reviewed examined approaches and criteria to classify and reclassify ELLs.

Definition of LEP Students

NCLB defines LEP students as students who speak a native language other than English by foreign birth or ancestry, who live in an environment where the primary language is other than English, and who have difficulties with speaking, reading, writing, or understanding the English language that interfere with social interactions and academic tasks (NCLB, 2001). Limited English Proficient (LEP) and English Language Learner (ELL) are the terms most commonly used by states to describe students who are not fully fluent in English. Since NCLB uses the term LEP in its documentation, most states have aligned their terms and state definitions of ELL with those provided by NCLB. Because many states use both terms in state literature, documents, and official state definitions (e.g. Nebraska, as seen in Wolf et al., 2008), these two terms are used interchangeably in the following sections of the review.

Identification of LEP students

The identification of LEPs and ELLs is based on two sources of information: students' language background information and their level of English proficiency (Abedi, 2008a). Therefore, identifying LEP students usually involves two steps: a Home Language Survey (HLS), and an English Language Proficiency (ELP) test (Abedi, 2008a; Goldenberg & Rutherford-Quach, 2010; Wolf et al., 2008).

HLS, a brief questionnaire on which parents indicate what language or languages are spoken at home, has been commonly used by schools to identify ELL students. According to results of the *Survey of the States' Limited English Proficiency students*, among the 54 state educational agencies surveyed, 90% asked parents to complete an HLS (Kindler, 2002). Previous research, however, questioned the use of the HLS as a single source for identifying ELL language proficiency and raised concerns regarding the validity of HLS data (Abedi, 2007). The validity of a self-reported home language survey was questioned on the basis that parents might not provide accurate information, being concerned about equity of opportunities for their children (Abedi, Lord & Plummer, 1997; Abedi, 2008a) or might not accurately comprehend the survey items (Littlejohn, 1998). Critics also argued that the sole use of home language surveys could over-identify students as ELLs because the use of a language other than English at home did not necessarily indicate that a student would have limited English proficiency (Littlejohn, 1998).

In addition to the information collected from an HLS, student outcomes from ELP tests are used as criteria for the classification and reclassification of ELL students. According to Kindler (2002), 94% of the surveyed state educational agencies (51 of 54) asked their students to take ELP tests for ELL classification placement. Although an ELP assessment is an important aspect of the ELL identification process, researchers are concerned about the validity and reliability of ELP assessment outcomes.

Disparities among the ELP tests developed prior to NCLB are one of the major causes of concern regarding the accuracy and consistency of the measures used to classify ELL students. Zehler, Hopstock, Fleishman, and Greniuk (1994) compared six most frequently used ELP tests for 1) content and structure, 2) test administration procedures, 3) theoretical bases, and 4) test validity and reliability. These tests were 1) Idea Proficiency Test (IPT)-I; 2) Language Assessment Scales (LAS) 1C (Oral); Language Assessment Battery (LAB) IIA; 4) Bilingual Syntax Measure (BSM) II; 5) Maculaitis Assessment Program (MAC), Level 2-3; and 6) Peabody Picture Vocabulary Test (PPVT)-Revised. They found these tests differed on all of the four areas compared. The tests also differed in types of tasks and specific item content. For tests that required the same type of response and similar item content, the test items were actually assessing different language skills as the scoring criteria focused on separate aspects of the response. This review also found that these tests differed on the definition of language proficiency, the target range of grade level, and the specific time limits (e.g. test length is fixed or determined by student's responses). Del Vecchio and Guerrero (1995) reviewed five major ELP tests and reported similar findings. These tests varied with respect to test purpose, target age and language groups, test administration procedures, theoretical bases, and test reliability and validity (see also in Abedi, 2008a).Even though these test reviews did not qualify or criticize the existing ELP tests, they suggested that states or districts should weigh the pros and cons of the various ELP tests and select the test best suited to the needs of their students, teachers, programs, and schools.

In order to comply with NCLB requirements, states have recognized the importance of alignment of constructs between their ELP standards and assessments used to classify ELL students. As a result, post-NCLB ELP assessments were development by four consortia of states: Mountain West Assessment Consortium (MWAC), Pennsylvania Enhanced Assessment Group (PA EAG), State Collaborative on Assessment & Student Standards (SCASS) Consortium, and World-Class Instructional Design and Assessment (WIDA) Consortium. These new ELP assessments include Mountain West Assessment Consortium (MWAC) assessment, Comprehensive English Language Learning Assessment (CELLA), English Language Development Assessment (ELDA), and Comprehension and Communication in English State-to-State for English Language Learners (ACCESS for ELLs®). The test developers have carefully examined and revised as necessary the content and psychometric properties of these post-NCLB ELP tests. In the case of ELDA, for example, ELP standards were developed based on previously existing state-level standards, which became the theoretical construct of the ELDA. Field tests data were analyzed to provide validity and reliability evidence. Additionally, relationships between the ELDA, the Language Assessment Survey (LAS), the New IDEA Proficiency Test (New IPT), and teacher ratings of student proficiency were examined using multitrait multimethod analyses. The fact that ELDA scores were more closely associated with the teacher ratings was considered evidence of test validity because ELDA was designed to measure classroom language ability. Item bias/fairness was first addressed at the item development stage where a trained review expert examined potential test items and recommended modifications or the removal of potentially problematic items. Differential item functioning (DIF) analyses were further conducted using the field test data. Relatively few items were found to have DIF and of these, only a few were removed after close consideration of the item content (Lara et al., 2007; Wolf et al., 2008).

In summary, test developers of these new assessments have addressed potential validity questions through applied research. They have also incorporated attention to bias at both the item design stage and at the field trial stage, and used a rigorous screening process. In addition, the developers compared the tests against multiple validity criteria rather than against a single criterion. All of these efforts have built a strong foundation for the new ELP assessments as a tool to identify and classify ELL students (Wolf et al., 2008).

Recent articles have addressed other issues that also influence the accurate identification of ELL students. According to Wolf et al. (2008), 16 states allow local districts to choose an ELP test from a list of approved assessments. This practice may raise issues of comparability regarding levels of proficiency determined using different tests. In other words, students may be classified at different proficiency levels within a state, depending on the test that a student takes. If districts/schools use multiple tests to classify ELL students, Wolf and her colleagues suggest that evidence for the comparability of the constructs and cut score settings among these tests should be provided to justify the local decisions on classification of ELLs. In addition, a number of states use the same ELP test to identify ELL students and to monitor the progress of students' English language development. This requires states to employ different validation procedures for their assessment systems. If one test is used to serve multiple purposes, validity evidence should be offered for each intended use (Wolf et al., 2008).

Early research indicated that the use of other factors also influenced decisions on ELL classification and reclassification. These factors included socioeconomic status (SES), ethnicity, parent education, teacher evaluation, parent opinion and consultation, and performance in basic

skills (Abedi, 2004; Grissom, 2004). Kindler (2002) found that, among the 54 state educational agencies surveyed, over 80% made use of teacher observation, teacher interviews, and parent information in the identification of ELL students, and 60% or more used student records, student grades, informal assessments, and referrals. She reported that sixteen states used state-developed achievement tests. Criterion-referenced tests were used in 21 states, and 19 states used other assessments, including portfolios, cognitive assessments, reading/writing evaluations, and a variety of locally-designed tests. States now rely on more data-driven sources to make decisions for classification and reclassification.

Although different systems or criteria have been used for classifying and reclassifying ELL students within and across states (Abedi, Lord, Hofstetter, & Baker, 2000; Rivera, Stansfield, Scialdone, & Sharkey, 2000; Linquanti, 2001), the CRESST report (Wolf et al., 2008) suggested that states and school districts adopt a uniform approach in determining a student's level of English proficiency to make sure that one student classified as ELL at one school will be similarly classified at another school.

Instructional Approaches

National achievement data indicate an achievement gap between ELLs and their non-ELL peers. For example, on the 2007 National Assessment of Educational Progress (NAEP), there is a 36-point discrepancy between the average reading score (188 out of 500) for fourth grade ELLs and the average score (224) of their non-ELL peers (U.S. Department of Educaiton, Institute of Education Science, National Center for Education Statistics [NCES], 2007; Slavin & Cheung, 2003; Movit, Petrykowska, & Woodruff, 2010). Although there is a clear need for research examining factors that contribute to the gap in student achievement, there is limited knowledge about the quality of instruction and the actual amount of support ELLs receive (Goldenberg, 2008). This section reviews the research focusing on:

1) instructional methods used to teach ELLs in diverse content areas,
2) description of different instructional models designed for ELLs,
3) the effectiveness of LEP program models.

Instructional Methods

Research on instructional practices has examined a wide variety of different methods, techniques, and strategies for promoting the literacy skills of ELLs. Genesee, Lindholm-Leary, Saunders, and Christian (2005) categorized ELL literacy instructions into three approaches: 1) direct instruction, which entails the explicit instruction of specific literacy skills and strategies; 2) interactive instruction, mediated through interaction with other learners or teachers; and 3) process-based instruction, which emphasizes engagement in the authentic use of written language for communication.

Direct and interactive instruction are frequently combined: direct skill instruction is often included in interactive learning environments (Doherty, Hilberg, Pinal & Tharp, 2003; Saunders & Goldenberg, 1999; see also in Genesee et al., 2005). Saunders and Goldenberg (1999) found that interactive and direct instruction should be combined in ELL classrooms, because they provide instruction in specific reading and writing skills within interactive contexts.

Direct instruction and interactive instruction have also been found effective in previous studies (e.g., McLaughlin et al., 2000; Saunders & Goldenberg, 1999; What Works Clearinghouse,

2006). August and Shanahan (2006) found that substantial instruction in key components of reading, such as phonemic awareness, phonics, fluency, vocabulary, and text comprehension, has a positive influence on the literacy development of ELLs. A reading study (Carlo et al., 2004) found that teaching academically useful words to ELLs enhanced fifth graders' knowledge of the words taught, depth of vocabulary knowledge, understanding of multiple meanings, and reading comprehension. The underlying principles of this intervention are that new words should be encountered in meaningful text, that words should be encountered in varying contexts, and that word knowledge should involve spelling, pronunciation, morphology, and syntax as well as depth of meaning. Baumann and Graves (2010) also examined the importance of teaching academic vocabulary and found that content-area teachers use a categorized academic vocabulary for instruction. Additionally, Bauer and Manyak (2008) found that that the social contexts and social interactions of language should be emphasized in a language-rich classroom environment. They concluded that classrooms should be places where students could take linguistic risks through classroom interaction. These findings suggest that it may be feasible to improve comprehension outcomes for ELL students in mixed classes by teaching word analysis and vocabulary learning strategies.

In contrast, the effectiveness of process-based instruction still remains questionable. Researchers who examined process-based instruction found that simply exposing students to literacy-rich learning environments was not sufficient to promote acquisition of the specific skills that comprise reading and writing. They found that focused and explicit instruction in particular skills and sub-skills was called for if ELLs are to become efficient and effective readers and writers (Kucer & Silva, 1999; Genesee et al., 2005).

Other researchers examined the impact of various ELL instructional methods on content area subjects other than reading and writing; e.g., social science, math, and science.Yet, the research community has not reached a consensus regarding the best practices for schools and districts (Thomas & Collier, 2002). A recent study from Chang, Singh and Filer (2009) examined the effects of classroom achievement grouping practices on the early mathematics performance of ELL students. This study indicated a significant negative effect of achievement grouping on ELL students. Cirillo, Bruna, and Herbel-Eisenmann (2010) investigated linguistic challenges that ELL students faced in their math education. Other researchers studied the use of cooperative learning (Campbell & Rowan, 1997); use of the native language (De la Cruz, 1998); development of language skills (Ron, 1999); and use of visual representations and manipulatives (De la Cruz, 1998).

Recent studies have focused on diverse approaches to developing an effective instructional environment for teaching social studies curriculum to ELLs. For example, Brown (2007) investigated the difficult nature of social studies texts and found that the following concrete strategies make social study texts more comprehensible for ELLs: 1) content maps, 2) outlines of a unit, 3) guiding questions, and 4) simplified text. Szpara and Ahmad (2007) examined a multi-tiered approach to meet the needs of ELLs in mainstream social studies classrooms. They concluded that teachers should 1) provide ELL students with social and cultural support, 2) use Cognitive Academic Language Learning Approach[2] to provide explicit instruction to help

---

[2] The Cognitive Academic Language Learning Approach (CALLA) is an ELL instructional model based on cognitive theory and research. The purpose of CALLA is to improve ELLs' academic achievement by teaching them essential academic content, fostering their language development and focusing on explicit instruction in learning strategies (Chamot & O'Malley, 1999).

students comprehend in-depth content, and 3) reduce cognitive load to make social studies curricula more accessible to ELL students.

Researchers have also examined methods that teachers can use to improve the effectiveness of science education for ELL students. Himmel, Short, Richards, and Echevarria (2009) found that lesson planning and delivery of science instruction should include approaches such as the Sheltered Instruction Observation Protocol Model (SIOP). The SIOP Model provides teachers a systematic way to address and accommodate the language and academic learning needs of ELL students. When delivering lessons, teachers using SIOP paid close attention to building background knowledge and ensuring that ELL learners comprehended the input (Echevarria, Vogt, & Short, 2004). Honnert and Bozan (2005) reported the positive impact of teaching summarization on ELL students' science achievement. They found that teaching summarization as a reading strategy increased students' abilities to acquire and use information and better comprehend science concepts. In combination with other vocabulary attainment activities, teaching summarization enhanced the students' ability to apply information to discussions, laboratory reports, and projects. Edmonds (2009) noted that cultural differences impeded the science education of ELL and suggested that teachers engage students through knowledge of their native culture. Teachers also should provide outlines and visual aids to make lectures and lessons more clear for ELLs. Her findings suggest that student collaboration and visual communication methods can encourage ELLs to engage in class discussions. Greathouse and Lincoln (2008) also found that the use of visual aids and an emphasis on students' prior experiences help convey information to ELL students. These findings suggest that experiential learning can help ELLs understand science terminology.

LEP Program Models

Zelasko and Antunez (2000) reviewed the program models for ELLs and put them into two major categories: bilingual models and English-as-a-second-language (ESL) programs. The former employs native language instruction as students develop English language proficiency, whereas the latter uses only English as a medium to provide comprehensive instruction. Among the existing program models, five are frequently used in schools in the United States (Zelasko & Antunez, 2000). These models differ regarding the underlying goals, the degree to which students' first language is used and maintained, resources and local conditions required, and the manner in which they are implanted at different grade level (Genesee, 1999). Most schools use a combination of approaches and adapt their instructional models to the size and the needs of their ELL population.

The following section reviews the features of five ELL program models:

1. Two-way immersion,
2. Developmental bilingual education,
3. Transitional bilingual education,
4. Sheltered Instruction or structured immersion, and
5. Pull-out ESL.

Bilingual Models

1. Two-way immersion programs, also known as *two-way bilingual* and *dual language immersion* programs (Genesee, 1999), strive to develop bilingualism in ELLs and English-proficient students. Two-way immersion programs usually start in kindergarten or first grade and continue

for four to six years. Most programs typically include 50% native English speakers and 50% ELLs who share the same native language and require that the non-English language be used for at least 50% of instruction (Genesee, 1999; Sugarman & Howard, 2001).

This model incorporates native speakers of English into ELL instruction; hence it results in language proficiency in both English and another language. Two-way immersion also promotes cultural awareness and the value of bilingualism. This model is feasible in schools with large populations of ELLs that share the same first language. It works best when English and the first language have equal status and are integrated into all domains of schooling. In addition, two-way immersion requires qualified staff member who are proficient in least one language, certified in content areas, and able to integrate content with language objective (Mikow-Porto, Humphries, Egelson, O'Connell & Teague, 2004).

2. Developmental bilingual education programs are also known as *late-exit bilingual* programs. Described as additive programs, developmental bilingual programs aim to develop ELL students' first language while helping them to acquire English proficiency and master academic content knowledge. To achieve this goal, this model usually starts in kindergarten or first grade, and gradually transitions the language of instruction from the first language of ELLs to English as students' English language skills develop (Genesee, 1999). This model works best in schools with a stable population of ELLs who speak a common native language. Since prior research indicates that ELLs may need four to seven years of bilingual instruction to close the achievement gap (Mikow-Porto et al., 2004), an ideal developmental bilingual program should be implemented through high school.

Research has found that developmental bilingual programs promoted the long-term cognitive development of ELL students, especially when their first language was developed through Piaget's formal operations stage at puberty (Genesee, 1999; Mikow-Porto et al., 2004). This model uses students' first language to engage them in cognitively challenging work and English as a second language to engage students in meaningful content. Studies have found that developmental bilingual programs were one of the most effective ELL models in leading to increased academic achievement (Thomas & Collier, 1997; Genesee, 1999).

3. Transitional bilingual education program, also known as *early-exit bilingual* education, is the most common form of bilingual program for ELLs in the United States (Genesee, 1999). The purpose of this model is to increase students' academic achievement while learning English as a second language. A transitional bilingual program usually starts in kindergarten, using students' first language and English for instruction at the beginning of the program. Students' first language skills are developed to a limited extent and are phased out once they obtain language proficiency in English. Most students in this program are mainstreamed into regular classrooms within three years (Genesee, 1999; Montecel & Cortez, 2002).

Like the developmental bilingual model, transitional bilingual programs require that a stable population of ELLs share the same first language. This model works well in schools where teachers in the transitional bilingual program are proficient in both languages and know how to employ sheltered instruction strategies to provide effective instruction and meaningful activities. Due to the transitional nature of this model, schools/districts also need to provide follow-up programs to ensure that former ELLs perform as well as their English-speaking peers after they are redesignated.

4. *Sheltered instruction* or *structured immersion* integrates academic content objectives and language development goals. This model is appropriate when serving ELLs with a variety of language backgrounds as well as students at different ELP levels. Instruction in this model is delivered in English and adapted to students' proficiency level. The focus of sheltered instruction is on the content-area curriculum. It incorporates content clues, such as demonstrations, graphic organizers and visual aids, into instruction to make content more accessible to ELL students. This model also emphasizes teacher-student interaction, which is characterized by discussion, sufficient wait time to encourage student responses, explicit teaching, and the use of meaningful activities to practice conversational skills (Genesee, 1999).

This program works well at different grade levels. As it integrates content and language objectives, this model requires districts to create grade-level curricula for each subject area with well-defined objectives and standards. In addition, sheltered instruction requires teachers to use strategies to make instruction comprehensible to ELL students. Teachers in this program need to receive training in sheltered instruction strategies, ESL methodologies, and second-language acquisition processes, as well as training in content areas. This model works best in schools and districts that allow students to enter mainstream classes one subject at a time, as well as those that offer programs to support former ELLs after they have been mainstreamed (Genesee, 1999).

5. *Pull-out ESL* focuses on students' English language acquisition only. In this program, students leave mainstream classrooms to spend part of their day receiving small-group ESL instruction. The amount of time ELLs spend in pull-out classes varies from 30 minutes per day up to half a day (Thomas & Collier, 1997). Like content-based ESL, this model works best when students are grouped by language proficiency level (e.g., Pre-functional, Beginning, Intermediate, Advanced, and Fluent English Proficient). Pull-out ESL is adaptable to changing populations or schools that have new ELLs at different grade levels; therefore, this model is often implemented in low-incidence schools or to serve students who do not share the same first language (Valdez, Svedkauskaite, & McNabb, 2002).

Rolstad, Mahoney and Glass (2005) found that pull-out ESL was among the least effective programs in providing comprehensive academic skill development. A disadvantage is that pulling students out can waste instructional time and students will miss class work while they are gone. If pull-out instruction is not closely coordinated with the content taught in the mainstream classroom, ELL students may fall behind in content areas while improving their English language skills. In this case, it is extremely important for ESL teachers to work closely with classroom teachers to ensure that the students are not being graded down for missed work.

A summary of program model features appears in Table 1.

Table 1: Features of Five ELL Program Models Implemented in the United States.

ELL Program Model	Program Features
Two-way immersion	- Result in language proficiency in both English and the first language. - Promote cultural awareness and the value of bilingualism. - Feasible in schools with large populations of ELLs that share the same first language. - Require qualified staff members proficient in least one language, certified in content areas, and able to integrate content with language objective.
Developmental bilingual education	- Promote the long-term cognitive development of ELL students. - Develop ELL students' first language while helping them to acquire English proficiency and master academic content knowledge - Can be difficult for schools with high student mobility. - Require qualified bilingual staff. - An ideal developmental bilingual program should be implemented through high school.
Transitional bilingual education	- Increase students' academic achievement while learning English as a second language. - Most students are mainstreamed into regular classrooms within three years. - Does not work well in schools with students from multiple language backgrounds. - Require qualified bilingual teachers. - Schools/districts also need to provide follow-up programs to ensure that former ELLs perform as well as their English-speaking peers after they are redesignated.
Sheltered Instruction or structured immersion	- Work well when serving ELLs with a variety of language background as well as students at different ELP levels - Promote teacher-student interaction. - Teachers need to receive training in sheltered instruction strategies, ESL methodologies, and second language acquisition processes, as well as training in content areas. - Works best in schools and districts that allow students to enter mainstream classes one subject at a time. - Schools/districts need to offer programs to support former ELLs after they have been mainstreamed.
Pull-out ESL	- Adaptable to changing populations or schools that have new ELLs at different grade levels. - Work well in low incidence schools or to serve students who do not share the same first language. - Can waste instructional time and students will miss class work while they are gone. - ESL teaches need to work closely with classroom teachers to ensure that the students are not being graded down for missed work.

Effectiveness of LEP Program Models

Prior research indicated that schools with language-related instructional programs help promote ELL students' academic achievement. More recent studies found that bilingual fluency has a

positive impact on student outcomes, including higher math and reading scores (Golash-Boza, 2005; Portes & Hao, 2002); greater cognitive flexibility and abstract thinking skills (Rumbaut, 1995); and access to positive cultural capital in students' families and communities (Portes & Rumbaut, 2001).

A meta-analysis conducted by Rolstad, Mahoney and Glass (2005) provided a comprehensive review of ELL program effectiveness. The analysis found that bilingual education was more beneficial for ELL students than all-English approaches. Students in long-term developmental bilingual programs had higher achievement scores than those in short-term transitional bilingual programs. After controlling for ELL status, this study reported a small positive effect for bilingual program with a .23 effect size, while the outcome measures in the native language program resulted in a positive effect size .86. This meta-analysis reported that bilingual education outperformed English-only approaches in increasing outcomes of student academic achievement both in English and in their native language. These findings were also found in a study by Collier and Thomas (2004), who concluded that such language programs should be implemented for at least six to eight years in order for the long-term benefits on ELL students' academic performance to become apparent.

Despite the research evidence in favor of long-term bilingual instruction (Collier & Thomas, 2004; Rolstad et al., 2005), the effectiveness of various ELL program models remains the subject of debate. Although there may be a basis to claim that one program model is more effective than another in certain situations, the choice is often made at the local level after careful consideration of the needs of the students involved and the resources available. Han and Bridglall (2009) reported that, given a small number of ELLs with diverse native languages, schools tended to favor the pull-out ESL program. In addition, they found that transitional or remedial language programs designed to quickly move ELL students into English-only instruction were still prevalent in the United States.

It is therefore important to consider various factors that ultimately influence the type of program most likely to be appropriate and effective in a given situation. These factors include school characteristics, student characteristics, and district or school resources. It is notable that schools with varying concentrations of ELL students have different characteristics at school, principal and teacher levels, hence affecting their selection of instructional models. For example, schools with a low concentration of ELL students are less likely to provide language instruction in students' native languages and adapt instruction to students' individual needs (Zehler et al., 2003).

This finding is supported by a study (Cosentino de Cohen, Deterding, & Clewell, 2005) that analyzed national data from the 1999-2000 *Schools and Staffing Survey* for public elementary schools. This study found that nearly 70% of ELL students were enrolled in only 10% of the U.S. elementary schools, with an average of 47% of ELLs in schools with more than 25% ELLs in the student population (High-LEP schools) and 5% in schools with fewer than 25% ELLs in the student population (Low-LEP schools). Low-LEP schools tended to be smaller suburban and rural schools where students were mostly white. As compared to High-LEP schools, Low-LEP schools offered fewer remedial programs, less parental outreach and support, and less native-language instruction. In addition, these schools were less likely to use standardized processes to identify LEP students and to provide specialized instruction. It was reported that High-LEP schools were three times more likely to use ELL students' native language in instruction than

11

Low-LEP schools. Principals and teachers from Low-LEP schools tended to be less racially and ethnically diverse and earned less than their counterparts in High-LEP schools. Teachers in Low-LEP schools were less likely to hold ESL/bilingual certification in addition to their main certification. These teachers also reported receiving less professional development in content, methods and assessments than their peers in High-LEP schools.

These findings point to the importance of improving teacher education programs for ESL/bilingual teachers and also to providing non-ELL teachers with training in working with LEP students (Menken & Antunez, 2001). This training is particularly urgent in Low-LEP schools, as the teachers in these schools are less racially diverse and receive less professional development related to teaching ELL students.

Pedagogical Knowledge and Skills in an Effective Teacher Education Program

Research has found that for ELL students to effectively comprehend and participate in curriculum and instruction, teachers must use methods tailored for these students (Abedi & Herman, 2010; Zehr, 2010). Additionally, there is a growing consensus among English language educators that teachers of ELLs require specialized skills in order to accommodate their students (Antunez, 2002). White, Makkonen and Stewart (2010) select six topics to organize the pedagogical knowledge and skills that appear important for general education teachers to know and be able to do to support ELL students. The knowledge and skills covered under each of these six topics and supporting literature are identified below.

1. *Recognizing or supporting diversity*. Teachers of ELLs need specialized knowledge and skills related to 1) making content comprehensible; 2) integrating language with content instruction; 3) respecting and incorporating first languages; 4) recognizing how culture and language intersect with classroom participation; and 5) understanding the needs of students with different levels of formal schooling (Antunez, 2002; Menken & Look, 2000).

2. *Differentiating instruction.* Teachers of ELL students should use differential instruction to ensure that the class is challenging and appropriate for students at multiple developmental levels (Adams & Pierce, 2003). Brooks and Thurston (2010) found that small-group work in middle school content area classes led to more frequent academic language production among ELL students than when using whole class instruction. Further, teaching mathematics with primary language assistance from an instructional aide or the teacher is positively correlated with higher mathematics achievement among English language learners (Williams et al. 2007).

3. *Selecting approaches, materials or curricula.* Language, as the medium of instruction, serves a critical function in education. Therefore, both ELL teachers and general education teachers need to be able to plan and implement instruction that simultaneously incorporates both language lessons and content lessons (August & Shanahan, 2006; Gersten, et al., 2007; Himmel, Short, Richards, & Echevarria, 2009).

4. *Knowing language acquisition theories and strategies.* General education teachers need an understanding of language and language acquisition if they are to support ELL students. These skills include knowing "basic constructs of bilingualism and second language development, the nature of language proficiency, the role of the first language

and culture in learning, and the demands that mainstream education places on culturally diverse students" (Clair & Adger, 1999, p. 2). Filmore and Snow (2000) suggest that teachers should have knowledge of a number of topics regarding oral and written language (e.g., the basic units of language, complexity of English spelling, patterns of rhetorical structure). Teachers not only need to be able to model and recognize proper grammar, they "need to be able to explain grammatical errors in ways that can help ELLs edit their writing" (Lucas & Grinberg, 2008, p. 617). Additionally, in order to teach the language of a discipline (e.g., vocabulary, narrative styles), teachers need to understand the particular linguistic characteristics, features and structure of the discourse used in any discipline area they are teaching (Aguirre-Munoz & Amabisca, 2010). Research also finds that teacher preparation coursework and experiences need to focus squarely on the linguistic characteristics, features and structure of the language used in effective teaching (Lucas & Grinberg, 2008).

5. *Communicating with students and families.* It is essential to communicate with families in order to develop the most comprehensive and accurate understanding of individual students' first and second language acquisition and school history. ELL students' backgrounds are so diverse that there is no simple one-size-fits-all solution to providing appropriate supports and interventions for ELL students (Short & Fitzsimmons, 2000). Communicating with students and families helps avoid misconceptions about language skills and academic knowledge and builds trust and understanding. Additionally, teachers should model respect for and an interest in cultural and language diversity. Teachers can inquire about and invite ELL students to bring their experiences into the classroom, for example, ask about trips back to their home countries and listen to ELL students talk proudly about what they did and saw on their trips back home. When the opposite occurs, that is, teachers or administrators criticizing ELL students and their families for leaving on trips to their homeland, ELL students feel alienated and unattached (Hall, Özerk, Zulfiqar & Tan, 2002).

6. *Assessing students' language status and development.* Informal assessment skills are particularly important for teachers to be able to recognize ELL students' expressive and receptive language proficiencies. To support ELL students, teachers need to be able to distinguish between "Basic Interpersonal Communication Skills" (BICS) and the ability to use language in "school subject matter learning" (Hakuta, Butler & Witt, 2000, p. 3). In conversational situations, ELL students can use BICS and appear to be proficient language users; however, learning through reading and writing in subject area courses requires different grammar, rhetoric, and vocabulary than is necessary for BICS. Teachers who can assess student strengths and needs in academic language skills will be better able to help ELLs engage in and with learning activities and materials. Additionally, teachers need to be able to uncover and help students build relevant background knowledge. Although English language learners may already possess background knowledge, it may be for other topics and/or hard to express in English (Short &Fitzsimmons, 2007).

In summary, teaching English language learners requires the coupling of effective principles of instruction with modifications reflecting the needs of non-native speakers (Gersten, Baker, Haager, & Graves, 2005).

Characteristics of an Effective ELL Program

Research studies have identified a number of program factors and instructional characteristics that promote the academic success of ELLs. Programs that are relatively effective had the following characteristics (Genesee et al. 2005, p. 377):

1. A positive school environment (Montecel & Cortez, 2002). School leadership teams can help create a school-wide environment that supports effective teaching and successful student outcomes (Movit, Petrykowska, & Woodruff, 2010).
2. A curriculum that was meaningful and academically challenging, incorporated higher order thinking (Doherty et al., 2003; Montecel & Cortez, 2002), was thematically integrated (Montecel & Cortez, 2002), established a clear alignment with standards and assessment (Doherty et al., 2003; Montecel & Cortez, 2002), and was consistent and sustained over time.
3. A program model that was grounded in sound theory and best practices associated with an enriched, not remedial, instructional model (Montecel & Cortez, 2002).
4. Teachers in bilingual programs who understood theories about bilingualism and second language development as well as the goals and rationale for the model in which they were teaching (Montecel & Cortez, 2002).
5. The use of cooperative learning and high-quality exchanges between teachers and students (e.g., Doherty et al., 2003; Montecel & Cortez, 2002).

Assessment of ELL Students' Progress Toward Mastering the English Language

ELL students need reliable and valid measures to determine their language progress, as the outcome determines ELL instruction, classification, progress reporting and promotion (Abedi, 2008a). In addition, ELL students' level of English proficiency is an important criterion in determining their readiness for participating in a state's academic assessments. The importance of adequately assessing ELL students' level of English proficiency is stated in the NCLB legislation, which required that schools receiving Title III funding assess ELL students using reliable and valid measures (NCLB, 2001). This section reviews measures of ELL students' English proficiency and their limitations, and discusses the use of assessment results as a criterion to redesignate ELL students.

Two types of assessments have been used to measure ELL students' language progress: standardized achievement tests and English Language Proficiency (ELP) tests. Kindler (2002) found that over 94% of state educational agencies used ELP tests results to classify and reclassify ELL students, whereas 76% of them (41 of 54) adopted achievement tests to measure ELL progress. More recent studies, however, have questioned the use of achievement test results for this purpose, saying that the content-based achievement tests were designed to assess only English-speaking students' content knowledge (Mahoney & MacSwan, 2005; Abedi, 2008a). As a result, few states are currently using standardized achievement tests to classify and reclassify ELLs.

Standardized Achievement Tests

Researchers have suggested that standardized achievement tests not be used to assess ELL students for classification and promotion purposes as the only criterion (Wolf et al., 2008). State content-based assessments that are used for No Child Left Behind Act (NCLB, 2001) Title I

accountability purposes are mainly constructed and field tested for English speaking students. Therefore, these tests may be subjective to linguistic factors that could seriously undermine their validity for ELLs. Abedi (2006) reviewed the current assessments for ELLs and found that unnecessary linguistic complexity of test items could be an additional source of measurement error in using standardized achievement tests for ELL students. Young (2009) found that construct-irrelevant variance was introduced when ELLs had difficulty in understanding the test item on a content assessment. He then summarized other researchers' work in which the use of language might not be fully accessible to ELLs in a content-based assessment (p. 124):

1. Unfamiliar vocabulary that is not related to the target construct (August, Carlo, Dressler, & Snow, 2005; Bailey et al., 2007; Martiniello, 2008),
2. Cultural references or idiomatic expressions that are not equally familiar to all students (Martiniello, 2008),
3. Syntax that may be confusing or ambiguous (Abedi, 2006; Martiniello, 2008),
4. Low-frequency, long, or morphologically complex words and long sentences (Abedi, 2006; Martiniello, 2008),
5. Sentence structure that may be confusing or difficult for students to follow (Abedi & Lord, 2001; Martiniello, 2008) and
6. Syntax or vocabulary that is above the test's target grade level (Borgioli, 2008).

In addition, language factors may interact with test items in achievement tests (Solano-Flores & Trumbull, 2003). This study found that linguistically complex items as a source of measurement error led students to misinterpret and misunderstand test questions.

Studies found that testing accommodations is a useful approach for mitigating the language effect on content-based achievement tests (Young et al., 2008; Abedi, Hofstetter & Lord, 2004). For ELLs, an accommodation refers to a change to testing, procedures, test materials, or the testing situation to allow students to participate in an assessment (Willner, Rivera, & Acosta, 2008). Although all states provide guidelines for testing accommodations for ELLs, researchers found that the guidelines often lack a sound research basis for recommending specific accommodations that are relevant to the educational needs of ELLs (Abedi & Gandara, 2006). Rivera, Collum, Willner, and Sia (2006) found that only 44 of the 75 testing accommodations recommended by states were relevant to the needs of ELLs.

Research findings on the effectiveness of testing accommodations for ELLs are mixed and inconclusive. Therefore, decisions about which accommodations to use, for whom, and under what circumstances are based on limited empirical evidence for their effectiveness and validity. Sireci, Li, and Scarpati (2003) conducted a meta-analysis of ELL testing accommodations and found that linguistic modification of test items is the most promising testing accommodation for ELLs. Abedi, Hofstetter and Lord (2004) did a comprehensive review of the current test accommodation strategies for ELLs and concluded that (p. 17):

1) The language assessment should match students' primary language of instruction. Therefore, translating test items from English into other languages is not an effective accommodation strategy when students have learned the subject in English.
2) The effectiveness of test accommodations is affected by students' background factors, such as English reading proficiency and the length of time in the United States.
3) Linguistic modification of test items is an effective accommodation strategy. The performance gap between ELLs and other students is narrowed by reducing the use of

      low-frequency vocabulary and complex language structures that are incidental to the
      content knowledge assessed.
4) Customized glossaries can be used as an effective and valid alternative to commercial
      dictionaries for all students, helping ELLs while not affecting the scores of English-
      proficiency students.

In addition to psychometric concerns with the use of achievement tests for ELL classification and reclassification, researchers could not reach consensus on the level of student performance above which a student is considered language proficient. For example, Grissom (2004) recommended that a cut-off score at the 36th percentile point or higher on a standardized norm-referenced test be used for a student to exit from ELL status, whereas a report by the U.S. General Accounting Office (2001) recommended that a student be considered English proficient with a score above the 50th percentile on a standardized achievement test. Another study from Linquanti (2001) reported that the ELL reclassification cutoff scores in seven California districts ranged from the 33rd percentile to the 40th percentile. There was never an agreement across these early studies on which cutoff scores imply an acceptable level of proficiency.

English Language Proficiency (ELP) Tests

NCLB requires that schools receiving Title I funding annually assess ELL students' level of English proficiency in reading, writing, speaking and listening. Additionally, these schools are required to use valid and reliable assessments that incorporate the concept of academic language, and align the content of ELP assessments with the states' ELP standards (NCLB, 2001).

Traditional language assessments may not be adequate to meet the mandates of NCLB. Del Vecchio and Guerrero (1995) reviewed five commonly-used ELP tests and questioned the construct definition and reliability of the classifications resulting from the tests. A wide range of disparities including test purposes, age and language group, administration, scoring, and theoretical foundations are reported in this handbook, which challenges the accuracy and consistency of these measures to classify ELLs.

The current focus of NCLB Title III ELP assessment is on "academic English" (Abedi, 2008a). Considering that ELP assessments may be used as one criterion in making a variety of academic decisions about ELL students, the constructs addressed by ELP assessments should reflect language ability required in an academic context. The document (U.S. Department of Education, Office of Elementary and Secondary Education, 2004) also indicated that ELP assessments should be aligned with the state's ELP standards.

More recently, language testing researchers have reported that many ELP tests fail in addressing the development of academic English language skills that ELL students need to succeed in school settings (e.g. Bailey & Butler, 2003; Garcia, McKoon, & August, 2006; Wolf et al. 2008). Stevens, Butler, and Castellon-Wellington (2000) examined the relationship between language proficiency and student performance on standardized content-area tests. They compared the type of language assessed in the Language Assessment Scales (LAS) and the language used in a standardized social study test for the seventh grade. The researchers found that the language assessment test used more general language, whereas the content test had more academic language characteristics. In other words, the LAS used less complex syntax, vocabulary that consisted generally of everyday words, and discourse less demanding to process. On the other hand, the content-area test had academically more demanding language including academic

vocabulary, various syntactic structures, and more specific linguistic registers. These findings pointed to the need for measures of language proficiency that were consistent with the language demands of standardized content assessments (Butler & Castellon-Wellington, 2000).

Further limitations of traditional language proficiency assessments were reported by Wolf et al. (2008, p. 22):

1. The construct of the assessment is concerned mainly with social, everyday language, and the results do not reflect whether the student is at the level of readiness or competency to perform in an academic setting (Butler, Stevens, & Castellon-Wellington, 2007);

2. A mismatch is likely between the language skills traditionally tested and the language demands that are expected in school (Stevens, Butler, & Castellon-Wellington, 2000);

3. There are both great variety and a lack of consensus in what areas of language ability are addressed and the types of tasks used in the assessments (Zehler, Hopstock, Fleishman, & Greniuk, 1994);

4. Existing assessments do not address all key language use activities (i.e., listening, reading, speaking, and writing);

5. The assessments are not systematically designed to measure progress in the attainment of English proficiency (Garcia, McKoon, & August, 2006).

While not tested by researchers, Abedi (2008b) provided a set of recommendations for selecting and using ELP assessments that were intended to positively impact instruction, assessment and state accountability systems for ELL students (pp. 210-211):

1. Use multiple methods for setting standards and defining cut scores for achievement levels;
2. Examine the comparability of the assessment used to establish the baseline with the newly adopted ELP assessment. If there is not a strong link between two assessments, both in terms of content and psychometric characteristics, then use caution in making firm judgments about the growth of students' English proficiency;
3. Examine the content of your state adopted ELP assessment and align the content with the state ELP content standards;
4. Examine the pattern of possible differential performance of ELL students on the ELP assessment to make sure that the ELP assessment items do not differentially or unfairly perform across the subgroups within the ELL population;
5. Use multiple criteria for assessing ELL students' level of English proficiency, particularly with high-stakes decision such as classification or reclassification of students;
6. Use ELP assessment results along with other sources to make informed decision about ELL student participation in Title I assessment, as the literature clearly suggests that assessments that are constructed for native speakers of English may not provide valid outcomes to ELL students at the lower level of language proficiency;
7. Train staff with high levels of knowledge and experience in measurement so that you can constantly review and monitor assessment issues, particularly in the area of English proficiency;
8. Incorporate a major measurement research component into your programs that can be supervised and run with your professionally trained staff;

Using Assessment Results to Redesignate ELL students

"Redesignation" is the process by which ELL students are assessed for continuation of special language support services (Wolf et al., 2008). Some states also use the term "exit of ELL status."A variety of criteria have been used by states to determine whether students are ready to be "redesignated". These criteria include state ELP test scores (e.g., scores from ELDA or CELLA), academic achievement test scores, school personnel evaluation, parent opinion and consultation, student grades, portfolio (work sample) assessments, interviews, and evaluations of classroom performance (Abedi, 2008a, Grissom, 2004, Wolf et al., 2008)

Within the framework of NCLB, ELP assessments have been used for identifying ELL students, determining levels of proficiency for instructional placement, redesignating ELL status, and tracking the progress of the development of language proficiency (Wolf et al., 2008). Almost all states (46 of 48 surveyed states) used ELP assessment scores as one of the criteria to determine if ELL students were ready for the mainstream classroom, yet a closer look at the use of multiple criteria in state policies highlights the complexity and variation in the redesignation process across states (Wolf et al., 2008). According to two CRESST reports (Wolf et al., 2008; Wolf, Herman, & Dietel, 2010), 34 states used between one and five different criteria in various combinations to make a decision about redesignation. Twelve states used ELP test scores as the single source to make a redesignation decision. Seven states redesignated ELLs based on test scores from both the ELP assessment and content tests. Thirty states required a single method or combination of several methods, while other states (18 of 48 states) provided local districts with guidelines for acceptable criteria and let districts to choose their own redesignation criteria (Wolf et al., 2008).

The Minnesota Department of Education (MDE) requires that a student be redesignated using district-established criteria based on measures including teacher recommendation, parent input, and ELP assessment outcomes in speaking, listening, reading and writing (MDE, 2006). Research has also suggested that ELP assessments results be used with caution or combined with other criteria for ELL redesignation. A study on the academic achievement of students who were recently redesignated English proficient yielded mixed results regarding the students' subsequent academic performance compared to English-speaking students (Abedi, Leon, & Mirocha, 2003). Grissom's study (2004) on ELL redesignation in California also found that the use of ELP assessments for ELL redesignation was problematic. In this study, the state of California employed multiple criteria collectively in its redesignation process for ELL students. The criteria included (a) an ELP assessment, (b) teacher evaluation, (c) parent opinion, and (d) academic achievement based on content-based standardized tests. Findings of this study suggested that the results of a language proficiency assessment as the only criterion were insufficient for redesignation decisions.

The CRESST researchers recommended that all states have the authority to set the criteria for redesignating ELL students and monitor their implementation in districts (Wolf et al., 2008; Wolf, Herman, & Dietel, 2010). The authors suggested that states should provide a clear guideline to instruct the district-level decision making process of ELL redesignation. If multiple criteria were used, policy guidelines should specify how each criterion should be employed at district level. Additionally, the CRESST reports argued that all states/districts examine their reclassification rates and evaluate the effects of their reclassification policies. The researchers recommended that states, districts, and schools monitor newly redesignated ELL students and maintain a tracking system to do so. Data from this practice could be used to evaluate potential

consequences for reclassified and long-term ELL students, hence providing evidence to support or modify the states' redesignation criteria.

An Evaluation to Determine the Effectiveness of the Elements of the District's LEP Plan

The No Child Left Behind Act (NCLB, 2001) requires that states review Local Educational Agency (LEA) programs that have been designed and implemented to increase the language acquisition and academic achievement of ELL students enrolled in their schools. To meet the evaluation requirement mandated by law, the district must conduct an annual evaluation of its ELL program that states the extent to which its goals have been met.

Despite the importance of evaluating the district's language instructional program for compliance with Title III requirements, there are few studies that have examined evaluation approach and methodologies for this purpose. One consequence of this lack of research on ELL program evaluation is that districts have to make important policy decisions to implement programs based on limited evidence. In addition, a lack of valid evaluation impedes a district's ability to make informed decision about program modification and continuous improvement (Llosa & Slayton, 2009).

The next section reviews ELL program evaluations conducted by various states and districts and discusses how program evaluation can be conducted in ways that meaningfully affect the achievement of ELL students.

Common Features in Current Evaluation Plans

Evaluation procedures of language assistance programs often involve multiple stakeholders to gather information through written reviews, surveys and/or interviews. For example, the Council of the Great City Schools (2009) conducted an evaluation that examined the district level policies and practices, and the programmatic contexts of school systems that demonstrated growth in ELL student achievement from 2002 to 2006. Data collection of this study consisted of compilation of materials from district staff and interviews and focus groups with key district leaders and ELL-related personnel. The 1999-2000 evaluation of the Austin Independent School District (AISD) Bilingual/ESL Programs (Gonzalez, 2001) suggested that the evaluation plan be reviewed and revised through an interactive process involving multiple stakeholders, including program director, instructional coordinator, and the evaluation staff. This approach was echoed in the program evaluation plan of Lake Oswego School District, Oregon. The district required that information on the ELL program implementation be obtained from ESL staff, classroom teachers and assistants that work with ELL students, and administrators and other personnel involved with duties related to ELLs (Lake Oswego School District, 2011).

Many evaluation plans include components that address the requirements of Title III. For example, The *Title III Program Evaluation* drafted by the Michigan Department of Education (2008) provided an evaluation checklist for local educators. Local educators were asked to provide evidence and documentation in response to questions that addressed the following six components: 1) administrative responsibilities, 2) student identification, 3) parent and community involvement, 4) instructional programs and assessment, 5) teacher qualifications and professional development, and 6) program evaluation and school improvement. *The LEP Program Evaluation Guidance* provided by the Idaho State Department of Education also

required that local schools or districts use a school improvement planning checklist to evaluate their LEP program. On this checklist, evaluation questions were organized under the following five components: 1) identification and assessment, 2) programming and educational approaches, 3) staffing and professional development, 4) parental involvement, and 5) program evaluation and review (Iribarren, 2009). Minnesota Department of Education (MDE) developed a review plan for English Learner (EL) program review and Title III Monitoring. Seven critical elements for EL programs were designed to guide local educational agencies (LEAs) in planning and evaluating their EL programs. These critical elements include: 1) identification, placement and program exit; 2) appropriate language instruction program; 3) appropriate staff and professional development; 4) parent involvement; 5) accountability requirements; 6) fiscal requirements; and 7) nonpublic school participation in language instruction education program (MDE, 2011a). MDE requires that LEAs submit EL program description and complete a self-assessment of the implementation of their EL programs. LEAs also need to provide evidence that their EL program is aligned with the critical elements. Finally, MDE makes onsite visits to LEAs and schools, meeting with LEA leadership, revising program documentations and interviewing stakeholders (MDE, 2010b).

In the evaluation plans reviewed, evaluation information was often obtained from multiple data sources. The decision making for LEP programs in schools/districts is therefore data driven. For example, the Idaho LEP Program Evaluation Guidance (2006) notes that test scores for each student should be only one way to evaluate the effectiveness of an ELL program. Evaluation data could also be collected from professional development, teaching training and instruction. Schools/districts needed to make sure that they collect data from all variables that inform instruction, such as student characteristics, LEP programs of service, students' LEP status (1st year, 2nd year monitoring), ELP assessment data, and academic achievement data. Gonzalez *(*2001) documented program effectiveness that was assessed through various data sources. A master file of LEP students was used to record students' LEP status, home language, language dominance, and program service dates. Former LEP students were observed on a yearly basis and their academic performance was also included in the current evaluation. A student master file provided basic information about students' grade level, ethnicity and socioeconomic status. Additional data sources included programmatic information and professional staff development details provided by the bilingual coordinators, prior-year information concerning LEP students, and Emergency Immigrant Program expenditures. Program effectiveness was assessed by comparing outcome indicators for LEP students being served and for LEP students whose parents refused program services (Gonzalez*,* 2001). Iribarren (2009) recommended that a comprehensive list of school-wide factors should be part of the evaluation plan. Factors to consider include "attendance rate, dropout rate, suspension rate, expulsion rates, participation rate in co-curricular/extracurricular activities, student attitude and interest survey results, longitudinal performance data, achievement data, special education placements and referrals and participation in Gifted and Talented programming" (p. 5).

Essential Elements for a Useful Program Evaluation

Llosa and Slayton (2009) conducted a large scale evaluation of an early-reading program in schools with a substantial population of ELLs in California. A quasi-experimental approach was used for the program evaluation design. This study found that four important features should be incorporated into the evaluation plan: 1) an investigation of the context, 2) the use of multiple

types and sources of data, 3) the use of appropriate analytic tools, and 4) the use of extensive qualitative data.

Llosa and Slayton (2009) found that it was important to design the evaluation in the context of other instructional activities at school. They argued that conducting an evaluation with a limited focus on the program variables alone would yield results of limited utility as they failed to consider the actual context where the LEP program was implemented. As Maxwell (2004) suggested, an adequate evaluation of educational phenomena should adopt methods that "can investigate the involvement of particular context in the processes that generate these phenomena and outcomes" (p. 7). It is therefore critical to collect contextual information such as school characteristics, student demographics, and teacher training and professional development. The REL Central North Central Comprehensive Center (NCCC) co-developed a "big picture" ESL audit with the South Dakota Department of Education, Huron School District and University of South Dakota. Instead of limiting it to an isolated program review, REL researchers took a systems approach and called it an ESL Systems Appraisal. In this approach, data are collected to address multiple contextual components, including leadership, ESL program design and delivery, general education curriculum and instruction, professional development and qualified staff, assessment and accountability, fiscal requirement and management, and parents and community (NCCC, 2011).

In order to fully understand the ESL program and the context in which it is implemented, multiple types and sources of data should be considered in the evaluation plan (Llosa & Slayton, 2009). In their study, Llosa and Slayton employed a muti-method approach to include both quantitative (e.g. achievement data) and qualitative data (e.g. classroom observation and teacher interview data) to get a better understanding of factors that might affect the effectiveness of the program.

Finally, Llosa and Slayton (2009) argued that appropriate analytic tools should be selected to analyze multiple data sources. In the current evaluation, they used Hierarchical Linear Modeling (HLM) to examine the differences between the treatment and control groups. This analysis tool also allowed them to look at the relationship between the implementation of the early-reading program and student achievement within the treatment group. It is therefore essential to choose the right analytic tools that would maximize the information gathered from an evaluation and directly address the evaluation questions. If multiple sources of data are included in an evaluation, multiple tools should be considered to adequately analyze both quantitative and qualitative data.

References

Abedi, J. (2004). The No Child Left Behind Act and English language learners: Assessment and accountability issues. *Educational Researcher*, *33*(1), 4-14.

Abedi, J. (2006). Language issues in item development. In S. M. Downing & T. M. Haladyna (Eds.), *Handbook of test development* (pp. 377-398). Mahwah, NJ: Erlbaum.

Abedi, J. (2007). *English Language Proficiency Assessment in the Nation: Current Status and Future Practice* (Ed.). Davis: University of California.

Abedi, J. (2008a). Classification System for English Language Learners: Issues and Recommendations. *Educational Measurement: Issues & Practice, 27*(3), 17-31.

Abedi, J. (2008b). Measuring Students' Level of English Proficiency: Educational Significance and Assessment Requirements. *Educational Assessment, 13*(2/3), 193-214.

Abedi, J., & Gandara, P. (2006). Performance of English language learners as a subgroup in large-scale assessment: Interaction of research and policy. *Educational Measurement: Issues and Practice, 25*(4), 36-46.

Abedi, J. & Herman, J. (2010). Assessing English language learners' opportunity to learn mathematics: Issues and limitations. *Teachers College Record, 112*(3), 723-746.

Abedi, J., Hofstetter, C., & Lord, C. (2004) Assessment accommodations for English language learners: Implications for policy based research. *Review of Educational Research, 74*(1), 1-28.

Abedi, J., Leon, S., & Mirocha, J. (2003). *Impact of student language background on content-based performance: Analyses of extant data* (CRESST Tech. Rep. No. 603). Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

Abedi, J. & Lord, C. (2001). The language factor in mathematics tests. *Applied Measurement in Education*, *14*, 219-234.

Abedi, J., Lord, C., Hofstetter, C., & Baker, E. (2000). Impact of accommodation strategies on English language learners' test performance. *Educational Measurement: Issues and Practice*, *19*(3), 16-26.

Abedi, J., Lord, C., & Plummer, J. (1997). *Language background as a variable in NAEP mathematics performance* (CSE Tech. Rep. No. 429). Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing.

Adams, C. M., & Pierce, R. L. (2003). Teaching by tiering. *Science and children, 41*(3), 30-34.

Aguirre-Munoz, Z. & Amabisca, A. A. (2010). Defining opportunity to learn for English language learners: Linguistic and cultural dimensions of ELLs' instructional contexts. *Journal of Education for Students Place at Risk, 15,* 259-278.

Antunez, B. (2002). The preparation and professional development of teachers of English language learners. *ERIC Digest*. Washington, DC: ERIC Clearinghouse on Languages and Linguistics. (ERIC Document Reproduction Service No. ED 477724).

August, D., Carlo, M., Dressler, C., & Snow, C. (2005). The critical role of vocabulary development for English language learners. *Learning Disability Research and Practice*, *20*(1) 50-57

August, D., & Shanahan, t. (Eds.). (2006). *Developing Literacy in Second-Language Learners: Report of the National Literacy Panel on Language –Minority Children and Youth* (Executive Summary). Mahwah, NJ: Lawrence Erlbaum Associates, Publishers. Retrieved from http://www.cal.org/projects/archive/nlpreports/executive_summary.pdf

Bailey, A. L., & Butler, F. A. (2003*). An evidentiary framework for operationalizing academic language for broad application to K-12 education: A design document* (CRESST Tech. Rep. No. 611). Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

Bailey, A. L., & Butler, F. A. (2004). Ethical considerations in the assessment of the language and content knowledge of English language learners K-12. *Language Assessment Quarterly, 1,* 177-193.

Bailey, A. L., Huang, B., H., Shin, H. W., Farnsworth, T., & Butler, F. A. (2007). *Developing academic English language proficiency prototypes for 5th-grade reading: Psychometric and linguistic profiles of tasks* (CSE Tech. Rep. No. 727). Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing.

Bauer, E., & Manyak, P.C. (2008). Creating Language-Rich Instruction for English-Language Learners. *The Reading Teacher*, *62*(2), 176-178.

Baumann, J. F., & Graves, M. F. (2010). What is academic vocabulary? *Journal of Adolescent & Adult Literacy, 54*, 4-12.

Borgioli, G. (2008). Equity for English language learners in mathematics classrooms. *Teaching Children Mathematics, 15*, 185-191.

Brooks, K., & Thurston, L. (2010). English language learner academic engagement and instructional grouping configurations. *American Secondary Education, 39*(1), 45-60.

Brown, C. L. (2007). Strategies for making social studies texts more comprehensible for English language learners. *The Social Studies*, 98(5), 185-188.

Butler, F. A., & Castellon-Wellington, M. (2000). *Students' concurrent performance on tests of English language proficiency and academic achievement.* Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

Butler, F. A., Stevens, R., & Castellon, M. (2007). ELLs and standardized assessments: The interaction between language proficiency and performance on standardized tests. In A. L. Bailey (Ed.), *The language demands of school: Putting academic English to the test* (pp. 27-49). New Haven, CT: Yale University Press.

Campbell, P. & Rowan. T. (1997). *Teacher questions + student language + diversity = mathematical power. Multicultural and Gender Equity in the Mathematics Classroom: The Gift of Diversity*. Reston, VA: National Council of Teachers of Mathematics.

Carlo, M.S., August, D., McLaughlin, B., Snow, C., Dressler, C., Lippman, D., et al. (2004). Closing the gap: Addressing the vocabulary needs of English-language learners in bilingual and mainstream classrooms. *Reading Research Quarterly*, *39*, 188-215.

Chamot, A. U. & O'Malley, J. M. (1999). The Cognitive Academic Language Learning Approach: A model for linguistically diverse classrooms. In Biacindo, K. (Ed.), *Perspectives: Educational Psychology*, pp. 39-51. Boulder, CO: Coursewise Publishing.

Chang, M., K. Singh, K. Filer, K. (2009). Language factors associated with achievement grouping in math classrooms: a cross-sectional and longitudinal study. *School Effectiveness & School Improvement*, *20*(1), 27-45.

Cirillo, M., Bruna K. R., Herbel-Eisenmann, B. (2010). Acquisition of mathematical language: suggestions and activities for English Language Learners. *Multicultural Perspectives* 12(1): 34-41.

Clair, N., & Adger, C.T. (1999). Professional development for teachers in culturally diverse schools. *ERIC Digest*. Washington, DC: ERIC Clearinghouse on Languages and Linguistics. (ERIC Document Reproduction Service No. ED435185)

Collier, V.P., & Thomas, W.P. (2004). The astounding effectiveness of dual language education for all. *NABE Journal of Research and Practice, 2*(1), 1-20.

Cosentino de Cohen, C., Deterding, N. & Clewell, B.C. (2005). *Who's left behind: Immigrant children in high and low LEP schools*. Washington, DC: Urban Institute.

De la Cruz, Y. (1998). Issues in the teaching of math and science to Latinos. In Gonzalez, M.L., Huerta-Macias, A. & Tinajero J.V. (Eds.), *Educating Latino students: A guide to successful practice* (pp. 161-176). Lancaster, PA: Technomic Publishing Company, Inc.

Del Vecchio, A., & Guerrero, M. (1995). *Handbook of English language proficiency tests.* Albuquerque, NM: EAC West, New Mexico Highlands University

Doherty, R. W., Hilberg, R. S., Pinal, A., & Tharp, R. G. (2003). Five standards and student achievement. *NABE Journal of Research and Practice, 1*(1), 1-24.

Echevarria, J., Vogt, M.E. & Short, D. (2004). *Making content comprehensible for English language learners: The SIOP model, second edition*. Boston: Allyn & Bacon.

Edmonds, L. M. (2009). Challenges and solutions for ELLs. *Science Teacher, 76*(3), 30-33.

Fillmore, L. W. & Snow, C. (2000). *What teachers need to know about language.* ERIC Clearinghouse on Language and Linguistics Special Report. (ERIC Document Reproduction Service No. ED444379).

Garcia, G., McKoon, G., & August, D. (2006). Language and literacy assessment of language-minority students. In D. August & T. Shanahan (Eds.), *Developing literacy in second-language learners* (pp. 597-624). Mahwah, NJ: Lawrence Erlbaum Associates.

Genesee, F. (1999). *Program alternatives for linguistically diverse students*. (Educational Practice Report No. 1.) Washington, DC: Center for Research on Education, Diversity, and Excellence.

Genesee, F., Lindholm-Leary, K., Saunders, W., & Christian, D. (2005). English language learners in U.S. schools: An overview of research findings. *Journal of Education for Student Placed at Risk, 10*(4), 363-385.

Gersten, R., Baker, S. K., Haager, D., & Graves, A. W. (2005). Exploring the role of teacher quality in predicting reading outcomes for first-grade English learners. *Remedial and special education*, *26*(4), 197-206.

Gersten, R., Baker, S.K., Shanahan, T., Linan-Thompson, S., Collins, P., & Scarcella, R. (2007). *Effective literacy and English language instruction for English learners in the elementary grades: A practice guide* (NCEE 2007-4011). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S.

Department of Education. Retrieved from http://ies.ed.gov/ncee/wwc/publications/practiceguides.

Golash-Boza, T. (2005). Assessing the advantages of bilingualism for the children of immigrants. *International Migration Review,* 39(3), 721-753.

Goldenberg, C. (2008). Teaching English language learners: What the research does—and does not—say. *American Educator, 33*(2), 8–44.

Goldenberg, C. & Rutherford-Quach, S. (2010). *The Arizona Home Language Survey: The identification of students for ELL services.* Los Angeles: Civil Rights Project / Proyecto Derechos Civiles.

Gonzalez, R. M. (2001). *Bilingual/ESL programs evaluation report, 1999-2000.* Austin: Independent School District. Office of Program Evaluation.

Greathouse, D. & Lincoln F. (2008). Using all available tools. *Science Teacher, 75*(5), 48-52.

Grissom, J. B. (2004). Reclassification of English language learners. *Education Policy Analysis Archives*, *12*(36).

Guidance for Developing the Limited English Proficiency (LEP) Plan. (2006). Idaho State Department of Education. Retrieved Jan. 23rd, 2011. http://www.sde.idaho.gov/LEP/docs/Curriculum/LEPPlanGuidance.pdf

Guidelines for the Identification and Assessment of Limited English Proficient Students. (2010) Ohio Department of Education Learning Supports. Retrieved Jan. 23rd, 2011. http://www.ode.state.oh.us/GD/Templates/Pages/ODE/ODEDetail.aspx?page=3&TopicRelationID=499&ContentID=16099&Content=90398

Hakuta, K., Butler, Y. G., & Witt, D. (2000). *How long does it take English learners to attain proficiency?* Santa Barbara, CA: University of California Linguistic Minority Research Institute. Retrieved from http://escholarship.org/uc/item/13w7m06g

Hall, K. A., Özerk, K, Zulfiqar, M., & Tan, J. E., C. (2002). 'This is our school': Provision, purpose and pedagogy of supplementary schooling in Leeds and Oslo. *British Educational Research Journal, 28*(3), 399-418.

Han, W-J. & Bridglall, B. L. (2009). Assessing school supports for ELL students using ECLS-K. *Early Childhood Research Quarterly*, *24*, 445-462.

Himmel, J., Short, D. J., Richards, C., & Echevarria, J. (2009). Using the SIOP model to improve middle school science instruction. *ICreate Brief*. Retrieved from http://www.cal.org/create/resources/pubs/CREATEBrief_SIOPScience.pdf

Honnert, A. M. & Bozan, S. E. (2005). Summary frames: Language acquisition for special education and ELL students. *Science Activities, 42*(2), 19-29.

Iribarren, J. (2009). *Reasons to evaluate a language assistance program: Are we making progress?* (Informational Bulletin No. 3.) Wisconsin Department of Public Instruction. Retrieved on Jan 23rd, 2011 http://dpi.wisconsin.gov/ell/doc/program-eval.doc.

Lara, J., Ferrara, S., Calliope, M., Sewell, D., Winter, P., Kopriva, R., Bunch,M., & Joldersma, K. (2007) . The English language development assessment (ELDA). In J. Abedi (Ed.), *English language proficiency assessment in the nation: Current status and future practice* (pp. 47-60). Davis: University of California.

Lucas, T. & Grinberg, J. (2008). Responding to the linguistic reality of mainstream classrooms: Preparing all teachers to teach English language learners. In M. Cochran-Smith, S.

Feiman-Nemser, and J. McIntyre (Eds.). *Handbook of research on teacher education: Enduring issues in changing contexts* (3rd edition, pp. 606-636). Mahwah, NJ: Lawrence Erlbaum.

Kindler, A. J. (2002). *Survey of the states' Limited English Proficient students and available educational programs and services: 1999-2000 Summary Report*. Washington, DC: National Clearinghouse for English Language Acquisition and Language Instruction Educational Programs.

Kucer, S.B., & Silva, C. (1999). The English literacy development of bilingual students within a transition whole-language curriculum. *Bilingual Research Journal*, *23*(4), 345–371.

LEP Program Evaluation – School/District Self Evaluation Guidance. Idaho State Department of Education. Retrieved on Jan 23rd, 2011 from http://docs.google.com/viewer?a=v&q=cache:LlDvya4lEMoJ:www.sde.idaho.gov/LEP/docs/Curriculum/LEPProgramEvaluation_001.doc+The+Idaho+LEP+Program+Evaluation+Guidance&hl=en&gl=us&pid=bl&srcid=ADGEESjdbmBCvG06Cl4yn3S1z4Wtwk5xJ4q9qrrKHjFc_E9OT5qBF52B73zAFnBDvyZxo19fK3IJINGChwag_HvTb3udF6QRUZUe7iRk-wIL5dpD68x-xhHulQvqc0p5XDx9BDXNn53&sig=AHIEtbT4lg4SkL836kgJdT7v98zImPdoCg

Littlejohn, J. (1998). *Federal control out of control: The Office for Civil Rights hidden policies on bilingual education*. (CEO Policy Brief.) Washington, DC: Equal Opportunity Foundation.

Llosa, L. & Slayton, J. (2009). Using program evaluation to inform and improve the education of young English language learners in US schools. *Language Teaching Research, 13*(1), 35-54.

Linquanti, R. (2001). *The redesignation dilemma: Challenges and choices in fostering meaningful accountability for English learners*. (Policy Report 2001-1). Santa Barbara: University of California, Linguistic Minority Research Institute.

Mahoney, K. S. & MacSwan, J. (2005). Reexamining Identification and reclassification of English language learners: A critical discussion of select state practices. *Bilingual Research Journal, 29*(1), 31-42

Martiniello, M. (2008). Language and the performance of English-language learners in math word problems. *Harvard Educational Review*, *78*, 333-368.

Maxwell, J. A. (2004). Causal explanation, qualitative research, and scientific inquiry in education. *Educational Researcher, 33*(2), 3-11.

McLaughlin, B., August, D., Snow, C., Carlo, M., Dressler, C., White, C., Lively, T., & Lippman, D. (2000). Closing the gap: Addressing the vocabulary needs of English Language Learners in bilingual and mainstream classrooms. *Reading Research Quarterly*, *39*(2), 188-215.

Menken, K., & Antunez, B. (2001). *An overview of the preparation and certification of teachers working with limited English proficient students*. Washington, DC: National Clearinghouse of Bilingual Education.

Menken, K., & Look, K. (2000). Meeting the needs of linguistically and culturally diverse students. *Schools in the Middle, 9*(6), 20-25

Mikow-Porto, V., Humphries, S., Egelson, P., O'Connell, D., & Teague, J. (2004). *English Language Learners in the southeast: Research, policy, & practice.* (Policy Report). Greensboro, NC: SouthEastern Regional Vision for Education.

Minnesota Department of Education [MDE] (2006). *ELL education program guidelines: Identifications and program basics*. Retrieved from http://education.state.mn.us/mdeprod/groups/EnglishLang/documents/Manual/002115.pdf

Minnesota Department of Education [MDE] (2011a). *Critical elements for English Learner (EL) programs.* Retrieved from http://education.state.mn.us/mdeprod/groups/EnglishLang/documents/Publication/007285.pdf

Minnesota Department of Education [MDE] (2011b). *English Learner (EL) program review and Title III monitoring protocol.* Unpublished manuscript.

Montecel, M.R., & Cortez, J.D. (2002). Successful bilingual education programs: Development and the dissemination of criteria to identify promising and exemplary practices in bilingual education at the national level. *Bilingual Research Journal, 26*.

Movit, M., Petrykowska, I., & Woodruff, D. (2010, May). *Using school leadership teams to meet the Needs of English Language Learners.* (Information Brief). Washington, DC: National Center on Response to Intervention.

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

North Central Comprehensive Center (2011). *ESL Systems Appraisal*. Unpublished manuscript.

Lake Oswego School District (2011). *Program evaluation, review and improvement*. Retrieved from http://org.loswego.k12.or.us/ELL/Evaluation.htm

Portes, A. & Hao, L. (2002). The price of uniformity: Language, family and personality adjustment in immigrants second generation. *Ethnic and Racial Studies*, *25* (6), 889-912.

Portes, A. & Rumbaut, R. G. (2001). *Legacies: The story of the immigrant second generation*. Berkeley, CA: University of California Press and Russell Sage Foundation.

Program Evaluation, Review and Improvement. Lake Oswego School District, Oregon. Retrieved Jan 22nd, 2011 from http://org.loswego.k12.or.us/ELL/Evaluation.htm

Rivera, C., Collum, E., Willner, L. S., & Sia, J. K., Jr. (2006). An analysis of state assessment policies regarding the accommodation of English language learners. In C. Rivera & E. Collum (Eds.), *State assessment policy and practice for English language learners: A national perspective.* Mahwah, NJ: Lawrence Erlbaum Associates.

Rivera, C., Stansfield, C. W., Scialdone, L., & Sharkey, M. (2000). *An analysis of state policies for the inclusion and accommodation of English language learners in state assessment programs during 1998–1999*. Arlington, VA: The George Washington University, Center for Equity and Excellence in Education.

Ron, P. (1999). Spanish-English language issues in the mathematics classroom. In W. Secada, L. Ortiz-Franco, N. G. Hernandez, & Y. de la Cruz (Eds.), *Changing the faces of mathematics: Perspectives on Latinos* (pp. 22-33). Reston, VA: National Council of Teachers of Mathematics.

Rolstad, K., Mahoney, K. & Glass, G. V. (2005). The big picture: A meta-ana1ysis of program effectiveness research on English language learners. *Educational Policy, 19*(4), 572-594.

Rumbaut, R. (1995). The new Californians: Comparative research findings on the educational progress of immigrant children. In R. Rumbaut and W. Cornelius (Eds.), *California's immigrant children* (pp. 17-69). San Diego: Center for U.S.-Mexican Studies.

Saunders, W., & Goldenberg, C. (1999). *The effects of instructional conversations and literature logs on the story comprehension and thematic understanding of English proficient and limited English proficient students.* Center for Research on Education, Diversity & Excellence Research Reports.

Schoonover, R. (2010). Consultation request for McREL.

Sireci, S.G., Li, S. & Scarpati, S. (2003). *The effects of test accommodation on test performance: A review of the literature.* (Centre for Educational Assessment Research Report No. 485). Amherst, MA: School of Education, University of Massachusetts.

Short, D., & Fitzsimmons, S. (2007). *Double the work: Challenges and solutions to acquiring language and academic literacy for adolescent English language learners – A report to Carnegie Corporation of New York*. Washington, DC: Alliance for Excellent Education.

Slavin, R.E. & Cheung, A. (2003). *Effective reading programs for English language learners: A best-evidence synthesis*. Baltimore, MD: Johns Hopkins University, Center for Research on the Education of Students Placed at Risk.

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

Stevens, R. A., Butler, F. A., & Castellon-Wellington, M. (2000). *Academic language and content assessment: Measuring the progress of ELLs* (CRESST Tech. Rep. No. 552). Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

Sugarman, J., & Howard, E.R. (2001). *Development and maintenance of two-way immersion programs: Advice from practitioners* (Practitioner Brief #2). Santa Cruz, CA and Washington, DC: Center for Research on Education, Diversity & Excellence.

Szpara, M. Y. & Ahmad, I. (2007). Supporting English-language learners in social studies class: Results from a study of high school teachers. *Social Studies, 98*(5), 189-195.

Thomas, W.P., & Collier, V. (1997). *School effectiveness for language minority students.* Washington, DC: NCBE. Retrieved from http://www.ncbe.gwu.edu/ncbepubs/resource/effectiveness/index.htm.

Thomas, W.P., & Collier, V.P. (2002). *A national study of school effectiveness for language minority students' long-term academic achievement.* Santa Cruz, CA: Center for Research on Education, Diversity and Excellence, University of California-Santa Cruz.

Title III Program Evaluation. (2008). Michigan Department of Education. Retrieved on Jan 23rd, 2011. http://www.misd.net/bilingual/MITitleIIIProgramEvaluation.pdf

U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics [NCES]. (2007). *The national assessment of educational progress*. Washington, DC: Author.

U.S. Department of Education, Office of Elementary and Secondary Education. (2004). *Standards and assessments peer review guidance: Information and examples for meeting requirements of the No Child Left Behind Act of 2001*. Washington, DC: Author.

United States General Accounting Office. (2001). *Meeting the needs of students with limited English proficiency* (GAO-01-226). Washington, DC: Author.

Valdez, G., Svedkauskaite, A., & McNabb, M. (2002). *Critical Issue: Mastering the mosaic-framing impact factors to aid limited-English-proficient students in mathematics and science*. Naperville, IL: North Central Regional Educational Laboratory.

What Works Clearinghouse (2006). *Instructional conversations and literature logs* (What Works Clearinghouse intervention report). Rockville, MD: Author.

White, E. W., Makkonen, R. and Stewart, K. B. (2010). *Updated multistate review of professional teaching standards* (REL Technical Brief, REL 2010—No. 014). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory West. Retrieved from http://ies.ed.gov/ncee/edlabs.

Williams, T., Hakuta, J., Haertel, E., Kirst, M.,Perry, M., Oregón, M., Brazil, N., & Levin, J. (2007). *Similar English learner students, different results: Why do some schools do better? A follow-up analysis, based on a large-scale survey of California elementary schools serving low-income and EL students.* Mountain View, CA: EdSource.

Willner, S., Rivera, L., Acosta, B. (2008). *Descriptive analysis of state 2006-2007 content area accommodations policies for English language learners (2008).* (Prepared for the LEP Partnership, U.S. Department of Education). Arlington, VA: The George Washington University Center for Equity and Excellence in Education. Retrieved from http://ells.ceee.gwu.edu.

Wolf, M. K., Kao, J., Herman, J. L., Bachman, L. F., Bailey, A. L., Bachman, P. L., Farnsworth, T. & Chang, S. M. (2008). *Issues in assessing English language learners: English language proficiency measures and accommodation uses—Literature review* (CRESST Tech. Rep. No. 731). Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

Wolf, M. K., Herman, J. L. & Dietel, R. (2010). *Improving the validity of English language learner* Assessment Systems (CRESST Policy Brief No. 10). Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

Young, J. W. (2009). A framework for test validity research on content assessments taken by English language learners. *Educational Assessment, 14*(3/4), 122-138

Young, J. W., Cho, Y., Ling, G., Cline, F., Steinberg, J., & Stone, E. (2008). Validity and fairness of state standards-based assessments for English language learners. *Educational Assessment, 13*, 170-192.

Zehler, A., Fleischman, H., Hopstock, P. Stephenson, T., Pendzick, M., & Sapru, S. (2003). *Descriptive study of services to LEP students and LEP students with disabilities* (Vol I). Washington, DC: Office of English Language Acquisition, Language Enhancement and Academic Achievement of Limited English Proficient Students (OELA), U.S. Department of Education.

Zehler, A., Hopstock, P., Fleishman, H., & Greniuk, C. (1994). *An examination of assessment of limited English proficient students*. Arlington, VA: Special Issues Analysis Center.

Zehr, M. (2010). Boston settles with rederal officials in ELL investigation. *Education Week, 30*(7), 10. Retrieved from Professional Development Collection database.

Zelasko, N., Antunez, B., (2000). *If your child learns in two languages: A parent's guide for improving educational opportunities for children acquiring English as a second language.* Washington, DC: National Clearinghouse for Bilingual Education, The George Washington University Center for the Study of Languages and Education.

Appendix A: Idaho LEP Program Evaluation Guidance

LEP Program Evaluation – School/District Self Evaluation Guidance

"A school that is willing to examine itself critically is one that will increase the odds that its students will succeed." - Hoachlander and Mandel (1998)

"Evaluation conducted in the service of learning can inform and improve the effectiveness of our schools. In short, it can add value to our work, not simply yield status reports." – Kathleen Fitzpatrick, Program Evaluation: ESL (2002)

"A successful language development program is focused on good instructional practices implemented consistently within all content classrooms and additional, intensified instruction for newcomers." – Molly Jo de Fuentealba, Boise Independent School District (2006)

1. Why evaluate the LEP Program?
2. Resources for LEP Program Evaluation
3. Using Data to drive decision making for LEP Programs
4. Using a School Improvement Planning Checklist for LEP Programs

1. Why evaluate the LEP program?

Office of Civil Rights (OCR)
In determining whether a recipient's program for LEP students complies with Title VI of the Civil Rights Act of 1964, OCR has used the standard set forth in Castaneda v. Pickard, 648 F. 2d 989 (5th Cir. 1981). Under this standard, a program for LEP students is acceptable if:
  (1) The school system is pursuing a program informed by an educational theory recognized as sound by some experts in the field or, at least, deemed a legitimate experimental strategy [A Sound Approach];
  (2) The programs and practices actually used by [the] school system are reasonably calculated to implement effectively the educational theory adopted by the school [Reasonable Implementation];
  (3) The school's program succeeds, after a legitimate trial, in producing results indicating that the language barriers confronting students are actually being overcome [Evaluate for Outcomes]. Id. at 1009-10.

No Child Left Behind Act (2001)
Sec. 3122. Achievement Objectives and Accountability
Districts and schools are measured according to the percent and number of students (1) making progress in English language proficiency, (2) attaining English language proficiency and (3) meeting AYP.

State of Idaho
House Bill No. 787 states that each District must "formulate a plan in sufficient detail that measurable objectives can be identified and addressed which will accomplish English language acquisition and improved academic performance. Moneys distributed to school districts based upon the population of limited-English proficiency students and distributed to school districts to support programs for students with non-English or limited-English proficiency shall be utilized in support of the district plan."

2. Resources for LEP Program Evaluation

1. **Annual LEP Plan Guidance Document** – District LEP Plans must follow the format in this document. The questions will help districts walk through what they need to consider in order to implement a successful language development program.
2. **OCR Resource Materials** - For Planning and Self-Assessments of Programs for English Language Learners**.**
3. **OCR Direct Assessment Guide** – A survey regarding equal educational opportunities for English language learners.
4. **Claiming Opportunities Handbook** – A Handbook for Improving Education for English Language Learners Through Comprehensive School Reform (The Education Alliance at Brown University).
5. **Program Evaluation: English as a Second Language, NSSE** – National Study of School Evaluation. **Based on the national standards for student learning, this guide provides a comprehensive framework for program evaluation and improvement planning.**

3. Using Data to drive decision making for LEP Programs

Data and test scores for each student should be only one way to evaluate if a program is successful. Professional development, teacher training and solid instruction all help to determine or drive the "data". Therefore, there must be an accountability plan in place for all teachers to begin to implement good teaching strategies aimed at reaching the English language learners in their classroom. But when a school/district does use the data, make sure to incorporate all factors/data that will help inform instruction. For example, many 3rd generation Hispanic/LEP students in a particular school are struggling…why? What will your school do to address this particular issue?

Below are some of the information that schools should consider when disaggregating data:
- Race/ethnicity
- Gender
- Migrant or refugee status
- Free or reduced lunch
- Individualized Education Plan (IEP) for special education
- Gifted
- LEP Program of service
- 1st year in a U.S. school (LEP1)
- LEP on 2 year monitoring status (LEPX)
- English language proficiency assessment data (IELA)

- o   Academic achievement data
    - Classroom grades
    - ISAT, IRI, DWA, DMA

4. Using a School Improvement Planning Checklist for LEP Programs

School/District_____ Date _____

Addressing the needs of LEP students is a challenge. Because LEP students are held accountable for language acquisition **and** knowledge of academic content area, how do we help students meet the language and academic requirements at the same time? As each school begins to assess its programming for LEP students, it is important to remember 3 main points:

- It takes the whole school, including administrators to address the LEP students. *LEP students are within the school and school system, not just in a "program".*

- It takes teachers willing and able to give differentiated instruction to meet the needs of LEP students. *LEP students are required to have content area instruction and are in regular classrooms in most districts.*

- All students will succeed if kept to high expectations. *LEP students are not necessarily limited in education.*

Please use the following questions below to evaluate your programs and services for LEP students. It is essential that each school keep accurate and up-to-date data for each LEP student. LEP student growth should be measured through language proficiency testing, as well as academic content testing and classroom grades.

Initial Questions to consider to assist LEP students succeed	Yes	No
IDENTIFICATION and ASSESSMENT		
Is the school aware of the District's LEP Plan? Has the school been in contact with the LEP District contact person? Is the LEP Plan being implemented within the school?		
Is the school administration aware of the legal requirements pertaining to identifying and placing LEP students? Does every student have a Home Language Survey (HLS) on file?		
Are all LEP students being assessed using the IELA: (1) IELA/Placement for new students and (2) IELA/spring to determine growth?		
Is the school administration aware of the requirement to address the individual needs of each LEP student?		

Do LEP students have an Educational Learning Plan (ELP), inclusive of language goals and benchmarks, if they are given accommodations or adaptations within the classroom and subsequently on assessments?		
Is the school using ISAT and language proficiency assessment data to guide instruction for individual students?		
PROGRAMMING and EDUCATIONAL APPROACHES		
Is the program addressing the needs of LEP students within each content area (i.e. Math, Science, as well as language acquisition)?		
Is the entire faculty and administration aware of the LEP students and their needs within the school?		
Are all teachers utilizing the English Language Proficiency (ELP) Standards as a tool and entry point in teaching LEP students?		
Are the LEP students placed in pull out ESL classrooms? If so, are they learning content area vocabulary and skills?		
If LEP students are taught within the classroom, are they receiving additional assistance with language instruction? Are LEP students being *served* within the regular classroom?		
Are LEP students learning content knowledge and skills, as well as making progress in learning the English language?		
Is the curriculum for LEP students challenging and academic based?		
Does the district have a content-based LEP Plan in place? How will the LEP students learn through their content classes?		
Is your school/district providing before/after or summer school programs for LEP students?		
Do the district and school make it a priority to allocate district funding to serve LEP students?		
STAFFING and PROFESSIONAL DEVELOPMENT		
Do regular classroom teachers have the resources, skills and knowledge to address the needs of LEP students in their classroom?		
Are content teachers trained in specific methodologies to target LEP students?		
If ESL teachers are teaching content area, do they have certification in the specific content area, as well as their ESL certification?		
Are middle and high school LEP students receiving specific attention in each class?		
Is there an accountability plan in place for all teachers to take ownership of the LEP students in their classroom and serve them with good instructional practices?		
Is the administration of the school encouraging of all teachers to implement good teaching practices to help LEP students?		

Does the school have mainly paraprofessionals serving the LEP students?		
Is the main service for LEP students just translating if possible? Is there academic learning in the translation?		
PARENTAL INVOLVEMENT		
Are parents of LEP students given notifications in their home language?		
Are parents of LEP students included in decisions within the school?		
Are parents informed and given education regarding the school system and helping their children at home?		
PROGRAM EVALUATION and REVIEW		
Is the school keeping complete data for each student in order to calculate growth in language proficiency from year to year?		
Is the school evaluating the programs and services annually?		
Who is responsible for monitoring services, determining if they are effective and making changes, if necessary?		

Appendix B: What State/Districts Are Doing

Program Evaluation: State Level

STATE	
CO	Colorado's *Guidebook on Designing, Delivering, and Evaluating Services for English Learners (ELs) (Revised January 2011),* provides school leaders with information on Evaluating and Managing Programs for ELs (p. 58) and school districts with guidance in designing their evaluation rather than mandates. The guide states that "To ensure a sound evaluation, the relationship between needs assessment, program or services design, program implementation, and evaluation should be clear" and suggests that school and district leaders answer four questions when deciding on the value of "procedure related to planning and implementing services for ELs", including: • Was an adequate needs assessment conducted? • Were the goals and objectives adequately formulated and appropriated to the student needs**?** • Was the design and delivery of services, procedures, practices, and programs adequately described and consistent with the goals and objectives? • Were the evaluation questions adequately defined and in keeping with the goals and objectives? For more information, see http://www.cde.state.co.us/cde_english/download/ELLGuidebook/Final_1-13-2011_Guidebook%202011.pdf
ID	The *Idaho LEP Program Guidance* (May 16, 2008) provides district and school leaders with: • Resources for LEP program evaluation • Information regarding using data to drive decision making for LEP program • A school improvement planning checklist for LEP programs. Idaho House Bill No. 787 (2004) required all school districts to "formulate a plan in sufficient detail that measurable objectives can be identified and addressed which will accomplish English language acquisition and improved academic performance. Monies distributed to school districts based upon the population of limited-English proficiency students and distributed to school districts to support programs for students with non-English or limited-English proficiency shall be utilized in support of the district plan." For more information, see http://www.sde.idaho.gov/LEP/guidance.asp.

KY	The *Kentucky ELL Handbook* (2007 provides guidance and support, specifically regarding assessment and evaluation of ELL students, to decision-makers, administrators, school psychologists, speech-language pathologists, teachers and other school personnel, who directly or indirectly work with underachieving English Language Learners (ELLs). Chapter 9 deals specifically with "Designing Culturally & Linguistically Appropriate Interventions and Measuring Their Impact and states "it is highly recommended that a problem solving team or TAT be involved in collecting and organizing information needed to clearly define the problem, propose, design and implement interventions, and supervise the systematic and repeated assessment of the impact of the interventions. For more information, see http://www.warren.k12.ky.us/~esl/Kentucky%20ELL%20Evaluation%20Handbook.pdf.
MA	Massachusetts's *Coordinated Program Review Procedures, School District Information Package, English Learner Education (ELE) in Public Schools (2010-2011),* provides "guidance to districts in preparing for the English Learner Education (ELE) portion of the Department of Elementary and Secondary Education's Coordinated Program Review". The document also contains the *ELE Program Director's Checklist* that provides a concise overview of all the activities to be completed. (p. 2). For more information, see www.doe.mass.edu/pqa/**review**/cpr/instrument/chapter71A.doc
MI	Michigan's *Title III Program Evaluation* (December 2008), provides districts with a Program Checklist and a Program Evaluation Guide. The evaluation document includes a self-study guide and provides a process to review the efficacy of school's/district's Title III program. The evaluation checklist is organized into six areas, including: • Administrative Responsibilities • Student Identification • Parent and Community Involvement • Instructional Programs and Assessment • Teacher Qualification and Professional Development • Program Evaluation and School Improvement Each of these areas requires documentation, evidence, and level of compliance for a series of well defined questions. For more information, see http://www.misd.net/bilingual/MITitleIIIProgramEvaluation.pdf.

MN	Minnesota's *Accountability for Implementation (n.d.),* provides school's with a quick self-evaluation of its efforts to meet the needs of its ELL students and provide a rough idea of where strengths and weaknesses may be found in order to focus program improvement efforts. Schools must measure ELL student's progress annually and this data is also utilized in evaluating the effectiveness of a school's existing ELL program. Other program components that "may" be evaluated are: • Effective use of technology • Teacher instruction • Materials • Coordination with the grade level/subject matter curriculum. The checklist provided is not intended to take the place of an evaluation instrument; rather the intent is to provide teachers/school administrators with a rough idea of where their ELL program's strengths and weaknesses lie in order to facilitate program improvement efforts. For more information, see http://www3.mpls.k12.mn.us/departments/ell/ell_manual/account_implement.pdf.
NY	New York's *LEP/ELL Program Evaluation Toolkit (PET): A Self Evaluation* provides education leaders with a comprehensive assessment tool developed to examine the quality of programs and services provided to ELL students. The self-evaluation requires schools to determine if, and to what degree, the school's programs: • Align with the core curriculum • Demonstrate the rigor and effectiveness of the professional development plan • Reflect the support teacher received from the district, BOCES and school instructional specialists in implementing best practices in the classroom • Comply with language allocation policies • Benefit from rigorous monitoring and assessment • Improve communication with parents and families of ELL students. For more information on the Program Evaluation Toolkit (PET), see For more information on the Evaluation Tool, see http://www.p12.nysed.gov/biling/docs/LEPELLPETInfo.pdf

| OK | Oklahoma requires a District On-Site Review and a School On-Site Review of ELL programs.

For more information on the state's District On-Site Review, see http://sde.state.ok.us/curriculum/Bilingual/pdf/District_On-Site.pdf

For more information on the state's School On-Site Review, see http://sde.state.ok.us/curriculum/Bilingual/pdf/School_On-Site.pdf |
|---|---|
| WI | Wisconsin's *Reasons to Evaluate a Language Assistance Program: Are We Making Progress?* requires school districts to modify their programs if they prove to be unsuccessful after a legitimate trial. The Wisconsin Department of Public Instruction provides districts with a sample "Program Evaluation: School Improvement Planning Checklist" that includes questions relative to:

- Identification & Assessment
- Programming & Educational Approaches
- Staffing & Professional Development
- Parental Involvement

For more information, see http://www.cesa7.org/ellcenter/Resources/documents/program-eval.pdf. |

Program Evaluation: District Level

DISTRICT	

Lake Oswego School District, Oregon	The District conducts an annual evaluation of its ELL program that states the extent to which its goals have been met and identifies program weaknesses with specific recommendations regarding actions that should be taken. This guide provides:
	
	• Key evaluation elements
	    o Program alignment with stated goals
	    o Collection of student performance data
	    o Review of results
	    o Plan for modification/or improvements
	    o Implementation of program changes
	    o Ongoing review
	
	For more information, see http://org.loswego.k12.or.us/ELL/Evaluation.htm.

Webster Groves School District, Missouri	The Webster Grove School District reviews several key components to make recommendations for program improvement, including: • Description of Program Services • Description of Identification Process • Enrollment over the past five years • Percent of ELL student scoring in the top two of MAP • Percentile Rank for Students on the MAC II • Results of Teacher Survey • Results of Parent Focus Group For more information, see http://www.webster.k12.mo.us/education/components/scrapbook/default.php?sectiondetailid=39706&&PHPSESSID=xarzunswij.

Ongoing Assessment & Reclassification

STATE	
CO	In Colorado, LEP student re-designation is determined "through valid and reliable assessments and documented through observation". Re-designated students must be monitored for two years. "While not required by law, it can be useful for districts to establish a category of "Exited" or "Formerly ELL" (FLEP) for purposes of tracking student progress and to help alert teachers to the fact that students began their education as ELLs and probably live in bilingual environments." For more information, see, http://www.cde.state.co.us/cde_english/download/ELLGuidebook/Final%20Guidebook07-08.pdf.

Using Student Data to Inform Program Development

STATE	
MA	Annually or at the least bi-annually, school districts in Massachusetts should evaluate the effectiveness of its language program. The school district should consider the progress of its ELL students in acquiring English and maintaining academic progress. Districts should also evaluate longitudinal data that compares the academic progress of the former ELL student who is now fully English proficient with that of other non-ELL students in categories that include grade point averages; national and state test score averages; and rates of retention, dropping out, graduation, and receipt of honors and awards. A district whose program is not demonstrably effective in meeting the needs of ELL students must modify its program in a timely manner. For more information, see *http://www.mde.k12.ms.us/acad/id/CurrEnr/ELL/What_to_Do_Five_Steps.pdf*.

Implementing Programs in Low Incidence Districts/Schools

STATE	
MA	Massachusetts provides guidance on *Designing and Implementing Sheltered English Immersion (SEI) Program in Low Incidence Districts (March 2006)*, Provides three scenarios illustrating how three low-incidence districts in Massachusetts have sought to implement SEI in their districts. State law, G.L.c.71A, requires that most LEP students be educated in a Sheltered English Immersion (SEI) program. An SEI program consists of both *sheltered subject matter instruction in English* and *English language instruction*. This requirement applies to all districts that enroll LEP students, regardless of whether there is one LEP student or hundreds of LEP students enrolled in the district. For more information, see *www.doe.mass.edu/ell/**sei/lowincidence**.doc*..