School Reform and Standards-Based Education: A Model for English-Language Learners

JANA ECHEVARRIA
California State University, Long Beach

DEBORAH SHORT
Center for Applied Linguistics

ABSTRACT The authors examined a model of instruction for English-language learners (ELLs) who were learning academic English while they tried to meet content standards required by the nation's education reform movement. In previous work (J. Echevarria, M. E. Vogt, & D. Short, 2000), the authors developed and validated a model of instruction (Sheltered Instruction Observation Protocol; SIOP model) for ELLs. In this study, the authors tested the model for its effects on student achievement. Findings revealed that students whose teachers implemented the SIOP model performed slightly better than did a comparison group on an expository essay writing task, which closely approximated academic assignments that ELLs must perform in standards-based classrooms.

Key words: English-language learners, instructional model, school reform, standards-based education

School reform is a focus of the national education agenda—high academic standards are implemented in every state, and federal legislation requires annual testing of Title I students in Grades 3–8 and again in high school. Moreover, English-language learner (ELL) students are tested every year until they are proficient in English. Standards and the assessments that are aligned with them have become the rallying principles for improved academic performance in schools (Tucker & Codding, 1998).

The No Child Left Behind (NCLB) Act of 2001, the cornerstone of the present education agenda, provides significant federal dollars to states through a formula funding grant. States, in turn, provide funds to their local school systems. The monies constitute the largest increase ever in the funding of Title I programs, which were designed to educate economically disadvantaged students. Approximately 125 of the nation’s poorest urban districts have received the increased funds. A similar increase in Title III funds has passed through states to districts with limited English-proficient (LEP) students. (Federal legislation refers to students who are not fluent speakers of English as LEP students. In practice, the term ELL is used more commonly. The term ESL, or English as a second language, refers to a type of class, instructional program, or curriculum.) Whereas increased funding indicates a considerable commitment to improved education for all children, it is unclear whether current instructional practices accomplish that goal for significant numbers of ELLs. Will many children inadvertently be left behind because educators missed the mark with the ELL population?

We tested a model of instruction for ELLs, the Sheltered Instruction Observation Protocol (SIOP) model, to ascertain its effects on academic literacy development. If schools are to provide a quality education for all children, it is critical that teachers implement empirically sound practices, especially for ELLs, who consistently underperform in academic settings (Moss & Puma, 1995; Snow & Biancarosa, 2003; Wainer, 2004).

The level of academic achievement for ELLs has lagged significantly behind that of their language-majority peers. For example, in California, where over 1.5 million ELLs attend public school, students who are not proficient in English perform less well on standardized tests than do students who are proficient in English. Eleven percent of 7th-grade ELLs who took the reading portion of the state test in 2002 scored at or above the 50th percentile, compared with 57% of English-proficient language-minority students and 48% of all students who took the tests (California Department of Education, Educational Demographics Unit, 2004).

Furthermore, most ELLs in U.S. schools are of Hispanic descent; recent national tests of reading and writing (i.e., the National Assessment of Educational Progress; NAEP) show that at the three grade levels tested (4, 8, 12), many more Hispanic students performed at the below-basic level than did White students and Asian/Pacific Islander students. Far fewer students performed at proficient or advanced levels than did those groups (National Center for Education Statistics, 2002). That finding is particularly
noteworthy because the NAEP examinations usually exempt students at beginning levels of ESL proficiency (Grigg, Daane, Jin, & Campbell, 2003).

It is also significant that ELLs have high dropout rates and are more frequently placed in lower ability groups and academic tracks than are language-majority students (Bennici & Strang, 1995; Cummins, 1994; “Latinos in Education,” 1999; Snow & Biancarosa, 2003). Waggoner (1999) reported that about 13% of “newcomer and linguistically different” youth have either never been enrolled in U.S. schools or have left before completing high school. A recent study of high school attrition in Texas (Johnson, 2004) showed that 49% of Hispanic students who were ninth graders in 2000–2001 left high school before graduation, compared with only 22% of White students. Another study of districts in the South showed similar discrepancies between Hispanic dropout rates and White dropout rates (Wainer, 2004). Many of the students in both studies were ELLs.

Since the NCLB Act has been implemented, many states have required that students pass particular subject-area tests to obtain a high school diploma. Unfortunately, there is now an increase in the number of ELLs who are not receiving a diploma because they have failed such high-stakes tests despite fulfilling all other graduation requirements (Snow & Biancarosa, 2003).

Students have difficulty in school for a number of reasons; one is the mismatch between student needs and teacher preparation. Although the NCLB Act calls for highly qualified teachers in every core academic classroom by 2006 (2003 for new Title I teachers), the supply of certified ESL and bilingual teachers is too small for the demand. National studies (e.g., National Commission on Teaching and America’s Future, 1996) and regional and district-level studies (e.g., Wainer, 2004) have reported significant shortages of teachers qualified to teach students with limited English proficiency and of bilingual teachers trained to teach in a second language. Fewer than 13% of teachers in the nation have received professional development to prepare them for teaching linguistically and culturally diverse students (National Center for Education Statistics, 2002). To compensate, principals hire less qualified teachers, use substitute teachers, cancel courses, bus students elsewhere, require reading specialists to fill the void, increase class size, or ask teachers to teach outside their field of preparation (Vogt & Shearer, 2003; Wainer). It is not uncommon to find untrained paraprofessionals acting as English-language teachers for ELL students (Lavadenz, 1994; Pickett, 1999; Rueda, Monzo, & Higareda, 2004).

Moreover, federal guidelines regarding highly qualified teachers focus only on core subject-area teachers, requiring them to have a deep understanding of their subject matter but not requiring such teachers who have ELLs in their classes to have a commensurate level of understanding of second-language acquisition, ESL methods, or sheltered teaching methods. In the 1999–2000 Schools and Staffing Survey (National Center for Education Statistics, 2002), 41.2% of 2,984,781 public school teachers reported teaching ELL students, but only 12.5% of those teachers had 8 or more hr of training in the previous 3 years. Yet, the consensus from research on effective professional development is clear that 8 hr is not even the minimum that is needed to learn new approaches to teaching, such as strategies to teach ELLs well (Borko, 2004; Gonzales & Darling-Hammond, 1997).

As a result of those policies and practices, many ELLs receive much of their instruction from content-area teachers or paraeducators who have not had appropriate preparation or professional development to address their second-language development needs or to make content instruction comprehensible. This situation hinders their academic success. Not only do teachers need more preparation to work with ELLs but they also need to know the type of instruction that is most effective for these students, a population whose growing numbers require that educators take a serious look at their instructional programs.

Student Demographics

Across school districts in the United States, the number of students from non-English-speaking backgrounds has risen dramatically, representing the fastest growing segment of the student population. From the 1992–1993 school year through the 2002–2003 school year, the number of LEP students in public schools increased 84%, whereas total enrollment increased only 11% (National Clearinghouse for English Language Acquisition [NCELA], 2005). In 18 states, LEP enrollment grew more than 200% between 1992–1993 and 2002–2003. The 2000 U.S. Census reported that one in five school-aged children is a nonnative English speaker (Jamieson, Curry, & Martinez, 2001). As of the 2002–2003 school year, there were more than 5 million LEP students in the nation’s pre-K–12 schools, approximately 10% of the total enrollment (NCELA).

The rise in the number of LEP students is related strongly to the increased immigrant population in the United States. The U.S. Census Bureau (Jamieson et al., 2001) determined that in 1999, 20% of school-aged children had at least one parent who was an immigrant, and 5% of the students were immigrants. However, when race or Hispanic origin is considered, the distinctions among students become more apparent. Sixty-five percent of Hispanic students and 88% of Asian and Pacific Islander students had at least one immigrant parent (Jamieson et al.). Although not all Hispanic students or Asian students are ELLs, Hispanic students comprise 75% of all students in ESL, bilingual, and other English-language support programs, according to the “Latinos in Education” report (1999) published by the White House Initiative on Educational Excellence for Hispanic Americans. The Office of English Language Acquisition, U.S. Department of Education, reported that in 2000, Asian English learners accounted for almost 10% of all ELLs in the nation.
Who Are ELLs?

ELLs come to U.S. schools with many resources to share in classrooms, including linguistic resources in their native language. Yet, educators, policy makers, and the public should understand that all students who are learning English as an additional language are not alike. They enter schools with a wide range of language proficiencies (in English and in their native language) and of subject-matter knowledge. ELLs differ in their education backgrounds, expectations of schooling, socioeconomic status (SES), age of arrival in the United States, and personal experiences coming to, and living in, the United States (Waggoner, 1999).

Among immigrant students, some ELLs have strong academic preparation. They are at or above equivalent grade levels in the school curricula and are literate in their native language. For the most part, those students need English language development so that as they become more proficient in English, they can transfer their educational knowledge to their courses. A few subjects, such as U.S. history, may need special attention, because the students have not studied them previously. Those students have the greatest likelihood of having educational success if they receive appropriate language and content instruction in their schools.

Other immigrant students arrive at U.S. schools with limited formal schooling—perhaps because of war or the isolated location of their home. Ruiz-de-Velasco and Fix (2000) found that 20% of all LEP students at the high school level and 12% of the LEP students at the middle school level have missed 2 or more years of schooling since age 6. Among Hispanic students aged 15–17, more than one third are enrolled below grade level (Jamieson et al., 2001). Those students are not literate in their native languages and have not had schooling experiences such as changing teachers according to subject or taking a standardized test. They have significant gaps in their educational backgrounds, lack knowledge in specific subject areas, and often need additional time to become accustomed to school routines and expectations. They need literacy, English-language development, and content-area knowledge (Boyson & Short, 2003).

Schools also enroll students who have been raised in the United States but speak a language other than English at home. Although most of those students learn English during their elementary school years, some, nonetheless, reach secondary levels without having mastered English or the home language and may be caught in a state of semiliteracy, which is hard to escape (MacSwan & Rolstad, 2003; MacSwan, Rolstad, & Glass, 2002).

For the most part, when nonnative English speakers enter school, they do not have the same language skills and background in English as do native speakers from English-speaking homes. When native English-speaking children enroll in school, they have some oral proficiency and an understanding of the grammatical system. Some have knowledge of the alphabet and may have initial skills in reading and writing. Curricula and instruction build from the expectation that students know some English when they start school and rely especially on oral language proficiency. Yet, ELLs who enter school at all grades rarely have the level of proficiency in English found in native English-speaking students in kindergarten or first grade.

Although not acknowledged or understood by many educators, age-appropriate knowledge of the English language is a prerequisite for attaining content standards. Without oral and written English language skills, students have difficulty learning and demonstrating their knowledge of mathematical reasoning, science skills, social studies concepts, and so forth (Slavin & Cheung, 2003). Students who lack proficiency in English are at a decided disadvantage in school.

Students with limited formal schooling and below grade-level literacy are most at risk for educational failure. Although most ELLs show sufficient growth in acquiring social language skills (i.e., basic reading and conversational skills geared to many out-of-school and survival interactions) in a few years (Collier, 1987; Thomas & Collier, 2002), they are less successful in acquiring the language needed for school tasks like reading textbooks, participating in content-related classroom discourse, and writing research reports. Some studies of students placed in traditional ESL and bilingual programs reveal that most of those students need 5–9 years of instruction before their academic scores are at the average level of native English-speaking students (Collier, 1995; Thomas & Collier). That longer length of time is particularly difficult for educators to provide for secondary students who face graduation course requirements and high school exit examinations.

High Academic Standards for All Students

Historically, ELLs had time to learn English before they were placed in mainstream content classes and completed standardized tests in subjects such as reading and mathematics. Schools offered ESL or bilingual education pro-
grams to ELLs with specially trained teachers, yet kept those teachers and students separate from regular school programs. State and local policies typically exempted limited-English-proficient students from standardized tests for up to 3 years (Rivera, Stansfield, Scialdone, & Sharkey, 2000). However, that situation has shifted significantly.

In the past, ELLs would generally exit an ESL or bilingual program when they were proficient in English and able to perform subject-area coursework in regular English classrooms. At present, students exit ESL or bilingual programs before they are proficient in academic English for several reasons: (a) The number of ELLs increased without a comparable increase in certified teachers, so schools could not relegate the education of these students to separate specialized classes; (b) state legislatures have enacted policies that limit the number of years that students are permitted access to language-support services (California, Arizona, and Massachusetts require that students move into regular classrooms after 1 year), and bilingual education (an option whereby some ELLs can keep up with grade-level content) is severely restricted; and (c) districts place students in content subjects taught in English before they determine that such instruction will enable them to perform well on state standards-based assessments (Goto-Butler, Orr, Bousquet Gutierrez, & Hakuta, 2000).

The standards movement that began in 1989 at the National Governor’s Association summit on education in Charlottesville, Virginia, resulted in an agreement on the need for national education goals (Tucker & Coddington, 1998). Shortly after the summit, the National Council of Teachers of Mathematics (NCTM, 1989) released its national mathematics standards document, and subsequent federal legislation encouraged a widespread movement among other professional associations to develop standards for specific content areas. The intent was to (a) promote high expectations for student learning and (b) use the national standards as guidelines for state and local curriculum and assessment design and for professional teacher development.

Those laudable goals offered substantial opportunities for improving the education of ELLs. Whereas ESL instruction in the United States had been unsystematic and varied often from district to district and from state to state (Sheppard, 1995), the focus on standards served as a catalyst for educators involved in ELL schooling. The Teachers of English to Speakers of Other Languages (TESOL, 1997) developed ESL standards for students in pre-K–12 programs, and states adopted or adapted these standards to develop curriculum frameworks. The ESL standards focused on academic English-language skills, along with social-language proficiency and cultural norms of language use. ESL and bilingual educators thus had guidelines to increase the academic rigor of their language instruction and curricula. However, the other national content area standards did not make accommodations for second-language learners.

Thus, the challenge of teaching grade-level content to ELLs is acute because they have to learn academic concepts, discourse, skills, and behaviors in a language that they do not speak, read, or write proficiently. ELLs are expected to study the content subjects and demonstrate proficiency on the assessments, even though language acquisition is a long-term process, and beginning speakers of English are likely to need 4 or more years of instruction before they are proficient (Thomas & Collier, 2002). Furthermore, too few content teachers have the skills to teach content and promote language development simultaneously.

Besides instruction, assessment practices reveal another concern. Most states are developing or have adopted high-stakes tests based on state standards. As a result of the NCLB Act, all states are required to have standards-based reading and mathematics tests for students in Title I programs. As of February 2004, the U.S. Department of Education allows states to “not count” the scores of newly arrived ELLs in their accountability reports for the 1st year but still encourages all students to be tested. That practice has a negative implication for ELLs because most of the tests now used were designed for native English speakers who have spent their educational careers in U.S. schools.

Overall, the stakes for ELLs have been raised significantly as states and federal programs have restructured their accountability measures. The NCLB Act is an admirable goal but one that requires a specific plan for ELLs. Such a plan must include teachers’ use of instructional practices that will facilitate students’ academic literacy development so they can meet the high standards in all academic areas.

### Academic Literacy Development

The focus of ELL education must be more than learning to speak and understand English. Students may learn to communicate in English, but there is a distinction between conversational English and academic English (Cummins, 1981, 2003; TESOL, 1997). Students who have a moderate command of spoken English in social settings require support when trying to comprehend academic English.

As an example of text that is difficult to comprehend, consider the following passage:

> Traditionally, individuals enjoy owembla, especially on weekends. In fact, lack of owembla can result in hard feelings when overlooked or neglected. However, it should be noted that some people shy away from owembla. All in all, owembla are enjoyed by most people in many different cultures.

The passage has one word, owembla, (a nonsense word that carries the meaning “parties” in this passage) that is unfamiliar to the reader and renders the passage difficult to comprehend. For many ELLs, additional words in the passage may be unfamiliar, such as “hard feelings” and “neglected.” Moreover, contextual clues such as “especially on weekends” and idioms like “shy away from” may make the meaning even more confusing if the student does not have the background schema to use them. When one considers how many words and sociocultural referents are unfa-
miliar to ELLs, we begin to understand the difficulties they experience in the U.S. education system.

For ELLs to succeed in school, they must master not only English vocabulary and grammar but also the way that English is used in academic subjects. “Academic English” includes semantic and syntactic knowledge, along with functional language use. For example, students who use English must be able to (a) read and understand expository prose, such as that found in textbooks; (b) write persuasively; (c) argue points of view; and (d) take notes from teacher lectures. Those students also must articulate their thinking skills in English—make hypotheses and predictions, express analyses, draw conclusions, and so forth. In their various content classes, ELLs must pull together their emerging knowledge of the English language with the content knowledge they are studying to complete academic tasks associated with the content area. They also must learn how to do those tasks, such as generate an outline, negotiate roles in cooperative-learning groups, and interpret charts and graphs. The combination of the three knowledge bases—English, content topic, and the manner in which tasks should be accomplished—constitute the major components of academic literacy (Short, 2002).

Another consideration for school success is the explicit socialization of students to the often-implicit cultural expectations of the classroom, such as turn taking, participation rules, and established routines. Erickson and Shultz (1991) reported that student comfort with the social participation structure of an academic task can vary according to culturally learned assumptions about appropriateness in communication and in social relationships, individual personality, and power relations in the classroom social system and in society at large. Many ELLs could therefore benefit from explicit guidance about expected appropriate classroom behaviors and interactional styles. However, Bartolomé (1994) argued that teachers need to engage in culturally responsive teaching so their instruction is sensitive to, and builds upon, culturally different ways of learning, behaving, and using language.

Finally, the methods that teachers typically use, especially in the upper elementary and secondary schools, tend to not facilitate learning or literacy instruction for ELLs (Tharp, Estrada, Dalton, & Yamauchi, 2000). Reliance on oral instruction through lectures makes understanding information difficult. Paper-and-pencil tasks, like worksheets, that do not provide learning scaffolds for students also may be challenging. Textbook features intended to aid student understanding may have the opposite result for students who do not know how to use features such as bolded words, headings, sidebars, and graphs. Many ELLs have difficulty tracking the flow of information on cluttered pages. Furthermore, students who arrive in the United States beyond the initial age for literacy instruction and are not literate in their native language find that teachers are underprepared to teach basic literacy skills (Fillmore & Snow, 2002; Rueda & García, 2001).

**Sheltered Instruction for Accessing Core Content and Meeting Standards**

One direction that educators have taken to accommodate the need for teaching more academic content to ELLs while they are still learning English has been to incorporate more sheltered instruction (SI), or specially designed academic instruction in English (SDAIE; as it is called in California), in their educational programs. SI is an instructional approach that makes grade-level academic content in areas such as social studies, mathematics, and science accessible for ELLs by incorporating specialized strategies and techniques that accommodate the second-language acquisition process (Genesee, 1999). SI teachers use the regular core curriculum and modify their teaching to make the content understandable for ELLs while promoting their English-language development.

Some of the techniques that characterize SI include slower speech and clear enunciation, use of visuals and demonstrations, scaffolded instruction, targeted vocabulary development, connections to student experiences, student-to-student interaction, adaptation of materials, and use of supplementary materials (Addison, 1988; Echevarría, 1995; Echevarría & Graves, 2003; Genesee, 1999; Kauffman, Sheppard, Burkart, Peyton, & Short, 1995; Short, 1991; Vogt, 2000). Such strategies for content-area teachers are promoted by school districts, teacher training programs, and the literature (Crawford, 2003); with the ever-growing number of ELLs in U.S. schools, all teachers need to be aware of these instructional approaches for content classes.

Although most educators agree on these features as being important for SI for ELLs, there has not been an explicit model for effectively delivering sheltered lessons nor many investigations in which researchers measure what constitutes an effective sheltered lesson (August & Hakuta, 1997). As a result, SI has been implemented unevenly across districts and schools (Sheppard, 1995). The underachievement of ELLs on state and national assessments indicates that for school success, teachers must do more than simply implement a few strategies from ESL methodologies, such as showing visuals or slowing down the rate of speech. Those strategies may help students access the content concepts, but without systematic language development, students never develop the requisite academic literacy skills needed for achieving success in mainstream classes, for meeting content standards, or for passing standardized assessments.

That situation can thus be a tremendous challenge for many regular classroom and content teachers who instruct ELLs in their classes. Most of those teachers are held accountable for their students’ performance but are without key training or experience in effective instruction for ELLs. Although teachers can learn numerous techniques for helping ELLs understand the gist of subject content, teachers’ overall successes to date have been limited with this population. What may contribute to the situation is the
lack of a coherent, scientific model of SI that would facilitate teachers' systematic implementation of features known to be effective instructionally.

To address that need, we undertook a national research project to develop such a model of SI for ELLs, then ascertained its effects on their academic literacy development.

SIOP Model Development

“The Effects of Sheltered Instruction on the Achievement of Limited English Proficient Students” is a 7-year research project (1996–2003) that we conducted for the Center for Research on Education, Diversity & Excellence (CREDE), a national research center funded by the U.S. Department of Education, Office of Educational Research and Improvement (now known as the Institute of Education Sciences). Housed at the University of California, Santa Cruz, CREDE supported 30 research studies throughout the United States.

We report on the following research questions that we addressed in the national research project:

1. Does SI improve the achievement of ELL students in content areas such as social studies?
2. Are there significant differences in achievement data (reading scores, writing samples, attendance) for students of project teachers versus students in sheltered classes whose teachers have not received SIOP development training?

In the first 2 years of the project, we collaborated actively with practicing middle school teachers to refine the SI model and to implement it in their classrooms. We identified, according to literature review and classroom research, effective teaching strategies involved in SI, such as scaffolding, learning strategies, literacy techniques, and meaningful curricula and materials.

With a process similar to a design experiment (Cobb, Confrey, diSessa, Lehrer, & Schauble, 2003; The Design-Based Research Collective, 2003), the SIOP model was developed through a cyclical process, wherein researchers and project teachers designed, used, analyzed, and redesigned features of the model. The model was developed initially as the SIOP, a research observation instrument. It is a rubric that allows researchers to score teachers along a continuum of performance for each feature, thus determining how well teachers were including the essential features of effective SI in their lessons (Echevarria & Short, 2004; Short & Echevarria, 1999). Figure 1 shows the evolution of the model and its various applications.

Teachers used the lesson planning guide to implement the model; researchers used the protocol to measure fidelity of the teachers’ implementation of the SIOP model. Design-based research requires collaboration between

![SIOP model](image)

**FIGURE 1.** SIOP model. Hypothesis for the present study was: Does implementation of the SIOP model affect student achievement? SIOP = Sheltered Instruction Observation Protocol; ELL = English-language learners.
researchers and practitioners to produce meaningful change in instructional practice by studying learning in context (Cobb et al., 2003); this was at the heart of our CREDE research.

The protocol evolved into a lesson planning and delivery approach known as the SIOP model (Echevarria, Vogt, & Short, 2000, 2004). It provides concrete examples of features of SI that can enhance and expand teachers’ instructional practice. In brief, the SIOP model offers a framework for teachers to present curricular content concepts to ELLs through strategies and techniques that make new information comprehensible to the students. While doing so, teachers develop student language skills across the domains of reading, writing, listening, and speaking. Teachers might accomplish that task in ways that are suited to a particular lesson by asking students, for example, to (a) engage in peer discussions or a class debate, (b) read textbook chapters or supplementary materials, (c) complete a graphic organizer, (d) write in a journal, or (e) compose an essay. Some teachers might implement reader or writer workshops or a teacher-directed reading lesson.

The observation protocol is composed of 30 items that are grouped into eight components essential for making content comprehensible for ELLs (see Appendix): (a) Preparation, (b) Building Background, (c) Comprehensible Input, (d) Strategies, (e) Interaction, (f) Practice/Application, (g) Lesson Delivery, and (h) Review/Assessment.

The six features under Preparation examine the lesson-planning process, including the incorporation of language and content objectives, use of supplementary materials, and the meaningfulness of activities. Building Background focuses on making connections with students’ background experiences and prior learning and developing their academic vocabulary. Comprehensible Input considers adjusting teacher speech, modeling academic tasks, and using multimodal techniques to enhance comprehension. Strategies emphasizes explicit teaching of (a) learning strategies so that students know how to access and retain information, (b) scaffolding instruction, and (c) promoting higher order thinking skills.

Interaction reminds teachers to encourage elaborated speech and to group students appropriately for language and content development. Practice/Application calls for activities that extend language and content learning, and Lesson Delivery ensures that teachers present a lesson that meets planned objectives. As part of Review/Assessment, four items consider whether the teacher (a) reviewed key language and content concepts, (b) assessed student learning, and (c) provided feedback to students on their output.

The SIOP model shares many features recommended for high-quality instruction for all students, such as linking lesson objectives to content standards, but adds key features for the academic success of students learning through a second language, such as inclusion of language objectives in every content lesson and development of background knowledge.

There is no required routine to follow when one implements the SIOP model. However, some specific attention to academic literacy development is required in each SIOP-designed lesson, particularly in the form of language objectives and contextual language practice. Moreover, the emphasis that SIOP lessons place on building vocabulary and background knowledge and developing learning strategies contributes to students’ literacy development.

One strength of the SIOP model is that it allows for natural variation in classroom implementation while it provides teachers with specific lesson features that, when implemented consistently and to a high degree, are likely to lead to improved academic outcomes for ELLs. Another strength of the model is that the protocol provides a rating scale so that lesson observations may be scored. That feature is important for teachers’ own professional growth and development; it allows administrators and university field supervisors to provide concrete feedback to the teachers they observe and evaluate. The rating scale also is effective as a coaching tool for schools in which peer coaching is used. A more detailed discussion of the SIOP model is found in Making Content Comprehensible for English Learners: The SIOP Model (Echevarria et al., 2004).

During the first months after the model was developed, Guarino and colleagues (2001) conducted a study to establish the validity and reliability of the SIOP instrument, which researchers use to rate teacher implementation of the SIOP model of sheltered instruction. A statistical analysis revealed an intrarater agreement of .99. Additional analyses indicated that the SIOP instrument is a highly reliable and valid measure of SI.

Research Methodology

After we developed the SIOP model and validated the protocol, we trained a group of teachers to implement the model to a high degree (Short & Echevarria, 1999). That cadre of teachers provided the classrooms in which student achievement could be measured.

Setting

The research occurred in one West Coast and two East Coast public school districts (see Table 1) from 1998 to 2000. The West Coast school district was a large, diverse, urban school district with approximately 90,000 students,

<table>
<thead>
<tr>
<th>TABLE 1. Setting</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group</td>
</tr>
<tr>
<td>Intervention</td>
</tr>
<tr>
<td>Comparison</td>
</tr>
</tbody>
</table>
89 schools, and a student population that was 45% Hispanic, 20% African American, 18% Caucasian, and 11% Asian/Pacific Islander (see Table 2). Approximately 32% of the students in the district were identified as ELLs. All West Coast data originated from students in three middle schools with similar demographics. Eight intervention teachers were from one middle school, two intervention teachers were at a second school, and three comparison teachers were from a third school.

The two East Coast districts were in a large metropolitan area with a combined number of approximately 166,000 students distributed across more than 220 schools. In one district, the student population was 41% Caucasian, 32% Hispanic, 17% African American, and 10% Asian/Pacific Islander. Twenty-three percent of the students in the district were identified as ELLs. In the other district, the student population was 61% Caucasian, 11% Hispanic, 11% African American, and 14% Asian/Pacific Islander. Eight percent of the students were identified as ELLs. The East Coast data came from the students of nine teachers at four middle schools; one additional teacher and school provided the comparison group.

The ELLs in the East and West Coast districts had a distinct program of study whereby they participated in sheltered classes for the core subjects—English-language arts, social studies, mathematics, and science. The students did not have a designated ESL class. On the East Coast, only ELLs were in the sheltered classes, and the teachers used district-developed sheltered curricula that were aligned with the regular grade-level curricula. On the West Coast, in a few teachers’ classes some native English-speaking students were in the sheltered classes with ELLs because of issues of scheduling, but they were not part of the data collection. Teachers in the West Coast sheltered classes used the same core curriculum as in mainstream classes, but the sheltered teachers modified and adapted the curriculum to make it more understandable for their ELL students. In addition, social studies and English-language arts teachers functioned in teams to reinforce and support concepts across the subject areas. On both coasts, ELL students were mixed with native English speakers for the elective classes.

All of the intervention classes were designated SI courses with the exception of one health class on the East Coast. In that class, only the ELL students participated in the data collection and analysis.

Participants

Three hundred forty-six students in Grades 6–8 participated in the intervention classes. Of those participants, 237 students were from the West Coast school district and 109 students were from the East Coast districts. Approximately 54% of the participants were boys, and 46% were girls. Among the intervention students, 56% were Hispanic, 41% Asian/Pacific Islander, 2% Caucasian, and 1% African American. Four percent of the students had identified disabilities and received special education services.

We identified a comparison group of students and, although not specifically matched student to student, the comparison group emanated from a different school on each coast with similar student populations in terms of ethnic makeup, proportion of ELLs, and free or reduced-price lunch status. Ninety-four students in Grades 6–8 participated in the comparison group and in the same instructional program as the students in the intervention classes (i.e., designated sheltered content classes). Approximately 57% of the participants were boys; 43% were girls. Among the comparison students, 69% were Hispanic and 31% were Asian/Pacific Islander. None of the students in the comparison group received special education services.

At the time of the study, all participants were designated as ELL by their school districts on the basis of their performance on the districts’ language-proficiency assessments. The students represented the full range of proficiency from nonliterate beginners to advanced-level ESL students. One issue inherent in research with ELLs is the variability in identification and designation procedures. Goto-Butler and colleagues (2000) concluded in their studies of ELL students that “Districts vary considerably in both whom they included as LEP and non-LEP students . . . and rates of redesignation vary from year to year, even within the same district” (p. 144). In the present study, each of the three districts used a different classification system, and the largest district used two classification systems. Thus, any comparisons among students’ English-proficiency levels would be somewhat arbitrary. However, in each classification system, the majority of students fell within the midrange, meaning that the comparison and intervention groups were mostly advanced-beginning to intermediate speakers. Also, students in both groups shared similarities in related factors, such as ethnicity and schoolwide test scores. Rather than presuming to equate classification systems, we relied on the fact that all students in the study were ELLs who were enrolled in sheltered content classes (i.e., classes structured to give students extra support in designated academic areas.

Other important similarities existed between the two groups. Most students were from low-SES backgrounds, as

<table>
<thead>
<tr>
<th>TABLE 2. Student Demographics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Characteristic</td>
</tr>
<tr>
<td>Gender</td>
</tr>
<tr>
<td>Female</td>
</tr>
<tr>
<td>Male</td>
</tr>
<tr>
<td>Race or ethnicity</td>
</tr>
<tr>
<td>Hispanic</td>
</tr>
<tr>
<td>Asian/Pacific Islander</td>
</tr>
<tr>
<td>White</td>
</tr>
<tr>
<td>Black</td>
</tr>
</tbody>
</table>
measured by eligibility for the free and reduced-price lunches program. The predominant native language among
the intervention and comparison students was Spanish (more than 50% in all districts); also, a large number of stu-
dents spoke Cambodian, Vietnamese, and Korean. Some students spoke other Asian, African, Caribbean, European,
or Native American languages. Ethnicities included Hispanic, Cambodian, Vietnamese, Arabic, and Korean. The participants' countries of origin were diverse and included Mexico, Cambodia, El Salvador, Pakistan, Vietnam, Bolivia, Somalia, Yemen, and others.

Procedures

In the 1st year of the study, participating school districts nominated the project teachers. Criteria for nomination included subject area, willingness to participate, and 2 or more years of experience with SI. In subsequent years, other teachers were invited to join the study or were recruited by current project teachers. Project teachers spent 1 or 2 years (some teachers joined the study in 1997 and others in 1998) learning and practicing the SIOP model through an extensive professional development process; we measured implementa-
tion with the SIOP. The project teachers and researchers formed a learning community to refine the model through an examination of the teachers' classroom practices and student response to the SIOP lessons.

On both coasts, teachers participated in 3-day staff development institutes on the SIOP model each summer and in several reunion meetings during each school year. During the workshop sessions, teachers and researchers (a) developed a common understanding of the SIOP model or practiced strategies that integrate language and content knowledge or (b) scaffolded instruction to enhance comprehension, examined state standards in their particular content areas to determine associated language objectives and academic tasks that ELLs could learn, and established a lesson planning process. Teachers discussed student reaction to SIOP lesson methods and shared samples of student work. Between meetings, teachers and researchers communicated via a closed electronic list. Teachers also wrote early reflections and evaluations.

Each intervention teacher implemented SIOP lessons in his or her designated course and was videotaped three times per year—fall, winter, and spring. We used the SIOP rating scale to analyze these videotaped lessons, as well as other classroom observations that we conducted. We shared the analyses with teachers on an ongoing basis to facilitate teacher growth and validate the research interpretations. We provided written feedback to the teachers. We subsequently analyzed SIOP data collected throughout the project to determine overall teacher change and significant development in specific areas of instructional practice (Short & Echevarria, 1999).

We selected comparison teachers to match the intervention teachers as closely as possible on the basis of their schools' demographics and setting, student demographics in their sheltered classes, and professional training with ELLs. The comparison teachers had credentials and teaching experience similar to the intervention project teachers (see Table 3) but did not participate in any SIOP training. Each of the comparison teachers was certified to teach ELLs, as was all but one of the intervention teachers. In addition, the ethnic background of the comparison teachers was similar to that of the intervention teachers (see Table 4). On the West Coast, the teachers used SDAIE techniques to teach content areas; on the East Coast, all the teachers (except one intervention teacher) were High Intensity Language Teaching (HILT)-trained teachers. Teaching experience for the intervention teachers ranged from 1–15 years and from 3–14 years for the comparison teachers.

We videotaped the comparison teachers and rated them on the SIOP two times per year—fall and spring. They did not receive feedback on their lessons or participate in any training. The ratings of all teachers provided important data on the fidelity and quality of SIOP model implementation.

Table 3 shows that project teachers had years of experience teaching ELLs that were similar to those of compar-
sion teachers, but their SIOP scores were considerably different. Teachers certified to teach ELLs, like the teachers in our study, are familiar with effective strategies and techniques that make sheltered lessons understandable for their students. However, the difference in mean scores between the intervention and comparison teachers' implementation of SI may be attributed to the lack of a model to guide the comparison teachers' instruction. Although some of the

<table>
<thead>
<tr>
<th>Teacher</th>
<th>Mean years teaching</th>
<th>Mean SIOP score (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SIOP project</td>
<td>9</td>
<td>77</td>
</tr>
<tr>
<td>Comparison</td>
<td>4</td>
<td>51</td>
</tr>
</tbody>
</table>

*Note. SI = sheltered instruction; SIOP = Sheltered Instruction Observation Protocol.*

<table>
<thead>
<tr>
<th>Race or ethnicity</th>
<th>SIOP teachers</th>
<th>Comparison teachers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Black</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>White</td>
<td>16</td>
<td>3</td>
</tr>
<tr>
<td>Hispanic</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

*Note. SIOP = Sheltered Instruction Observation Protocol.*
comparison teachers were rated highly on certain features, they did not exhibit consistent use of best practices for ELLs that the intervention teachers used, as reflected in the SIOP model.

Qualitative data retrieved from written feedback from observations, electronic discussion, teacher evaluations and reflections, as well as periodic documentation in journals elucidate the kinds of instructional practices that account for some of the difference in intervention and comparison classrooms.

A comparison of observational notes from a SIOP sheltered English class and a non-SIOP sheltered English class indicates that both teachers began their lessons with a content objective, “Today we are going to . . . .,” but the SIOP teacher explicitly added a language objective that students would learn about editing their work. Furthermore, the SIOP teacher received the highest score of “4” for pacing, whereas the comparison teacher received a “2” because he “speeds through sections of the materials he may be assuming they understand” (Researcher field notes, March 15, 1999). Under the component Comprehensible Input, the SIOP teacher received a score of “4” because he “uses a variety of words to explain a concept” (Researcher field notes, April 14, 1999). In contrast, the comparison teacher’s speech was somewhat inappropriate for students’ proficiency level; it received a score of “2.” To illustrate the professional development process that the SIOP English teacher experienced, his journal notes from March 26, 1999, stated,

The last two years have been very rewarding. I have the luxury of having looped with the ELL group. Their writing and critical thinking skills have increased. It has been easier because of SIOP. Again, it has helped to focus the lessons in order to set objectives and methods which I teach.

Although teachers in the comparison group were competent, experienced teachers of ELLs, they did not have the specific instructional goals that the SIOP project afforded (i.e., to systematically examine features of instruction, highlight academic language needs of students, and work to improve practice in discreet ways).

Student Outcome Measures and Data Collection

To measure students’ academic literacy development over time, we used an expository writing assessment. We needed to use a measure that approximated the type of academic task that ELLs are regularly asked to perform in standards-based classrooms. Therefore, students wrote to a prompt that resembled typical writing tasks in content-area classes, in this case, social studies. We field tested the prompts in a school with a similar student body but in a nonparticipating district. For the pre- and posttest assessments, students wrote to the same prompt.

We had expected originally to use standardized test scores as an additional measure of student achievement. However, because the study preceded NCLB, most of the ELLs in the districts were exempted from the districts’ testing process of using norm-referenced standardized tests. So, because of limitations in the education policies at that time, we could not rely on reading or writing test scores of that nature. Furthermore, the district-developed writing assessments that were given to the students for promotion within the district ESL programs called for narrative rather than expository text, thus they were less suited to the content-area classes being studied. Our goal was to investigate ELL students’ performance on tasks that are critical to academic success; the writing assessment that we used was an important measure of academic literacy.

Given the limitations of this study, we used the writing assessment from the Illinois Measurement of Annual Growth in English (IMAGE) test. The IMAGE is the standardized test of reading and writing that the state administrators use to measure the annual growth of these skills for ELLs in Grades 3 and higher. The test is valid and reliable and has been correlated to and has predicted scores on the IGAP (the former Illinois test of achievement) that was administered to all students in Illinois, except those exempted because of linguistic development issues or learning disabilities.

Data collection for the pretest occurred within the first 6 weeks of class. The posttest assessment occurred within the last 6 weeks of the course. Seventy percent (n = 241) of all students in classes with SIOP-trained teachers were present for the pretest and posttest administrations. Eighty-two percent (n = 77) of the students in comparison classes were present for the pretest and posttest.

The testing procedures were similar to, but slightly modified from, those used in Illinois with the IMAGE (www.isbe.net/assessment/IMAGE.htm) to accommodate the schedules of participating schools. The test was untimed, whereas the IMAGE is normally given during two 40-min sessions. Most of the intervention and comparison students completed the assessment within a class period of 40–50 min. As with typical IMAGE administration, participants could draft their responses with a graphic organizer before writing a final product. Participants responded to only one prompt (Figure 2); for the IMAGE, they respond to four prompts. The researchers (authors and their research assistants) administered the test and monitored the classrooms during testing.

All of the pre- and posttest writing samples from the intervention and comparison classes were evaluated by an independent rater who had 2 years of experience scoring writing assessments for a school district and who was a former ESL teacher. The rater was blind to the conditions of the study. Because the IMAGE was unfamiliar to personnel in all districts, we opted for one rater who could learn about the test’s specifications and scoring rubric and thus maintain interrater reliability. The researchers randomized writing responses and essays of both groups of students and sent them to the rater for scoring.

Over 640 writing samples were scored with the IMAGE writing rubric. The rubric has a 6-point scale and provides
individual subscores for five dimensions of writing—language production, focus, support or elaboration, organization, and mechanics—as well as an overall score for each student. For each subtest, a student may have a score from 0 to 5; the total maximum score is 25 points. The five writing dimensions measure the following items:

1. Language production. Degree to which English-language acquisition is demonstrated in the written passage.
2. Focus. Degree to which the main idea is clear and maintained.
3. Support or elaboration. Extent that the main idea is explained or elaborated with specific evidence and supporting details.
4. Organization. Extent that the flow of ideas is logical and the text is connected.

Results

Table 5 shows the mean, standard deviation, and sample size for the performance of each group on the pre- and posttest writing samples. Each of the five IMAGE subtest scores appears after the total score is listed. The total score analysis was based on 315 students—those who had a pretest and posttest score for every subtest; however, subtest analyses were based on 318 students who had a score for each particular subtest, a slightly larger group for some subtests. Students in the SIOP classes performed less well on all pretests when compared with the comparison students, but they performed better on the posttests (see Table 6).

Because of differences between the two groups in their pretest scores prior to implementing the SIOP model, we conducted analyses of covariance (ANCOVAs) to determine whether students from the intervention classes made greater gains in writing than did students in the comparison classes. We calculated the adjusted means for the posttest results on the basis of the pretest covariate; Table 7 shows the ANCOVA results.

As indicated in Table 7, we found significant main effects for the intervention condition on the total score and on three of the five subtest scales. Comparisons between intervention and comparison groups on the total scores (i.e., aggregated across the five scales) indicated that the participants whose teachers were trained in the SIOP model made significantly better gains in writing than did the comparison group, $F(1, 312) = 10.79, p < .05$. In follow-up analyses on student performance on the writing-assessment subtests, we found that the intervention group performed at a significantly higher level in language production, $F(1, 314) = 5.00, p < .05$, organization, $F(1, 315) = 5.65, p < .05$, and mechanics, $F(1, 315) = 4.10, p < .05$, than did those in the comparison group, whose teachers had not received the research-developed training and support to deliver SI. The intervention group did not make statistically significant gains over the comparison group in their performance on the writing focus and elaboration subtests.

The effect size of the intervention, calculated for the intervention group as $d = \frac{M_1 - M_2}{\sigma_{\text{pooled}}}$ with $M_1 =$ posttest total writing and $M_2 =$ pretest total writing scores, was +.833. That effect size is considered large by most indexes; however, it may overestimate the effects of the intervention because some growth is expected as a result of maturation.

Discussion and Conclusions

The results of this study reveal positive effects of the SIOP model on student literacy achievement as measured with the IMAGE writing assessment. The intervention group scored lower on the total score and on all subtests of the pretest but higher on the posttest than did the comparison group. Thus, the intervention group made greater gains during the school year. The intervention group gained an average of 2.9 points (out of 25 points) between pretest and posttest administrations compared with a gain

| Rivers are important in the lives of people and in the buildings of towns and cities. Look at the two pictures and think about the rivers, the land, and the people. Directions: Write an essay to tell what is the same and what is different about these two pictures. Discuss how people use the river and how they use the land. Before you write the essay, you may want to take notes in a chart below. Then write your essay comparing the two pictures. |
| --- | --- |
| **Same** | **Different** |

FIGURE 2. Writing type: Expository—Comparison or contrast; Topic: Rivers and land use.
of 0.7 points for the comparison group, according to posttest mean scores adjusted for pretest performance (Table 6). The effect size of the intervention ($d = +.833$), estimated from the posttest gains made by the intervention group, suggest significant gains in students’ overall writing performance. The differences between the intervention and comparison groups were statistically significant for the total score and the language production, organization, and mechanics subtests. Those results indicate that the SIOP model offers a promising approach for helping ELLs develop academic literacy skills needed for success in school, in this case, academic writing.

In this study, we offer a preliminary examination of the SIOP model. Further research is needed on the effects of the SIOP model on reading and other content-area test scores. In the future, as we learn more about effective professional development for teachers on the SIOP model, researchers can design a scale-up study in terms of an experimental quantitative evaluation (e.g., with randomized student and teacher assignment and larger subject sets for greater effect size).

If no child is to be left behind in school, regardless of English proficiency level or academic background, substantial changes must be made in the way that ELLs are educated. The goal of high academic standards for all students is laudable, but the way to accomplish that goal must be reexamined because the achievement of ELLs is poor at

### TABLE 5. Mean Scores, Standard Deviations, and Sample Size for Intervention and Comparison Groups

<table>
<thead>
<tr>
<th>Variable</th>
<th>Intervention group</th>
<th></th>
<th></th>
<th>Comparison group</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pretest</td>
<td>Posttest</td>
<td>Pretest</td>
<td>Posttest</td>
<td>Pretest</td>
<td>Posttest</td>
</tr>
<tr>
<td>Total score</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$M$</td>
<td>13.55</td>
<td>16.36</td>
<td>14.61</td>
<td>15.81</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$SD$</td>
<td>3.42</td>
<td>3.33</td>
<td>3.36</td>
<td>3.45</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$N$</td>
<td>238</td>
<td>238</td>
<td>77</td>
<td>77</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Language production</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$M$</td>
<td>2.65</td>
<td>3.22</td>
<td>2.77</td>
<td>3.09</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$SD$</td>
<td>.78</td>
<td>.79</td>
<td>.78</td>
<td>.73</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$n$</td>
<td>240</td>
<td>240</td>
<td>77</td>
<td>77</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Focus</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$M$</td>
<td>2.81</td>
<td>3.30</td>
<td>3.01</td>
<td>3.17</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$SD$</td>
<td>.87</td>
<td>.98</td>
<td>.88</td>
<td>.94</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$n$</td>
<td>239</td>
<td>239</td>
<td>77</td>
<td>77</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Support and elaboration</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$M$</td>
<td>2.65</td>
<td>3.26</td>
<td>2.83</td>
<td>3.18</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$SD$</td>
<td>.78</td>
<td>.72</td>
<td>.70</td>
<td>.81</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$n$</td>
<td>241</td>
<td>241</td>
<td>77</td>
<td>77</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Organization</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$M$</td>
<td>2.77</td>
<td>3.31</td>
<td>3.16</td>
<td>3.21</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$SD$</td>
<td>.96</td>
<td>.78</td>
<td>.92</td>
<td>.71</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$n$</td>
<td>241</td>
<td>241</td>
<td>77</td>
<td>77</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mechanics</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$M$</td>
<td>2.72</td>
<td>3.28</td>
<td>2.84</td>
<td>3.17</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$SD$</td>
<td>.88</td>
<td>.87</td>
<td>.86</td>
<td>.94</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$n$</td>
<td>241</td>
<td>241</td>
<td>77</td>
<td>77</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### TABLE 6. Adjusted Means for Posttest Results

<table>
<thead>
<tr>
<th>Variable</th>
<th>Intervention</th>
<th>Comparison</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total score</td>
<td>16.52</td>
<td>15.34</td>
</tr>
<tr>
<td>Language production</td>
<td>3.24</td>
<td>3.04</td>
</tr>
<tr>
<td>Focus</td>
<td>3.32</td>
<td>3.10</td>
</tr>
<tr>
<td>Support and elaboration</td>
<td>3.28</td>
<td>3.13</td>
</tr>
<tr>
<td>Organization</td>
<td>3.34</td>
<td>3.12</td>
</tr>
<tr>
<td>Mechanics</td>
<td>3.30</td>
<td>3.11</td>
</tr>
</tbody>
</table>

### TABLE 7. Analysis of Covariance of Posttest Writing Results, by Intervention Condition

<table>
<thead>
<tr>
<th>Variable</th>
<th>$M^2$</th>
<th>$F$ ratio</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total score</td>
<td>78.276</td>
<td>10.785</td>
<td>.001*</td>
</tr>
<tr>
<td>Language production</td>
<td>2.133</td>
<td>5.004</td>
<td>.026*</td>
</tr>
<tr>
<td>Focus</td>
<td>2.904</td>
<td>3.706</td>
<td>.055</td>
</tr>
<tr>
<td>Support and elaboration</td>
<td>1.247</td>
<td>2.680</td>
<td>1.03</td>
</tr>
<tr>
<td>Organization</td>
<td>2.842</td>
<td>5.651</td>
<td>.018*</td>
</tr>
<tr>
<td>Mechanics</td>
<td>2.065</td>
<td>4.101</td>
<td>.044*</td>
</tr>
</tbody>
</table>

*Note. Pretest scores for each measure served as the covariate for posttest dependent measures. $df = 1$ for all variables. $^{*}p < .05.$
present (Olson, 2003). The impact of the status quo may be profound on our schools and society as the growth in numbers of ELLs continues to outpace the number of students who enter school with requisite background knowledge and fluent English proficiency (Fillmore & Snow, 2002).

We determined that certain features must be present in instruction so that content concepts are made comprehensible at the same time that academic English-language development is promoted (Echevarria et al., 2000, 2004). Deep understanding of the subject matter being taught, although necessary, is not sufficient when one works with ELLs. Effective SI, therefore, draws from and complements high-quality instructional methods advocated for regular classrooms but adds specific strategies for developing English-language skills. That is a shift in the teaching–learning relationship that requires teachers to engage students in listening, speaking, reading, and writing about the content in meaningful ways. In highly effective SI classrooms, explicit language instruction targeted to and slightly beyond students’ level of English proficiency also is presented in every lesson.

SI provides useful strategies and techniques for making instruction more understandable for ELLs, but without a scientifically validated model to guide teachers in lesson planning and lesson delivery, it will not be consistent within and across classrooms. For change to occur in teaching practice that leads to improved academic performance by students, teachers cannot simply select their favorite techniques; implementation of high-quality instruction must be systematic and steered by research. Without a sheltered model, language development is apt to be disregarded as a result of the pressure that teachers face to cover the curriculum (Short, 2002). Teachers report that having a well-articulated model of high-quality instruction empowers them to work more effectively with ELLs (Hulquist, 2002; Short & Echevarria, 1999).

Our research has revealed that when certain features of instruction were consistently and systematically used with ELLs through the SIOP model, their performance in expository writing improved significantly over that of similar students whose teachers were trained in teaching ELLs but did not implement the features found in the model. We believe that finding is especially important because expository writing is one of the most challenging academic areas for ELL students, and it is the type of writing that is fundamental to academic literacy.

As changes in the allocation of federal funds for ELLs and economically disadvantaged students bring increased attention to the achievement of all students, including ELLs, educators need to examine the interaction between the SIOP model, teacher decision making, implementation procedures, settings, student populations, and other variables. Although the model is effective, it is not a panacea for the challenge of helping ELL students meet high academic standards. As researchers who endeavor to conduct research in “messy settings” have stated, “Even useable knowledge will not make complex educational problems simple” (The Design-Based Research Collective, 2003, p. 7).

REFERENCES
Fillmore, L. W., & Snow, C. (2002). What teachers need to know about


The Sheltered Instruction Observation Protocol (SIOP)  
(Echevarria, Vogt, & Short, 2000, 2004)

Observer: _______________________________ Teacher: _______________________________
Date: _______________________________ School: _____________________________________
Grade: _______________________________ ESL level: ___________________________________
Class: _______________________________ Lesson: Multi-day Single-day (circle one)

Directions:  
Circle the number that best reflects what you observe in a sheltered lesson. You may give a score from 0–4.
Cite under “Comments” specific examples of the behaviors observed.

Total Score: [ ]  %Score [ ]  Tape #:___________

<table>
<thead>
<tr>
<th>Highly Evident</th>
<th>Somewhat Evident</th>
<th>Not Evident</th>
<th>NA</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

Preparation
1. Clearly defined content objectives for students [ ] [ ] [ ] [ ] [ ]
2. Clearly defined language objectives for students [ ] [ ] [ ] [ ] [ ]
3. Content concepts appropriate for age and educational background level of students [ ] [ ] [ ] [ ] [ ]
4. Supplementary materials used to a high degree, making the lesson clear and meaningful (graphs, models, visuals) [ ] [ ] [ ] [ ] [ ]
5. Adaptation of content (e.g., text, assignment) to all levels of student proficiency [ ] [ ] [ ] [ ] [ ]
6. Meaningful activities that integrate lesson concepts (e.g., surveys, letter writing, simulations, constructing models) with language practice opportunities for reading, writing, listening, and/or speaking [ ] [ ] [ ] [ ] [ ]

Instruction
Building Background
7. Concepts explicitly linked to students’ background experiences [ ] [ ] [ ] [ ] [ ]
8. Links explicitly made between past learning and new concepts [ ] [ ] [ ] [ ] [ ]
9. Key vocabulary emphasized (e.g., introduced, written, repeated and highlighted for students to see) [ ] [ ] [ ] [ ] [ ]

Comprehensible Input
10. Speech appropriate for students’ proficiency level (e.g., slower rate, enunciation and simple sentence structure for beginners) [ ] [ ] [ ] [ ] [ ]
11. Explanation of academic tasks clear [ ] [ ] [ ] [ ] [ ]
12. Uses a variety of techniques to make content concepts clear (e.g., modeling, visuals, hands-on activities, demonstrations, gestures, body language) [ ] [ ] [ ] [ ] [ ]

Strategies
13. Provides ample opportunities for student to use strategies [ ] [ ] [ ] [ ] [ ]
14. Consistent use of scaffolding techniques throughout lesson, assisting and supporting student understanding such as think-alouds (see Glossary) [ ] [ ] [ ] [ ] [ ]
15. Teacher uses a variety of question types throughout the lesson including those that promote higher-order thinking skills (e.g., literal, analytical, and interpretive questions) [ ] [ ] [ ] [ ] [ ]

Comments: (appendix continues)
### Interaction

16. Frequent opportunities for *interactions* and discussion between teacher/students and among students, which encourage elaborated responses about lesson concepts

17. *Grouping configurations* support language and content objectives of the lesson (see Glossary)

18. Consistently provides sufficient *wait time for student response*

19. Ample opportunities for students to *clarify key concepts in L1* (see Glossary)

### Practice/Application

20. Provides *hands-on* materials and/or manipulatives for students to practice using new content knowledge

21. Provides activities for students to *apply content and language knowledge* in the classroom

22. Uses activities that integrate all *language skills* (i.e., reading, writing, listening, and speaking)

### Lesson Delivery

23. *Content objectives* clearly supported by lesson delivery

24. *Language objectives* clearly supported by lesson delivery

25. Students engaged approximately 90–100% of the period (see Glossary)

26. *Pacing* of the lesson appropriate to the students’ ability level

### Review/Assessment

27. Comprehensive *review* of key vocabulary

28. Comprehensive *review* of key content concepts

29. Regularly provides feedback to students on their output (e.g., language, content, work)

30. Conducts *assessment* of student comprehension and learning of all lesson objectives (e.g., spot checking, group response) throughout the lesson (see Glossary)

---


Newton, MA: Allyn & Bacon.
Copyright of Journal of Educational Research is the property of Heldref Publications and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.