Evaluation of Quality Teaching for English Learners (QTEL) Professional Development

ies NATIONAL CENTER FOR
EDUCATION EVALUATION
AND REGIONAL ASSISTANCE

Institute of Education Sciences

Evaluation of Quality Teaching for English Learners (QTEL) Professional Development

Final Report

March 2012

Authors:

Johannes M. Bos, Principal Investigator
Berkeley Policy Associates

Raquel C. Sanchez, Project Director
Berkeley Policy Associates

Fannie Tseng, Principal Analyst
Berkeley Policy Associates

Nada Rayyes, Senior Analyst
Berkeley Policy Associates

Lorena Ortiz, Analyst
Berkeley Policy Associates

Castle Sinicrope, Analyst
Berkeley Policy Associates

This report is available on the Institute of Education Sciences website at http://ncee.ed.gov and the Regional Educational Laboratory Program website at http://edlabs.ed.gov.

Alternate Formats Upon request, this report is available in alternate formats, such as Braille, large print, audiotape, or computer diskette. For more information, please contact the Department's Alternate Format Center at 202-260-9895 or 202-205-8113.

Disclosure of potential conflict of interest

Regional Educational Laboratory West, housed at WestEd, contracted with Berkeley Policy Associates to conduct a third-party evaluation of the Program for Quality Teaching for English Learners, a WestEd intervention. None of the authors or other staff involved in the study from Berkeley Policy Associates and its subcontractors, American Institutes for Research, or any members of the Technical Work Group for the study, has financial interests that could be affected by the content of this report. The evaluation was conducted independent of WestEd staff, who developed and implemented the Quality Teaching for English Learners program.[1]

---

[1] Contractors carrying out research and evaluation projects for IES frequently need to obtain expert advice and technical assistance from individuals and entities whose other professional work may not be entirely independent of or separable from the tasks they are carrying out for the IES contractor. Contractors endeavor not to put such individuals or entities in positions in which they could bias the analysis and reporting of results, and their potential conflicts of interest are disclosed.

Acknowledgments

Contents

Figures

Tables

Executive summary

English language proficiency is critical to academic achievement in the United States. For several decades, educators and policymakers have explored strategies to ensure that English language learner students have access to rigorous academic content as much as non–English language learner students. Educating English language learner students is a challenge that has become a civil rights issue and a topic of federal legislation. In 1968, Congress passed the Bilingual Education Act, followed by the Equal Education Opportunity Act of 1974, which requires school districts to remove language barriers to instructional programming. More recently, the No Child Left Behind Act of 2001 mandated that all students, including English language learner students, demonstrate proficiency in English language arts and mathematics by 2014 (Abedi and Dietel 2004). This focus on academic success for all student subgroups is a priority in the Obama administration's *A Blueprint for Reform: The Reauthorization of the Elementary and Secondary Education Act* (U.S. Department of Education 2010).

The demographics of students in the United States have shifted significantly over the last few decades. The population has become increasingly diverse over the last 40 years in terms of race, ethnicity and linguistic background. According to U.S. Census figures, the proportion of children of immigrants among the school-age population grew from 6 percent in 1970 to 19 percent in 2000 (Capps et al. 2005). And the nation's overall K–12 school population grew less than 3 percent from 1995 to 2005, but the population of English language learner students increased 56 percent during that period (Batalova, Fix, and Murray 2007). It is important to note that the English language learner population includes students who are immigrants as well as U.S. born citizens who speak a language other than English at home.

The English language learner student population is growing, but its academic success is not (Working Group on ELL Policy 2009, 2010). Approximately half of English language learner students nationwide leave high school without a diploma (Hopstock and Stephenson 2003), compared with 11 percent for students overall (National Center for Education Statistics 2002). The No Child Left Behind Act of 2001 requires all students to master the same curriculum, regardless of their baseline English language proficiency; however, the diverse levels of academic preparation students bring to the classroom create complex pedagogical challenges for secondary teachers (August and Hakuta 1997; Parsad, Lewis, and Farris 2001; Ruiz-de-Velasco and Fix 2000; Walqui and van Lier 2010).

Overview of Quality Teaching for English Learners

Quality Teaching for English Learners (QTEL), developed by WestEd, is an approach to improving the teaching of English language learner students at the secondary level. It aligns with the broader democratic goals of equal access and equal opportunity for all students (Walqui and van Lier 2010). QTEL targets the teachers of English language learner students classified as limited English proficient and those reclassified as fluent English proficient and placed in mainstream classrooms. By enhancing the ability of teachers to work with English language learner students, the intervention also seeks to increase the quality of instruction for all other students in the mainstream classroom.

QTEL is a nonscripted intervention tailored to the needs of particular schools, teachers, and students. Five fundamental principles guide all QTEL activities (Walqui and van Lier 2010):

1. Sustain academic rigor.
2. Hold high expectations.
3. Engage in quality interactions.
4. Sustain a language focus.
5. Develop a quality curriculum.

These principles permeate the three core components of QTEL: summer institutes, individualized teacher coaching, and collaborative lesson design meetings.

The summer institutes consist of seven days of professional development group sessions to provide a foundation for using new tools and processes for the academic and linguistic development of adolescent English language learner students. To promote continuity across school years, three days are offered at the end of a school year (June/July) and four days are offered before the start of the next (August/September).

Four to six cycles of individualized coaching are offered to teachers participating in QTEL each year. Coaches help teachers develop academically and linguistically rigorous lessons that implement QTEL principles, tools, and processes. These coaching cycles consist of a one-on-one lesson design meeting, an observation of the lesson's implementation, and a debriefing.

The collaborative lesson design meetings, a series of monthly planning sessions, are held at the school sites and facilitated by QTEL staff to provide support for QTEL implementation.

The present study

The goal of this study was to determine whether QTEL is effective in improving academic outcomes for English language learners in U.S. middle schools. The results intend to inform policy decisions on professional development for teachers of English language learner students.[2]

The study is a school-level randomized controlled trial to test the effectiveness of QTEL using an intent-to-treat model. This model tests the effectiveness of offering an intervention rather than that of participating in it. Teachers eligible for the intervention included those of English language arts and English language development in the schools assigned to the intervention group. Teachers in the control group participated in non-QTEL professional development, as if not involved in the study. These professional development activities, and other implementation contexts, are discussed in chapter 3.

From a sample of 52 middle schools, Berkeley Policy Associates randomly selected 26 for the intervention group. Teachers in intervention schools were offered QTEL; teachers in control schools were not. The study team estimated QTEL's effects on student outcomes in English language arts and English language development, as measured by the California Standards Test

---

[2] In this study, *English language learner* includes students classified as limited English proficient and those classified as such in the past but have been reclassified as fluent English proficient.

(CST-ELA) and the California English Language Development Test (CELDT). The current study measured secondary impacts on teacher knowledge, attitudes, and practice. The sample included middle schools in urban and suburban areas of three Southern California counties, the English language arts and English language development teachers in these schools, and their students. The sample teachers and students included those who moved into the schools during the study. The study was conducted for the 2007/08, 2008/09, and 2009/10 school years.

Six primary research questions focused on students:

1. What is the impact of QTEL on students' standardized test scores in English language arts among all grade 8 students attending intervention schools at the end of Year 3 (2009/10)?

2. What is the impact of QTEL on students' standardized test scores in English language arts among all grade 7 students attending intervention schools at the end of Year 3 (2009/10)?

3. What is the impact of QTEL on English language learner students' standardized test scores in English language arts among all grade 8 English language learner students attending intervention schools at the end of Year 3 (2009/10)?

4. What is the impact of QTEL on English language learner students' standardized test scores in English language arts among all grade 7 English language learner students attending intervention schools at the end of Year 3 (2009/10)?

5. Focusing on the subgroup of English language learner students who were classified as limited English proficient in 7th grade in study Year 2 (2008/09) and who were still in intervention schools and took the CELDT in 8th grade (in the fall of study Year 3, 2009/10), what is the impact of QTEL on standardized test scores in English language proficiency (i.e., on 8th grade CELDT scores)?

6. Within the subgroup of English language learner students who were classified as limited English proficient in 6th grade in study Year 2 (2008/09) and who were still in intervention schools and took the CELDT in 7th grade (in the fall of study Year 3, 2009/10), what is the impact of QTEL on standardized test scores in English language proficiency (i.e., on 7th grade CELDT scores)?

Three secondary research questions focused on teachers:

1. What is the impact of QTEL on teacher instructional knowledge?

2. What is the impact of QTEL on teacher attitudes toward English language learner students?

3. What is the impact of QTEL on teacher practice, as measured by the Sheltered Instruction Observation Protocol (SIOP)?

Sample characteristics

The analyses included test scores of students in grades 6, 7, and 8. These data included students who moved into the schools after the study started. The study began with test scores for 6,382 grade 6 intervention students and 6,000 grade 6 control students in Year 1. In Year 2, the sample included test scores for 9,230 grade 7 intervention students and 8,272 control students. In Year 3, the sample included test scores for 9,555 grade 8 intervention students and 8,625 grade 8 control students. Attrition occurred for both study groups; however, it was not statistically significantly different between the intervention and control groups.

For the teacher sample, the study began with 338 eligible (English language arts or English language development) teachers in the intervention group and 303 in the control group.[3] Teachers who moved into the target grades after the study started were included. Teacher mobility in and out of districts over the three years resulted in overall attrition of 41 percent for the intervention group and 38 percent for the control group; the differences in teacher mobility were not statistically significant.

Data collection and analysis

The data included student standardized test scores and various teacher measures. To examine student outcomes for the primary research questions, California Standards Test for English Language Arts (CST-ELA) and California English Language Development Test (CELDT) scores were collected for students in grades 6, 7, and 8 in 2007/08, 2008/09, and 2009/10 (see chapter 2). To address teacher outcomes, a teacher survey was administered in the spring of 2008, 2009, and 2010. A teacher knowledge test was administered at these same times, but an administrative error precluded a link from Year 1 tests to teacher identifiers; thus, only data for Year 2 and Year 3 could be analyzed. The Sheltered Instruction Observation Protocol (SIOP) was used to rate classroom observations of teachers in spring 2007, 2008, 2009, and 2010. The SIOP is a measure of teacher performance designed to rate teachers on the delivery of instruction to English language learners in K–12 settings. It measures teacher performance along eight dimensions: lesson preparation; building background; comprehensible input; strategies; interaction ; practice/application; effectiveness of lesson delivery; and lesson review/assessment (Echevarria, Vogt, and Short 2004).

QTEL implementation for this study varied from the intended design. To accommodate 26 middle schools, QTEL coaches focused on one department (English language arts) and targeted only English language arts and English language development teachers. Usually, QTEL is implemented as a schoolwide intervention in which coaches work with teachers from all content areas, including mathematics, social studies, and science. Also, to reach the needed number of schools, implementation was staggered—focusing on grade 6 teachers in Year 1, grade 7 teachers in Year 2, and grade 8 teachers in Year 3. Ideally, QTEL coaches would work with the same teachers each year.

Various contextual factors also caused QTEL implementation to vary from the intended design. Budget crises and teacher layoffs were the biggest challenges to consistent implementation.

---

[3] The numbers of teachers in this paragraph refer to teachers of grades 6, 7, and 8 pooled together.

Non-QTEL-related reforms, changing academic calendars, professional culture, principal leadership, and teacher buy-in also contributed. Teacher participation in QTEL was generally low, and missing or incomplete implementation data contributed to an incomplete picture of implementation.

Impacts

No significant effects were found on student achievement, as measured by the CST-ELA, or on English language development, as measured by the CELDT. That is, there were no meaningful or significant differences in academic performance or language proficiency skills, as measured by these assessments, between the intervention students and the control students.

No significant effects were found on teacher attitudes, teacher knowledge, or teacher practice, as measured by the teacher survey, teacher knowledge assessment, and the SIOP, respectively.

Exploratory analysis

Exploratory analyses were conducted to more fully understand whether and how QTEL might have affected the intervention group. These analyses focused on the effects of QTEL on subgroups of teachers and students, additional outcomes (using a tool developed by Berkeley Policy Associates aligned with QTEL principles), and dose response.

Research questions for the effects of QTEL include:

1. Do impacts on students' CST-ELA scores vary by student English language learner status (English only, initially fluent English proficient, redesignated fluent English proficient, or limited English proficient)?

2. Do impacts on students' CST-ELA scores vary by teacher characteristics and baseline classroom quality?

   a. Does the impact of QTEL on student achievement vary by the level of experience of the teachers in their schools?

   b. Does the impact of QTEL on student achievement vary by the extent to which teachers in their schools have an advanced degree (master's or above)?

   c. Does the impact of QTEL on student achievement vary by the baseline quality of their school's English language arts and English language development classrooms?

Additional outcome questions:

3. Does QTEL improve areas of teacher practice expected to aligned with the QTEL program, as measured with the Program Aligned Classroom Observation (PACO) instrument?

4. Does QTEL improve different areas of teacher practice, as measured by subscales of the SIOP?

Dose-response analysis questions:

5. What is the potential effect on student achievement and other outcomes of extending teacher professional development to an additional 10 percent of teachers in middle schools in a district?

6. What is the potential effect on student achievement and other outcomes of extending intensive teacher professional development (at least seven days or more) to an additional 10 percent of teachers in middle schools in a district?

Results of exploratory analysis

No statistically significant impacts of QTEL were found on the English language arts achievement of any of the four English language learner status subgroups in either grade 7 or grade 8.

For teacher characteristics, QTEL was estimated to have increased the test scores of grade 8 students who were in schools where more than 43.5 percent of the teachers had an advanced degree. The difference of 10.40 points translated to an effect size of 0.17 standard deviation in these scores. This impact estimate was statistically significant ($p = .027$) and statistically significantly different from the estimated QTEL impact on the test scores of grade 8 students in schools with less highly educated teachers (effect size = 0.01, $p = .167$). No other impact estimates related to teacher characteristics were statistically significant or differed significantly across the subgroups.

The exploratory analysis of the four subscales of the PACO instrument resulted in one statistically significant positive impact estimate: the intervention–control difference on the *student interaction* subscale translated to an effect size of 0.445 ($p = .005$). This subscale measures the degree of academically oriented student–student interaction within the classroom. QTEL had no statistically significant impacts on any Sheltered Instruction Observation Protocol subscales.

Limitations

The internal validity of the findings is limited by the following:

1. Students and teachers left the schools between random assignment and when outcome data were collected. For example, as chapter 2 discussed, 29.5 percent of grade 6 students in 2007/08 were no longer in the grade 8 impact sample in 2009/10. For English language learner students, these grade 6–8 attrition rates differed between intervention schools (41.5 percent) and control schools (28.9 percent). If there are systematic differences between the expected outcomes of students and teachers who leave the intervention schools and those who leave the control schools, such differences would bias the resulting impact estimates.

2. Three schools (two intervention and one control) were consolidated during the study period. To maintain the integrity of random assignment, one-third of the teachers and students in the consolidated school were randomly selected and considered control group

members in the outcome analyses, even though they were treated as intervention group members by QTEL after the consolidation took place. As a result, any estimated QTEL impact on this school would likely be attenuated (biased toward zero), causing a small bias in the overall impact estimates across the entire study sample.

3. Teachers in the control group continued to have access to their regular professional development activities, as provided and prescribed by their school or district. The data on the control group are limited to survey responses from administrators and teacher self-reports. As a result, control teachers might have had undetected crossover exposure to QTEL or similar content.

4. Classroom observations were conducted using a convenience sample. The classrooms observed did not necessarily represent their schools and grades and observed classrooms in intervention schools might have differed from those in control schools, possibly biasing the findings based on these observations.

5. Nonresponse on teacher surveys might have caused nonresponse bias in estimates based on data from those surveys. That is, teachers who responded in intervention schools may have systematically differed from teachers who responded in control schools.

The external validity of the findings is limited by the following:

1. The sample of school districts and schools was not a random sample in the United States, California, or Southern California. There is no way to know whether the results generalize beyond it.

2. Schools and districts participating in the study volunteered for a study in which schools were randomized to receive QTEL or not to receive it. This means that the results may not be representative of schools that are fully committed to the QTEL intervention, in which case they may have avoided randomization and contracted for QTEL (or similar services) directly.

3. Participation in QTEL services was not universal. Most teachers eligible to participate did not receive all services as intended. Thus, the findings do not generalize to a setting in which all participants receive all intended services.

4. The fidelity of implementation was limited. The delivery of some QTEL services was compromised by school staffing and logistical issues, and tracking of service receipt by QTEL staff was limited. As a result, the findings do not generalize to a setting with complete implementation fidelity.

5. While the classroom observation instruments demonstrated acceptable levels of internal consistency[4] and interrater reliability, there is insufficient data available to establish external validity of either the SIOP or the PACO instrument. Prior research establishing the external validity of the SIOP was not available.

---

[4]We considered a Cronbach's alpha of 0.70 as the minimum acceptable for internal consistency for the classroom observation measures. The 30-item SIOP demonstrated a Cronbach's alpha of 0.94. The 22-item PACO demonstrated a Cronbach's alpha of 0.92.

Four data quality issues limit the reliability of the findings:

1. Program data on coaching and professional development attendance were incomplete and poorly documented. As a result, the description of implementation might be unreliable.

2. The main student outcome measures captured only part of the anticipated impact on student outcomes, because standardized test scores do not measure the full range of skills and competencies required for success in grade-level content area courses.

3. Baseline data on students and teachers were limited. Students entered grade 6 after random assignment. And no elementary school data were available to establish their baseline equivalence, both in grade 6 and after attrition in grades 7 and 8. Teacher baseline data were limited to classroom observations, which were not linked to individual teachers and could be used only to establish baseline equivalence at the school level.

4. Because teacher rosters were either unavailable or unreliable, individual students could not be linked to individual teachers. This prevented analyses of the direct relationship between a teacher's receiving QTEL services and that teacher's students' outcomes.

Implications and future research

This study exemplifies the challenges in examining the effectiveness of a nonscripted professional development intervention tailored to participant needs. The study tested a version of QTEL implemented under specific conditions. These analyses did not detect significant impacts on student achievement, teacher knowledge, teacher attitudes, or teacher practice. Future research on schoolwide implementation of the more complete QTEL model would be beneficial.

The exploratory analysis suggests that QTEL might have differential effects on the students of teachers with different levels of education. It also suggests that there might be a positive relationship between QTEL and the amount of student–student interaction within classrooms, as measured by an observation instrument aligned with the intervention. These findings point to the need for more research on how QTEL interacts with teachers' prior learning and how it affects teacher practice.

In-depth case studies may be useful for improving QTEL implementation and offering the field promising practices. It is also possible that interventions like QTEL take more than three years to show impact on the achievement of English language learner students. Longitudinal studies may provide information about these potential effects.

Chapter 1. Study background

English language proficiency is fundamental to academic achievement in the United States. Addressing the academic needs of English language learner students has been a major goal of U.S. education policymakers since the Civil Rights Movement. The Bilingual Education Act of 1968 was reinforced by the Equal Education Opportunity Act of 1974, which requires school districts to remove language barriers to instructional programming. More recently, the No Child Left Behind Act of 2001 has focused on the academic outcomes of all students, including English language learner students, requiring that they demonstrate proficiency in English language arts and mathematics by 2014 (Abedi and Dietel 2004). That focus remains a priority in the Obama administration's *A Blueprint for Reform: The Reauthorization of the Elementary and Secondary Education Act* (U.S. Department of Education 2010).

This study assesses the impacts of Quality Teaching for English Learners (QTEL) professional development on the academic performance of middle school students. The goal of QTEL, developed by WestEd, is to increase the capacity of secondary-level content area teachers to simultaneously address the language needs and academic needs of English language learner students within mainstream secondary school settings.

Regional Educational Laboratory West (REL West) contracted with Berkeley Policy Associates, an employee-owned small business with more than 35 years in social policy research and evaluation under federal contract, to conduct this independent evaluation. Berkeley Policy Associates designed the study, collected and analyzed the data, and wrote the report. REL West provided them with implementation data, such as teacher participation records and coaching logs, and REL West QTEL developers and coaches interviewed with researchers from Berkeley Policy Associates about specific implementation contexts, challenges, and successes.

The study was a school-level randomized controlled trial testing the effectiveness of QTEL using an intent-to-treat model.[5] From a sample of 52 middle schools, Berkeley Policy Associates randomly selected 26 for the intervention group. Teachers in intervention schools were offered QTEL; teachers in control schools were not. The study team estimated the intervention's effects on student learning in English language arts and English language development, as measured by the California Standards Test (CST-ELA) and the California English Language Development Test (CELDT). As detailed in chapter 2, the current study also measured secondary impacts on teacher knowledge, attitudes, and practice. The study sample included middle schools in urban and suburban areas of three Southern California counties, the English language arts and English language development teachers in these schools, and the students of those teachers. The study took place in the 2007/08, 2008/09, and 2009/10 school years.

Need for the study

The linguistic landscape of classrooms in the United States has changed in recent decades. According to U.S. Census figures, the share of children of immigrants among the school-age

---

[5] An intent-to-treat study measures the effects of making available or offering an intervention, not the effects of receiving or participating in it.

population grew from 6 percent in 1970 to 19 percent in 2000 (Capps et al. 2005). And the nation's overall K–12 population grew less than 3 percent overall from 1995 to 2005, but the population of English language learner students increased 56 percent during that period (Batalova, Fix, and Murray 2007).

However, while population growth of English language learner students is robust, academic success for these students is not (Working Group on ELL Policy 2009, 2010). Approximately half of English language learner students nationwide leave high school without a diploma (Hopstock and Stephenson 2003), compared with less than 11 percent for students overall (National Center for Education Statistics 2002). Current accountability standards established under the No Child Left Behind Act of 2001 require all students to master the same curriculum, regardless of their baseline English language proficiency. These requirements, as well as the diverse levels of academic preparation that students bring to high school classrooms, create complex pedagogical challenges for secondary teachers (August and Hakuta 1997; Parsad, Lewis, and Farris 2001; Ruiz-de-Velasco and Fix 2000; Walqui and van Lier 2010).

The traditional model of English as a second language instruction in U.S. secondary schools is to separate English language learner students from their native English-speaking counterparts— creating an English as a second language "track," where students must demonstrate a certain level of English language proficiency prior to admission into mainstream English language arts classes (Valdes 2001). This model rests on a structural view of language that separates language learning from academic content learning. The underlying premise is that a students' lack of English fluency is a deficiency to be remedied before they are allowed to tackle grade-level academic content (Clegg 1996). For example, as recently as 2009, Arizona instituted a policy that mandates a daily minimum of four hours of Structured English Immersion for all English language learner students until they perform at the advanced level on the state's assessment of English language proficiency. According to this policy, "[t]he primary determinant of the appropriate student grouping for [Structured English Immersion] classrooms is the English proficiency level of the students" and not their grade level or academic preparation (Arizona English Language Learners Task Force 2007, p. 4).

Moving away from approaches that separate English language learner students from their English-speaking peers, California schools and districts have implemented a variety of strategies to support English language development, including structural approaches to teaching language. However, these strategies typically lack a coherent theoretical base (Walqui and van Lier 2010). QTEL offers a different approach, targeting not only the English language learner students classified as limited English proficient but also those who have been reclassified as fluent English proficient and placed in mainstream classrooms.

The current study intends to determine if QTEL is an effective alternative to other approaches to teaching English language learner students, such as the variety of strategies used in the control schools, including the traditional structural approach. Teachers in control schools received a variety of professional development activities aimed at supporting English language learner students, but those activities were not delivered through the QTEL model.

It is particularly important to conduct QTEL research in the western United States because of the inconsistency among state policies within the region (such as the policies in Arizona and California discussed above) on how best to educate English language learner students. The

results of this study intend to inform policy decisions on professional development for teachers of English language learner students.

Previous research on QTEL

The QTEL developers and other WestEd researchers piloted the professional development model and materials for QTEL in a variety of settings. Expert panels reviewed the QTEL model, design, and materials (Farr 2006). A 2004/05 randomized controlled trial in New York City found that QTEL had substantial beneficial impacts on teachers' pedagogical content knowledge (Farr 2006). However, that study found few measurable impacts on instructional practices and teacher attitudes and no effect on student achievement in literacy. The study was limited by a short implementation period (four months). The authors recommended that QTEL be rigorously tested over an intervention period of at least two academic years, with improved measurement of instructional practice through classroom observations and multiple years of student achievement data (Farr 2006).

Overview of QTEL

Developers in WestEd's Teacher Professional Development Program were concerned that the structural approach to teaching English as a second language lowered expectations and simplified curricula for adolescent English language learner students. In addition, the lack of a coherent theoretical foundation for the strategies employed in California classrooms left teachers with no framework for evaluating the appropriateness of a strategy for a specific instructional goal.

WestEd designed QTEL to ground its professional development for teachers of English language learner students in a coherent theory of learning. The intervention prepares teachers to create lessons that develop students' potential through challenging academic content, rather than through mastering discrete and isolated linguistic skills as a prerequisite to academic content. By providing teachers with these skills, the developers intended to improve instruction for English language learner students and for all other students in the mainstream classroom. Because all students, regardless of their first language, need to develop academic language skills and vocabulary, the principles of QTEL were designed to improve student outcomes for all students, including English language learner students.

Based on Vygotsky's (1978) notion that learning precedes development, QTEL focuses on how teachers can plan challenging activities just beyond what a student can do independently. QTEL helps teachers provide instructional scaffolding so that students can participate in lessons and engage with grade-appropriate concepts in a new language. "In this view, deliberate, well-constructed teaching drives [academic and language] development" (Walqui and van Lier 2010, p. 7).

Drawing on theory and practice from research in sociolinguistics, cognitive psychology, and sociocultural learning, QTEL offers an academic framework rich in intellectual challenge combined with highly supported tasks designed to develop teacher expertise and student achievement for English language learner students within the mainstream secondary school program (Walqui and van Lier 2010).

According to Walqui and van Lier (2010), "[t]he definition of quality teaching must account for the many diverse ways of teaching that can address students' needs with excellence" (p. 81). Because all good teaching responds to the needs of individual students, the developers identify five fundamental principles that guide all QTEL activities:

1. Sustain academic rigor.

2. Hold high expectations.

3. Engage in quality interactions.

4. Sustain a language focus.

5. Develop a quality curriculum.

These five principles are an extension of the sociocultural theory "that student development is a consequence of (and not a prerequisite for) carefully planned opportunities for students to participate in meaningful and demanding academic activity with others and that learning is primarily a social and cultural, rather than individual, phenomenon" (p. 83). These fundamental principles are manifested through lesson goals and objectives and can guide teachers in lesson planning (table 1.1).

The professional development session components focus first on engaging teachers as participants in QTEL lessons so that they experience learning as the students would. Teachers are then guided through a study of that lesson to identify the essential components of the lesson and relate them back to the QTEL principles. The teachers then work in small groups to evaluate and adapt existing lessons and texts to incorporate the QTEL instructional tools and processes.

In the first module QTEL emphasizes the instructional practices teachers can employ to scaffold the reading of narrative and informational texts that are beyond the students' independent reading level. As teachers become more familiar with the structures and practices that underlie the QTEL lesson, they are expected to engage students in the reading and writing of increasingly complex informational and expository texts.

QTEL is delivered through three components[6] of professional development:

1. *Summer institutes.* QTEL staff offer seven days of summer professional development group sessions to build understanding and pedagogical knowledge to support implementation of new tools and processes for the academic and linguistic development of adolescent English language learner students. To promote continuity across school years, three days of group sessions are offered at the end of a school year (June/July) and four days are offered before the start of the next (August/September).

2. *Individualized coaching and in-classroom support.* During the school year, QTEL staff offer four to six individual coaching cycles, helping teachers develop academically and linguistically rigorous lessons that implement QTEL principles, tools, and processes. These coaching cycles consist of a one-on-one lesson design meeting, an observation of the lesson's implementation, and a debriefing.

---

[6] See chapter 3 for each component described in more detail.

3. *Lesson design meetings.* QTEL staff facilitate monthly collaborative planning sessions or study groups at the school sites to support implementation of QTEL tools and processes.

The QTEL developers designed these components in support of the idea that teacher expertise is developed over time and along a continuum—from apprenticeship to mastery—engaging teachers in reflective practice. Reflective practice involves the capacity to analyze one's own actions in the professional setting and to engage in a continual process of improvement. It involves the teacher's analysis of her own instructional behavior in the classroom and conscious decision making about what strategies and methods to attempt, refine or discard in future lessons (Walqui, 2000). Thus, within the QTEL model, reflective practice takes place within a variety of formats, including workshops, classroom coaching, lesson study, and collaborative planning. QTEL was designed to be integrated across all these formats and to encourage teachers to attend summer workshops, receive coaching, and participate in collaborative lesson design activities to create a professional culture in which teachers support each other to implement lessons that are consistent with the principles of QTEL (Walqui and van Lier 2010).

The research design did not include a rigorous process study of the program implementation. A general discussion of the program implementation and treatment contrast is provided in chapter 3 of this report.

Table 1.1. QTEL principles, goals, and objectives

Principles	Goals	Objectives
Sustain academic rigor	Promote deep disciplinary knowledge	• Develop central ideas in the discipline first, postponing interesting but secondary details • Establish interconnections among central ideas of the discipline • Deepen understanding of themes over time
	Engage students in generative disciplinary concepts and skills	• Have students anchor new knowledge to central concepts in order to build understanding • Have students apply familiar central ideas or strategies to their emerging understanding of new concepts • Invite students to build increasingly complex explanations of disciplinary concepts and processes
	Engage students in generative cognitive skills (higher order thinking)	• Have students combine facts and ideas to synthesize, evaluate, and generalize • Have students build arguments, solve problems, and construct new meanings and understandings
Hold high expectations	Engage students in tasks that provide high challenge and high support	• Provide students with activities that are academically challenging, but flexible enough to allow multiple entry points: all students, regardless of where they start, will benefit from participation • Scaffold students' ability to participate in activities • Ensure that students are asked to engage in increasingly more complex tasks • Treat students as if they already have the abilities you are seeking to develop

Principles	Goals	Objectives
	Engage students in the development of their own expertise	• Conduct metacognitive activities so that students gain knowledge of how to learn, how to monitor their progress, and how to self-correct • Provide practice in the use of academic tools and activities[a] so that students appropriate them over time • Encourage students to support each other in their development • Encourage students to support each other in building academic stamina
	Make criteria for quality work clear to all	• Use rubrics to spell out expected quality of work • Encourage students to take risks and to work hard to master challenging academic work
Engage students in quality interactions	Engage students in sustained interactions with teacher and peers	• Invite students to go beyond brief, single responses and to elaborate, illustrate, and connect to their interlocutors' ideas
	Focus interactions on the construction of knowledge	• State explicitly that constructing new understandings is hard work that requires listening intently to interlocutors, making sense of what they are saying, and deciding how to respond (either by agreeing and providing further evidence, or by disagreeing and stating why that is the case) • Ask students to focus on the coherence of what they are saying (Are they stating main ideas? Are they making sense?) and to deepen their understanding by making connections to related ideas
Sustain a language focus	Promote language learning in meaningful contexts	• Provide explicit examples (formulaic expressions) of how to mark agreement, disagreement, and other moves in response to an interlocutor or text
	Promote disciplinary language use	• Focus on the social purpose of genre, audience, structure, and specific language of disciplinary texts; have students practice deconstructing and creating similar texts
	Amplify rather than simplify communications	• Give rich and varied examples, looking at difficult concepts from several angles
	Address specific language issues judiciously	• Focus corrective feedback on EITHER fluency, complexity, OR accuracy (BUT not at the same time)[b]
Develop quality curriculum	Structure opportunities to scaffold learning, incorporating the above goals	• Set long-term goals and benchmarks • Use a problem-based approach with increasingly interrelated lessons • Use a spiraling progression[c] • Make connections between subject matter and students' reality • Build on students' lives and experiences

a. Examples of academic tools and activities in English language arts include reading grade-level narrative and expository texts that are written in the original, natural language of the authors (rather than texts that are adapted or modified for less proficient readers of English), taking notes, using appropriate reference materials, participating in critical discussion or debate, analyzing a character's motivations, etc.

b. Walqui and Van Lier (2010) warn against overwhelming language learners with "red ink" by trying to correct every flaw in their language at once. Instead they recommend that the feedback from the teacher "mirror" the focus of the assignment so that the student can connect the feedback to the main purpose of the lesson at hand (p. 74).

c. The concepts of scaffolding and a spiraling curriculum comes from the work of Jerome Bruner (1966, 1996) and refer to a teaching approach in which each subject or skill area is revisited at intervals, at a more sophisticated level each time.

Source: Adapted from Walqui and van Lier (2010, pp. 84–85).

Logic model for studying the impact of QTEL

QTEL's goal is to improve the academic outcomes of secondary students who are English language learner students or who share mainstream classrooms with those students. However, the professional development is a teacher-level intervention. The researchers therefore expected teacher-level outcomes to precede any student-level academic impacts. This hypothetical causal pathway is illustrated in figure 1.1.

Figure 1.1. Logic model
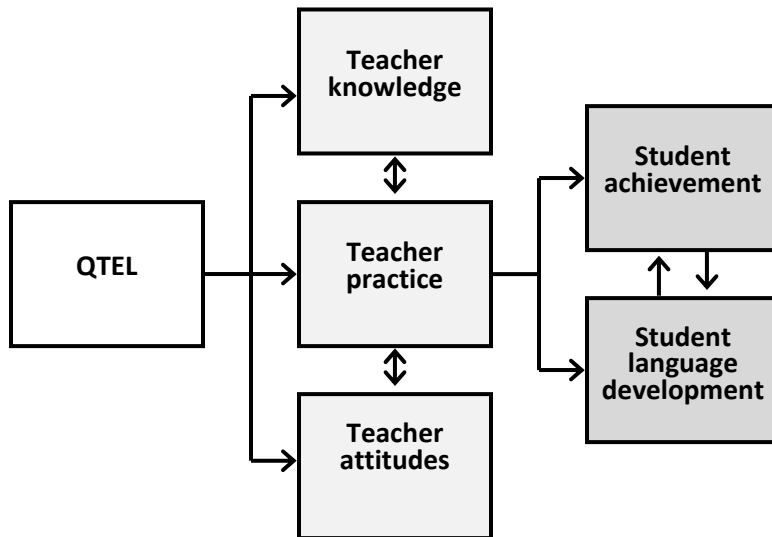
Source: Authors' construction.

According to the theory of action in figure 1.1, successful classroom implementation of QTEL principles depends on effective teachers who thoroughly understand the curriculum and how second languages are learned and developed. Teachers must believe their students can master the target language and curricular content. Teachers must also create and implement lesson activities that enable students at different levels of English proficiency (including those whose primary language is English) to meaningfully engage each other on the academic content. The box on the far left of the logic model represents QTEL professional development (summer institutes, individual coaching, and lesson design meetings), which is intended to stimulate changes at the teacher level. The center boxes represent three aspects of predicted change in teachers: knowledge about second language acquisition and pedagogy; classroom practice or use of sound pedagogy in the classroom; and attitudes toward English language learner students.

If teacher knowledge, belief, and practice change as intended, those changes would be expected to impact the two student outcomes in the boxes on the right: students would accelerate their language development and improve their academic achievement, as measured by state standardized tests. The bidirectional arrows between these student-level outcomes represent the hypothesized interdependence of student language development and student achievement, as described in the sociocultural theories of learning that are the foundation of the QTEL model (Walqui and van Lier 2010).

Research questions

To evaluate QTEL, the study team focused on two domains: student-level outcomes and teacher-level outcomes. Student achievement outcomes were the primary outcomes of interest. This primary focus acknowledges that for QTEL to be successful and replicable, it must improve the achievement of students in schools that implement it. As discussed in "Overview of QTEL," the QTEL model was developed to enhance teacher skills and improve the outcomes of all students, with anticipated impacts being more pronounced for English language learner students. Hence, research questions 1 through 6 focused on student achievement for all students whose teachers were offered QTEL, as well as for English language learner students in those classrooms.

The student-level research questions[7] are:

1. What is the impact of QTEL on students' standardized test scores in English language arts among all grade 8 students attending intervention schools at the end of Year 3 (2009/10)?

2. What is the impact of QTEL on students' standardized test scores in English language arts among all grade 7 students attending intervention schools at the end of Year 3 (2009/10)?

3. What is the impact of QTEL on English language learner students' standardized test scores in English language arts among all grade 8 English language learner students attending intervention schools at the end of Year 3 (2009/10)?

4. What is the impact of QTEL on English language learner students' standardized test scores in English language arts among all grade 7 English language learner students attending intervention schools at the end of Year 3 (2009/10)?

5. Focusing on the subgroup of English language learner students who were classified as limited English proficient in 7th grade in study Year 2 (2008/09) and who were still in intervention schools and took the CELDT in 8th grade (in the fall of study Year 3, 2009/10), what is the impact of QTEL on standardized test scores in English language proficiency (i.e., on 8th grade CELDT scores)?

6. Within the subgroup of English language learner students who were classified as limited English proficient in 6th grade in study Year 2 (2008/09) and who were still in intervention schools and took the CELDT in 7th grade (in the fall of study Year 3, 2009/10), what is the impact of QTEL on standardized test scores in English language proficiency (i.e., on 7th grade CELDT scores)?

Research questions 1, 3, and 5 focus on impacts on student outcomes in the year that teachers have access to the full QTEL; research questions 2, 4, and 6 ask whether QTEL has impacts on student outcomes in the year following the implementation year. Research questions 1 and 3 potentially capture three-year cumulative effects on students in grade 8, and research questions 2

---

[7] Having six confirmatory student achievement research questions may reduce the likelihood of finding significant impacts after adjusting for multiple comparisons (see chapter 2).

and 4 potentially capture two-year cumulative effects on students in grade 7. Because of the timing of the CELDT, research questions 5 and 6 capture only two-year and one-year effects on those outcomes, respectively. All these distinctions are discussed in greater detail in chapter 2.

As previously discussed, the theory of action linking QTEL professional development to students' achievement and language proficiency posits that the intervention affects several teacher-level outcomes first. A final set of confirmatory research questions focuses on the secondary teacher-level outcomes:

1. What is the impact of QTEL on teacher instructional knowledge?

2. What is the impact of QTEL on teacher attitudes toward English language learner students?

3. What is the impact of QTEL on teacher practice as measured by the Sheltered Instruction Observation Protocol (SIOP)?

In this study, teachers "exposure to QTEL" was defined as teachers having access to QTEL professional development (that is, teaching the target grade and content area in an intervention school). For students, it was defined as being assigned to a teacher in the intervention group for core English language arts or English language development courses. Target teachers[8] decided whether to participate in QTEL activities and whether to implement QTEL-style lessons. Teachers were encouraged by the QTEL coaches to participate in as many professional development activities as possible and to accept in-classroom support from the coaches.

Exploratory research questions on subgroup variation were also included. While the intervention's primary focus was improving the academic outcomes of English language learner students (those formally classified as limited English proficient and those who have been classified as such in the past), non–English language learner students were also included in the study impacts, estimated at the school level, to capture potential impacts for them as well.

Additional subgroup analyses disaggregated and compared impact findings across student, teacher, and school characteristics, including through a detailed breakdown of four different English language learner statuses among students; teacher certification; the aggregate baseline quality of classrooms in the school; and school-level student composition. Because the instructional needs of English language learner students vary, as do the teaching approaches and experience of teachers who work with them, these subgroup analyses determined the extent to which QTEL can succeed with different kinds of students and teachers and in different school settings. The exploratory research questions include:

1. Do impacts on students' CST-ELA scores vary by student English language learner status (English only, initially fluent English proficient, redesignated fluent English proficient, or limited English proficient)?

---

[8] Target teachers were English language arts and English language development teachers targeted in each year of the study. In Year 1, these teachers who taught grades 6 and 7 were offered the summer institutes; in Year 2 and Year 3, grades 7 and 8 teachers were offered the summer institutes. Teachers were targeted for coaching in the following manner: grade 6 in Year 1, grade 7 in Year 2, and grade 8 in Year 3.

2. Do impacts on students' CST-ELA scores vary by teacher characteristics and baseline classroom quality?

   a. Does the impact of QTEL on student achievement vary by the level of experience of the teachers in their schools?

   b. Does the impact of QTEL on student achievement vary by the extent to which teachers in their schools have an advanced degree (master's or above)?

   c. Does the impact of QTEL on student achievement vary by the baseline quality of their school's English language arts and English language development classrooms?

Effects on several additional outcomes were estimated to determine whether and how QTEL impacted different alternative measures of classroom quality. As teacher practice is an outcome that can be measured in many different ways, the purpose of these additional analyses was to estimate the impact of QTEL on different measures of this outcome. These measures included an instrument that was more closely aligned with QTEL (and potentially more sensitive to resulting changes in the classroom) and several subscales of the Sheltered Instruction Observation Protocol, which capture different aspects of the classroom environment. Estimating effects on such subscales could have uncovered offsetting program effects (in different directions) on different aspects of the SIOP.

The research questions addressing these additional outcomes were:

3. Does QTEL improve areas of teacher practice expected to aligned with the QTEL program, as measured with the Program Aligned Classroom Observation (PACO) instrument?

4. Does QTEL improve different areas of teacher practice, as measured by subscales of the SIOP?

Finally, QTEL effects were estimated based on intensity of teacher exposure to QTEL. Research questions for this analysis included:

5. What is the potential effect on student achievement and other outcomes of extending teacher professional development to an additional 10 percent of teachers in middle schools in a district?

6. What is the potential effect on student achievement and other outcomes of extending intensive teacher professional development (at least seven days or more) to an additional 10 percent of teachers in middle schools in a district?

Structure of the report

This report has six chapters. Chapter 2 describes the study design, including sample recruitment (schools and teachers), random assignment, data collection, the final study sample, and the data analysis methods. It also explores sample attrition and details baseline equivalence at both the teacher and student levels. Chapter 3 describes QTEL as implemented in the study. Chapter 4 reports the results of the impact analyses for the experimental findings consistent with the

established research domains. Chapter 5 presents the exploratory analysis and findings to further contextualize the impacts discussed in chapter 4. Finally, chapter 6 summarizes the key findings and recommendations for what the results might mean to educators, policymakers, and researchers. Five appendixes are also attached: appendix A presents the statistical power analysis; appendix B presents the results of random assignment; appendix C provides examples of the data collection instruments; appendix D provides the unadjusted means for the primary student-level outcomes; appendix E presents the sensitivity analyses results; and appendix F reports detailed tables and discussion of student sample.

Chapter 2. Study design and methodology

This study measured the effects of QTEL on all students in particular grades at the schools in the study, regardless of whether their teachers participated in specific QTEL activities, such as summer institutes, coaching, or lesson design meetings. This design allowed the effect of offering QTEL to all English language arts and English language development teachers in a school to be determined. Thus, the current study did not measure the effects of QTEL on participating teachers and their students. This design, an intent-to-treat impact study, produces unbiased estimates of the offer of intervention but does not produce unbiased estimates of actual receipt of the intervention (Shadish, Cook, and Campbell 2002; Angrist, Imbens, and Rubin 1996; Orr 1999). Because intent-to-treat estimates reflect real-life conditions, where participation is usually not universal and teachers and students are mobile, these estimates are thought to be good indicators of effects in real-life circumstances (Shadish et al. 2002). Data collection focused on school- and grade-level samples of teachers and students to produce representative school-level estimates of the impacts of offering QTEL to a school's English language arts and English language development teachers for three years (with different grades of teachers being offered the full intervention in different years).

Study design

The evaluation of QTEL uses an experimental design in which middle schools in eight districts were randomly assigned to the intervention group or the control group. In intervention schools, all English language arts and English language development teachers were offered the opportunity to participate in the three components of QTEL professional development: summer institutes, coaching, and lesson design meetings.

The intervention was staggered across three years: grade 6 teachers were the target group in Year 1, grade 7 teachers were the target group in Year 2, and grade 8 teachers were the target group in Year 3. In the 12 middle schools that offered grade 7 and grade 8 only,[9] grade 7 teachers were the targets of implementation in Years 1 and 2 and grade 8 teachers were the targets in Year 3. Thus, in all intervention schools, students in each of the school's grades were exposed to English language arts and English language development teachers who were offered QTEL. Teachers in control schools continued to have access to their regular professional development activities as provided and prescribed by their school or district. These teachers were expected to continue their usual instructional practices in their classrooms during the three-year study period.

The researchers designed the study to follow two cohorts of students. The first cohort included students whose teachers had access to QTEL summer professional development and coaching in each year the students were in middle school. The second cohort enrolled in middle school a year later and encountered the same teachers the year following the teachers' first access to QTEL (when its impact on teacher instruction might have either matured or worn off). This design

---

[9] During the baseline school year (2006/07), there were 13 schools with a grades 7 and 8 configuration in the sample. However, by Year 1 of implementation, one of these schools had converted to a grades 6 through 8 configuration, leaving 12 schools with a grades 7 and 8 configuration.

therefore allowed the cumulative impact of QTEL on students to be tested both during and immediately after the year in which their teachers were first offered it. This design is summarized in table 2.1; the shaded boxes in Year 3 of the study show how the final achievement outcomes of the two student cohorts were captured. Note that even though different cohorts of students were followed over time, this design is not longitudinal in nature. All inferences about program impacts are based on cross-sectional comparisons of student outcomes within a given year and cohort. The table also shows that teacher outcomes were measured, a process discussed more fully in "Key outcomes and measurement."

Table 2.1. QTEL evaluation experimental design

	Year 0 2006/07		Year 1 2007/08		Year 2 2008/09		Year 3 2009/10	
	Summer/fall	Spring	Summer/fall	Spring	Summer/fall	Spring	Summer/fall	Spring
Teachers								
Grade 6 ELA/ELD								
Intervention		O^A	SI/CO	O^B	NI		SI	
Control		O^A	NI	O^B	NI		NI	
Grade 7 ELA/ELD								
Intervention		O^A	SI		SI/CO	O^B	SI	
Control		O^A	NI		NI	O^B	NI	
Grade 8 ELA/ELD								
Intervention		O^A	NI		SI		SI/CO	O^B
Control		O^A	NI		NI		NI	O^B
Students								
Grade 6								
Intervention		O^C	QT	O^C	QT	O^C	QT	O^C
Control		O^C	NI	O^C	NI	O^C	NI	O^C
Grade 7								
Intervention		O^C	QT	O^C	QT	O^C	QT	O^C
Control		O^C	NI	O^C	NI	O^C	NI	O^C
Grade 8								
Intervention		O^C	NI	O^C	QT	O^C	QT	O^C
Control		O^C	NI	O^C	NI	O^C	NI	O^C

ELA/ELD = English language arts/English language development; O^A = Classroom observation only; O^B = Classroom observation and teacher survey data; O^C = Student standardized test scores; SI = QTEL summer institute only; SI/CO = QTEL summer institute and coaching; QT = Exposure to QTEL-trained teacher; and NI = No intervention.
Note: Shaded areas correspond to student cohorts tracked across multiple years. Bold outlines identify classrooms in which observational data are collected.
Source: Authors' construction based on implementation data from WestEd.

To measure the effects of QTEL on student achievement, the scores of grade 7 and grade 8 students in intervention schools on the 2009/10 (Year 3) California Standards Test of English Language Arts (CST-ELA) were compared with those of students in control schools. The difference for the grade 7 students is a cumulative estimate of the effect of two years of QTEL on students whose teacher was exposed to QTEL the year before teaching them (or the effect of one year of QTEL on students for whom grade 7 was the first middle school year in the study school and whose grade 7 teacher was exposed to QTEL the year before teaching them). The difference for the grade 8 students is a cumulative estimate of the effect of three years of QTEL on the English language arts achievement of students whose teacher was exposed to QTEL during the

year teaching them (or the effect of two years of QTEL on students who began middle school in grade 7 or of one year of QTEL on students who transferred into the study school in grade 8).

A comparison of these two sets of student outcome measures (those for grade 7 and grade 8 students in the 2009/10 school year) confounds the effect of more exposure (grade 8 students having potentially been exposed to an extra year of QTEL) with the effect of delayed implementation by teachers (grade 7 students having potentially experienced teachers whose QTEL participation was in the prior year). It is not possible to compare these two impact estimates and identify which of these two factors explains the differences.[10]

To measure the effects of QTEL on student achievement in English language development among English language learner students, the scores of grade 7 and grade 8 students in intervention- and control-group schools on the 2009/10 (Year 3) California English Language Development Test (CELDT) were compared. Because this test is administered at the beginning of the school year, it measures students' knowledge gains in English language development in the previous school year. The difference in the CELDT scores for the grade 7 students is, therefore, an estimate of the effect of one year of QTEL on students whose grade 6 teacher was exposed to QTEL the year before teaching them.[11] The difference in the CELDT scores for grade 8 students in 2009/10 is a cumulative estimate of the effect of two years of QTEL on the English language development achievement of students whose grade 6 and grade 7 teachers were exposed to QTEL during the year teaching them (or the effect of one year of QTEL on the English language development achievement of students who began middle school in grade 7).

As is the case for students' English language arts achievement, a comparison of the two student outcome measures for English language development achievement (those for grade 7 and grade 8 students in the 2009/10 school year) confounds the effect of more exposure (grade 8 students having potentially been exposed to an extra year of QTEL) with the effect of delayed implementation by teachers (grade 7 students having potentially experienced teachers whose QTEL participation was in the prior year). It is not possible to compare these two impact estimates and identify which of these two factors explains the differences.

Using random assignment to evaluate QTEL prevented potential biases from affecting the study's conclusions about the intervention's effectiveness. Using QTEL was the only systematic difference between intervention schools and control schools. Any remaining differences were the result of random sampling error and were controlled for using baseline variables, such as aggregate student assessment scores for prior cohorts of students (Bloom, Bos, and Lee 1999). Including these baseline variables also increased the statistical precision of the impact analyses. These variables were included in the analyses regardless of whether there were statistically significant differences between the intervention and control groups in the baseline characteristics.

---

[10] Within the time constraints of this study, it was not possible to follow the second cohort of students for a third year. Therefore, the study was unable to estimate the cumulative three-year effect of delayed exposure to a teacher who had access to QTEL on students' English language arts achievement.

[11] This estimate was available only for students who attended middle schools that offered a grade 6. That is, it excluded the 12 schools that offered only grades 7 and 8.

Important for the integrity of a randomized controlled trial is keeping the initial sample intact throughout the study. If outcome data cover only part of the sample, the initial balance between the intervention and control groups might not remain intact (Shadish et al. 2002). That is, schools that dropped out of one group could have differed from those that remained in the other. This issue is addressed in detail in "Study sample." Aside from a single two-school district that dropped out of the study shortly after random assignment, sample attrition in the current study was limited to nonresponse to teacher surveys and knowledge tests. The student outcome samples were defined as including only students present for and eligible to take standardized achievement tests in English language arts or English language development and who actually took the tests. The key student-level outcomes were thus universally available for all students in the sample. As detailed in "Study sample," the researchers examined whether there were differences in reported rates of test-taking between the intervention and control schools and did not find any.

Study timeline

A timeline describing the research, recruitment, and data collection activities is in table 2.2. These activities and related data collection instruments are described in greater detail in the next four sections.

Table 2.2. Timeline of the QTEL evaluation, May 2006–December 2010

Date	Task
May 2006–October 2006	Identification of potential districts and preliminary discussions with key district leaders
November 2006–May 2007	District meetings to describe study
March 2007	Memorandum of Understanding with school districts
April 2007	First round of random assignment of schools to intervention and control groups
April 2007	First round of notification to districts of intervention and control status
April 2007–June 2007	Baseline classroom observations
May 2007–April 2008	Year 1 intervention implementation and implementation data collection
June 2007	Second round of random assignment of schools to intervention and control groups
July 2007	Second round of notification to districts of intervention and control status
March 2008–May 2008	Teacher outcome observation and survey data collection for Year 1
August 2008–December 2008	Student-level demographic and test data collected from the districts for baseline (2006/07) and for Year 1 (2007/08)
May 2008–April 2009	Year 2 intervention implementation and implementation data collection
March 2009–June 2009	Teacher outcome observation and survey data collection for Year 2
August 2009–December 2009	Student-level demographic and test data collected from the districts for Year 2 (2008/09)
May 2009–April 2010	Year 3 intervention implementation and implementation data collection
February 2010–May 2010	Teacher outcome observation and survey data collection for Year 3
August 2010–December 2010	Student-level demographic and test data collected from the districts for Year 3 (2009/10)

Note: All student-level data collection included tests administered at two points during the school year: the CELDT in the fall and the CST-ELA in the spring. All these data were collected from the school districts the following fall.
Source: Authors' summary of QTEL evaluation timeline.

Target population

This study targeted middle schools in districts in Southern California in which English language learner students classified as limited English proficient comprised at least 10 percent of the student population. The 10 percent cutoff was chosen during district recruitment, to increase the likelihood that QTEL would be relevant to participating schools and districts and to maintain a large pool of potentially eligible schools from which to recruit. In all eight school districts targeted for recruitment, the study team asked district personnel to identify middle schools that met the student population requirements, had principals and school leadership interested in participating in QTEL and in this evaluation, and were not engaged in or committed to other professional development activities or studies focused on English language learner students. Because of these restrictions, the sample schools are not representative of all middle schools in their districts.

The final evaluation sample consisted of 52 middle schools in Southern California. This sample size was informed by a statistical power analysis (see appendix A), which found that to detect effects on student achievement as small as 0.2 standard deviation, the study required a sample size of 50 schools. To guard against reductions in statistical power that might be caused by attrition during follow-up, two schools were added as a safety margin.

Recruitment

Recruitment relied on several strategies. First, the WestEd QTEL team identified district sponsors in each location. These sponsors were responsible either for district policy for the instruction of English language learner students or for coordinating teacher professional development. They helped WestEd staff arrange in-person meetings to discuss QTEL and the research study with other district staff, middle school principals, and assistant middle school principals. Senior WestEd staff and senior researchers from Berkeley Policy Associates attended the meetings and conducted separate but coordinated presentations on QTEL and the study. The teams also used the in-person meetings to answer questions and discuss the pros and cons of participating in QTEL and the study.

Once middle school principals agreed to participate, the district sponsor obtained the appropriate administrative approvals. In one district, a research review panel had the authority to grant approval; in another, an additional presentation to the school board was required before approval was granted. WestEd and Berkeley Policy Associates attended all requisite meetings and provided the required documentation to support the research applications.

According to district staff, district-level interest in the study was motivated by QTEL's focus on the instructional needs of English language learner students at the middle school level. These students constitute a low-achieving subpopulation of students nationwide (Batalova, Fix, and Murray 2007). All the districts expressed a need to provide teacher professional development focused on the needs of these students. Between November 2006 and March 2007, WestEd and Berkeley Policy Associates visited 11 school districts in a large county, and 7 agreed to participate. After this initial round of recruitment, but before random assignment, two schools in the largest district (District 6; table 2.3) dropped out when they were awarded a grant that committed them to another study. After random assignment, a small suburban district (District 9) also opted out, resulting in the loss of its two schools.

Due to the loss of these schools, the study team continued recruitment efforts in other counties in the region. In May 2007, two school districts in two additional counties agreed to participate in the study. That same month, a charter school from District 6 in the first county agreed to participate, resulting in a final sample of 52 schools, none of which subsequently left the study.

Consent

Because recruitment was at the district and school levels, teacher consent forms to participate in the study were not initially obtained. Instead, the study team developed and signed memoranda of understanding at the district level and asked teachers for their informed consent with each data collection activity they participated in. No formal memoranda of understanding or informed consent were obtained at the school level. Also, no informed consent was obtained from individual students or their parents, because no individually identifiable data were collected directly from students.[12]

Random assignment of schools

Random assignment was conducted in April and June of 2007 (one round each month). Schools were randomly assigned to the intervention group, which was offered the opportunity to participate in QTEL, or to the control group, which was not offered the opportunity to participate in QTEL but could participate in any other available professional development. Berkeley Policy Associates used a SAS® program developed specifically for this study to conduct the first round of randomization on April 3 (see appendix B). This round of randomization included 46 schools from a large county. After two of these schools dropped out and eight new schools were recruited in two additional counties, Berkeley Policy Associates conducted a second round of randomization, using the same SAS® program, on June 20. Before each round, the schools were stratified by district, to balance the sample of schools between the intervention group and the control group.[13,14] As previously discussed, District 9, which included one intervention school and one control school, dropped out of the study shortly after the first round of random assignment but before implementation. The two schools that dropped out represented a single random assignment cluster, thus maintaining the integrity of randomization in all the other clusters.

All 11 schools in District 7 and 1 school in District 6 offered only grades 7 and 8. These schools were not treated differently in randomization. Blocking by district ensured that the 11 schools in District 7 were distributed in a 5/6 or 6/5 ratio between the intervention and control groups. That is, each school had an equal probability of being in the intervention group or the control group.

---

[12] Student test scores were provided by the districts in accordance with Family Educational Rights and Privacy Act regulations.

[13] To avoid creating an extra random assignment cluster, the single school in the original county that was recruited between the first and second rounds of random assignment was added to one of the later county clusters and randomized with that cluster.

[14] In the largest and most urban district in the study (District 6), the sample of schools was divided into three strata before random assignment, based on the proportion of limited English proficient students formally identified in school-level data. Using these strata ensured an equitable distribution of intervention schools in different parts of the "distribution of need" within the district and minimized the likelihood of an imbalance in key demographic background characteristics between intervention and control schools in this district.

The one school in District 6 with a grades 7 and 8 configuration was randomly assigned to the control group.

Table 2.3. Middle schools in the study sample

District	Schools in district	Schools randomly assigned in study	Schools in final study sample before consolidation of three schools	Schools in final study sample after consolidation of three schools
District 1	6	4	4	4
District 2	2	2	2	2
District 3	4	4	4	4
District 4	4	3	3	3
District 5	5	5	5	5
District 6	33	19	19	17[a]
District 7	11	11	11	11
District 8	5	4	4	4
District 9	2	2	0	0[b]
Total	72	54	52	50

a. After random assignment, three schools in District 6 were consolidated into one school, which was treated as an intervention school. In the impact analysis, this school was treated as three schools: two intervention schools and one control school.
b. After random assignment, District 9 opted out, resulting in the loss of their two schools from the study.
Source: Authors' summary based on data collected from school district websites and recruitment documentation.

School consolidation

At the beginning of Year 2 (2007/08), three schools in District 6 were consolidated. At random assignment, two were in the intervention group and one was in the control group. After consolidation, the entire consolidated school was treated as an intervention school for QTEL delivery. However, to maintain the integrity of the randomization, one-third of the students and teachers in this consolidated school were randomly assigned and treated as being in the control group in all subsequent impact analyses. Therefore, the composition of the two research groups remained as it was at the time of random assignment, and one-third of the students and teachers in the consolidated school were control to intervention crossovers.

Study sample

The student sample for English language arts achievement consisted of all students who were attending a study school at the end of the school year, when the CST-ELA was administered, and who took the test. The sample included only students of teachers who were exposed to the intervention or who would have been exposed to the intervention if their school had been assigned to the intervention group. The student sample for English language development consisted of the limited English proficient students of teachers exposed to the intervention who were attending a study school at the beginning of the subsequent school year, when the CELDT was administered. The sample only includes students who actually took the test. The teacher sample consisted of all English language arts and English language development teachers of the target grades for each year in all intervention and control schools. The following section describes each sample and how they changed as the study progressed.

Student sample

A sample flow diagram for English language arts achievement for the full sample of students who were in grade 8 in 2009/10 and who participated in the CST-ELA in that year is shown in figure 2.1. The study began with 54 randomly assigned schools. Of these, as discussed in "Random assignment of schools," both the schools in a single district with two participating schools (one intervention school and one control school) dropped out shortly after random assignment, eliminating the entire district from the study.

Figure 2.1. Sample flow diagram for grade 8 students

Note: Final student counts reflect grade 8 students with CST-ELA scores. During the baseline year, one school district containing two schools dropped out of the study. During Year 2, three schools from the same district were consolidated. At baseline and Year 1, two of these schools were in the intervention group and one was in the control group. The consolidation during Year 2 resulted in one intervention school. In the chapter 4 impact analyses, one-third of the students in this school were moved to the control group for model estimation. Grades 6–8 stayers are students who potentially had three years of exposure to QTEL teachers. Since 12 schools did not have a grade 6, the grades 6–8 stayers include only students who attended study schools with a grade 6.
Source: Student-level data for participating districts, analyzed by authors.

Next, the figure shows counts of individual students who participated in the CST-ELA.[15] A detailed discussion of these counts and statistical tests of the differences in attrition over time is included in appendix F. The overall three-year sample attrition among Year 1 grade 6 students (i.e., between their enrollment in grade 6 and their participation in the 2009/10 Test of English Language Arts as grade 8 students) was 29.5 percent, and the difference in attrition across the intervention and control groups was not statistically significant (p= 0.439). Although it cannot be ruled out that this level of attrition caused bias in the impact estimates, it falls within acceptable limits according to the standards of the What Works Clearinghouse (U.S. Department of Education, 2008).

As discussed in "Study design," the study focused on both grade 7 and grade 8 students in Year 3. An additional sample flow diagram for students assessed with the CST-ELA in grade 7 in the 2009/10 school year is shown in appendix F figure F2. Of the grade 7 students who took that test, 75.9 percent of students in the intervention group (4,686) were in grade 6 in the same school in Year 2 and 80.3 percent of students in the control group (4,546) were in grade 6 in the same school in Year 2. The difference in one-year student retention for these students was not statistically significant ($p=.789$).

Student sample for English language arts achievement outcomes for English language learner students

A sample flow diagram for the 8,098 English language learner students who were in grade 8 in 2009/10 and who participated in the CST-ELA in that year is shown in appendix F figure F3.[16] For this study, English language learner students were defined as students classified as either limited English proficient or redesignated fluent English proficient at the beginning of the school year. This classification was based partly on students' test scores on the CELDT, which is administered at the beginning of each school year (in September or October).[17] Students who entered a grade limited English proficient were required to take the test, as were students who were new to the school system and who reported that their primary language at home was not English. These students were then classified as initially fluent English proficient, redesignated fluent English proficient, or limited English proficient, based partly on the results of the test. Because this study's English language learner student subsamples included both limited English proficient and redesignated fluent English proficient students, reclassification of students over

---

[15] The samples for student-level English language arts outcome measures include only students who participated in the CST-ELA. The study team has no data for students who did not participate in these tests. Using published data from the California Department of Education, the study team determined that aggregate rates of assessment completion ranged from 83.4 percent to 98.5 percent in the study schools, with an average rate of 92.2 percent. There were no statistically significant differences in assessment rates between the intervention and control groups ($p = .543$).

[16] As described previously, the samples for student-level English language arts achievement measures include only students who participated in the California Standards Test of English Language Arts. The study team has no data for students who did not participate in these tests. The California Department of Education provides enrollment information for different classifications of English proficiency at the middle school level but not at the grade level. Therefore, it was not possible to estimate the assessment completion rate for different English language subgroups by grade.

[17] More details about the process through which students are classified as limited English proficient or redesignated fluent English proficient are in "Key outcomes and measurement."

time did not affect the overall composition of these samples. (No students are ever redesignated into initially fluent English proficient status.)

The figure in appendix F shows counts of individual English language learner students who participated in the CST-ELA. The overall three-year sample attrition among Year 1 grade 6 students (i.e., between their enrollment in grade 6 and their participation in the 2009/10 Test of English Language Arts as grade 8 students) was 35.0 percent. The difference in attrition across the intervention and control groups was not statistically significant (p=.735), but at 41.5 percent in the intervention group and 28.9 percent in the control group it was substantial. Because the overall attrition rate was 35.0 percent and the difference in attrition between the intervention and control groups was 12.6 percentage points, this level of attrition is a cause for concern and presents a possible limitation on the validity of the three-year study results for English language learner students in grade 8 in 2009/10.

The remaining grade 8 English language learner students in Year 3 (3,424 in the intervention group and 2,480 in the control group) were students who were not in the same study school during their grade 6 or Year 1 of this study. Among these students, 52.1 percent joined the school in grade 7 (55.7 percent for the intervention schools and 47.8 percent for the control schools; the difference was not statistically significant; $p = .344$) and 21.0 percent joined the school in grade 8 (22.8 percent for the intervention schools and 19.0 percent for the control schools; the difference was not statistically significant; $p = .605$).

Because the study focused on both grade 7 and grade 8 English language learner students in Year 3 (2009/10), a sample flow diagram for the 7,699 English language learner students assessed using the CST-ELA in grade 7 in the 2009/10 school year is also included in appendix F figure F4. Retention rates for grade 7 English language learner students in 2009/10 were as follows: 73.0 percent of students in the intervention group (1,858) were in grade 6 in the same school in Year 2 and 83.4 percent of students in the control group (1,900) were in grade 6 in the same school in Year 2. The difference in retention among grade 7 students between the intervention and control schools was not statistically significant ($p = .942$).

Student sample for English language development outcomes

A sample flow diagram for grade 7 students who had been classified as limited English proficient in the 2008/09 school year, who were therefore required to attend English language development classes during that school year, and who were assessed using the CELDT at the beginning of the 2009/10 school year is shown in appendix F figure F5. These 2009/10 CELDT scores constituted an outcome for the English language development instruction students received in the 2008/09 school year.

Of these students, 82.8 percent in the intervention group (773) and 80.7 percent in the control group (763) were in grade 7 in the same schools in Year 2 and were still classified as limited English proficient. The difference in one-year student retention was not statistically significant ($p = .924$). Neither was the difference in one-year student reclassification from limited English proficient to redesignated fluent English proficient ($p = .840$). Moreover, at 17.2 percent, the rate of reclassification was not so large that it could have had a large effect on any subsequent outcomes measured for students who were not reclassified. That is, the inherent English language proficiency of those reclassified in the intervention group and those reclassified in the

control group would have had to be very substantial for such a bias to materially affect the impact estimates based on the remaining 82.8 percent of the limited English proficient sample.

As with the analysis of English language arts achievement samples, the study also focused on both grade 6 and grade 7 students classified as limited English proficient in 2008/09 and tested using the CELDT in 2009/10. A sample flow diagram for grade 6 students classified as limited English proficient in 2008/09 and assessed with the CELDT in 2009/10 (when they had just entered grade 7) is shown in appendix F figure F6. The percentages of students classified as limited English proficient in 2008/09 who were tested in 2009/10 were 70.3 percent in intervention schools and 76.6 percent in control schools. The difference was not statistically significant ($p = .721$).

Teacher sample

The teacher sample consisted of all teachers who were exposed to QTEL or would have been if their school had been assigned to the intervention group. Depending on the implementation schedule, data collection focused on grade 6 teachers in 2007/08, on grade 7 teachers in 2008/09, and on grade 8 teachers in 2009/10 (see chapter 3). During each school year, all English language arts and English language development teachers in study schools in these grades were considered part of the teacher sample for that year. In schools with only grade 7 and grade 8 classrooms, the study focused on grade 7 teachers in both 2007/08 and 2008/09 and on grade 8 teachers in 2009/10.

At the teacher level, there are sources of attrition, selection, and potential bias. Teacher mobility (the movement of teachers in and out of intervention and control schools) can affect both the implementation and the outcomes of a research study. On average, across all 50 study schools, 29 percent of teachers left between Year 1 and Year 2, and 28 percent left between Year 2 and Year 3 (table 2.4). Across the three years cumulatively, 40 percent of teachers did not remain in the same school. Year-to-year teacher attrition was not statistically significantly different between intervention and control schools (Years 1–2, $p = .472$; Years 2–3, $p = .614$; Years 1–3, $p = .489$).

If this teacher mobility had been related to QTEL (if teachers with access to QTEL were more inclined to remain in the same school), the estimated impact of QTEL on teacher outcomes would have combined its impact on teacher knowledge or practice with its impact on teacher mobility. It would not have been possible to disentangle these effects because there would have been no way to control for differential attrition across intervention and control schools. Even though the overall rates of teacher mobility were not statistically significantly different across the two groups of schools, the measured and unmeasured background characteristics of the stayers and leavers might differ in meaningful ways. Therefore, consistent with the intent-to-treat nature of the impact estimates, this study conceptualizes impact estimates as capturing the combined impact of changes in teacher practice and teacher quality; and changes in the composition of the teacher corps in participating schools. This approach eliminates teacher mobility, differential or not, as a source of potential selection bias in the impact estimates.

A special case of teacher attrition in a study like this occurs when teachers in one research group move to a school assigned to the other. Such crossover weakens the intervention-induced difference in the aggregate school-level experiences of students and teachers. At the outset of the study, the participating school districts agreed to not reassign teachers from intervention to

control schools and vice versa. There was no evidence from review of the teacher rosters that any teacher-level crossover occurred. However, the teacher rosters were incomplete and such crossover thus cannot be ruled out.

Table 2.4. Number of teachers in all grade levels in each year of the study

Year	Intervention	Control	Difference	*p*-value	Total
Year 1	338	303			641
Year 1 to Year 2	Lost 99 (29%)	Lost 81 (27%)	0.026	472	180 (29%)
Year 2	326	281			607
Year 2 to Year 3[a]	Lost 93 (29%)	Lost 75 (27%)	0.018	614	168 (28%)
Year 3	326	282			608
Year 1 to Year 3[b,c]	Lost 138 (41%)	Lost 115 (38%)	0.027	489	253 (40%)

a. Figures do not include District 2 (missing rosters for Year 3), which comprises approximately 1–3 percent of the total sample, depending on the year of the study.
b. While eligible teachers left the study schools between Year 1 and Year 2, many were replaced; therefore, the figures in this row represent the number of eligible teachers (the number of teachers in the study) at the beginning of Year 2. Likewise, the number of eligible teachers in Year 3 represents partial replacement of the sample in between Year 2 and Year 3.
c. Year 1 to Year 3 attrition is calculated as the number lost between Year 1 and Year 3 divided by *n* in Year 1; for the intervention group, attrition is $(138 / 338) * 100 = 41$ percent.
Source: Authors' analysis of teacher rosters from participating schools.

A source of potential bias in the teacher sample more serious than teacher mobility was nonresponse to data collection. It is possible that teachers whose outcomes are observed no longer fully represent the schools where they teach. This would be especially problematic if QTEL were found to affect the likelihood that teachers responded to surveys or allow observation of their classroom. As detailed in "Key outcomes and measurement," there is no evidence of differential nonresponse across the research groups, though response rates for teacher follow-up surveys and knowledge tests were relatively low (69.5 percent). Thus, while response rates among teachers in both research groups were similar, the possibility remains that there were meaningful differences between the two research groups in who responded to the data collection attempts.

Another source of potential bias concerns the nonrandom selection of classrooms for observation. Classrooms were selected for observation based on a convenience sample of classrooms, which was derived from master schedules and teacher willingness in collaboration with school administrators in each school. For practical reasons, it was not feasible to observe all English Language Arts and English Language Development teachers in the selected grade exposed to QTEL in each school. The selective sampling method used would be problematic if the research team and school administrators had selected classrooms to observe using criteria other than those exogenous to QTEL. Thus, for example, administrators in intervention schools might have tried to steer the researchers to observe teachers who participated in QTEL, potentially creating a selection bias. The sample selection was based on master schedules, a neutral criterion. The research team recognizes, however, that sampling was also based on teacher willingness to participate, which may introduce a source of bias.

Key outcomes and measurement

This evaluation collected outcome data in four broad areas: student standardized achievement measures; teacher instructional knowledge; teacher attitudes toward English language learner students; and teacher practice.

The primary outcome measures of the study were student achievement measures from standardized statewide achievement tests, the CST-ELA and the CELDT. The student outcome measures were independent of the intervention. Neither their development nor their administration involved any QTEL or research staff associated with this study. Teacher instructional knowledge and teacher attitudes toward English language learner students served as secondary outcomes and were measured using surveys administered to target teachers in each study year. Teacher practice was another secondary outcome and was measured through classroom observations of teachers in the target year, using the Sheltered Instruction Observation Protocol (Echevarria, Vogt, and Short 2004).

The study team supplemented the student and teacher outcome data with baseline Sheltered Instruction Observation Protocol data, baseline student achievement data, QTEL participation data, interviews with developers and coaches, and a survey of district administrators to gain a better understanding of implementation contexts. The study team also conducted classroom observations in both intervention and control schools, using the Program Aligned Classroom Observation instrument developed by Berkeley Policy Associates.

Primary student-level outcomes

The following sections describe the two standardized assessments that provide student-level data for the primary outcomes: the CST-ELA and the CELDT.

California Standards Test of English Language Arts (CST-ELA)

The CST-ELA is an end-of-grade/end-of-course assessment of student achievement of California's English language arts content standards (California Department of Education 2002). The test is a part of California's Academic Performance Index, which rates the relative performance of each school. In the index's calculations, the California Standards Test carries the most weight. It is administered to all students in grades 2 through 11. The state uses these tests as the core of the assessment system to identify students who achieve desired (proficient) performance levels.

The specific analysis variables for the CST-ELA are the California Standards Test English Language Arts scale scores. The scale scores for each grade and subject area range between 150 (low) and 600 (high). Scale scores compare the tests from year to year and determine performance levels. Half the test's questions change each year, and scale scores adjust for any differences in difficulty that result from question replacement. Although the average number of questions answered correctly should not be compared year to year, scale scores and percentages of students scoring at each performance level may be compared across multiple years within grade level and subject area (for example, grade 4 in 2006 with grade 4 in 2007; California Department of Education 2007). This study does not estimate impacts for combined grades because, as discussed in chapter 1, the hypothesized impacts differed across grades. The California Standards Tests have been reported to be valid and reliable indicators of English

language arts achievement among California students (California Department of Education 2009b).

California English Language Development Test (CELDT)

The CELDT was the measure of student achievement for all sample students under the federal classification of limited English proficient. This test measures English language proficiency and is not strictly a measure of student achievement. However, because it measures learning gains associated with English language development classes, it was treated as a student outcome.

The CELDT is required for all students whose parents report that their primary language at home is not English. It serves three purposes in the state accountability system: to identify limited English proficient students; to determine the level of English language proficiency of limited English proficient students; and to assess the progress of limited English proficient students in acquiring skills of listening, speaking, reading, and writing in English (California Department of Education 2008a). The test is aligned with the California English language development standards, which identify five proficiency levels through which English language learner students progress toward English proficiency. The state provides cutoff scores to districts each year for each proficiency level. Students are deemed English proficient when their overall CELDT score reaches the threshold for the early advanced level or higher.

The CELDT measures students' English proficiency and accounts for one component of the determination of whether an English language learner student is eligible for initial federal classification as limited English proficient or initially fluent English proficient. Once reclassified from limited English proficient to fluent English proficient,[18] English language learner students' California Standards Test scores and grades are monitored by the school district and state for three years, but these students are no longer assessed with the CELDT. Because of this reclassification, the sample of students for whom the CELDT data are available changed over time.

Since some reclassification of students took place after random assignment, it is possible that assignment to QTEL may have affected the reclassification process, causing differences across the research groups in the background characteristics of the subsamples of students who took the test. However, as discussed above, the difference between the intervention and control group in one-year student reclassification from limited English proficient to redesignated fluent English proficient was not statistically significant ($p = .840$), and the overall reclassification rate was 17.2 percent. While this does not rule out differences in the makeup of the not reclassified subsamples in the two research groups, those differences would have to be large to make a considerable difference in the estimated effects on this outcome. Still, the potential for bias remains. And though it is unlikely to be considerable, it is a limitation to the study findings for this outcome.

---

[18] California Education Code Section 313(d) requires the use of multiple criteria in the reclassification of limited English proficient students. The criteria must include the student's score on the California English Language Development Test; the student's score on the California Standards Test of English Language Arts; teacher evaluation, including a review of the pupil's curriculum mastery; and parent input and consultation. Even with teacher recommendation, the redesignation process takes over a year. Therefore, this process is unlikely to have introduced bias into the sample for the California English Language Development Test.

As its outcome variable, the study focuses on the CELDT overall scale score, calculated by averaging four subdomain scale scores: listening, speaking, reading, and writing. For grades 6, 7, and 8, the scale scores range from 248 (low) to 741 (high). The 2009/10 results can be compared with the 2008/09, 2007/08, and 2006/07 results, as the scores are on a common scale, but they cannot be compared with CELDT results earlier than 2006/07 (California Department of Education 2009a). CELDT scores have been reported to be valid and reliable indicators of English language development among California students (California Department of Education 2008b).

Secondary teacher-level outcomes

The study team also measured QTEL's impacts on a set of secondary teacher-level outcomes in the following three areas to better understand QTEL's impact on teachers and classrooms: teacher instructional knowledge; teacher attitudes toward English language learner students; and teacher practice. See appendix C for the instruments used to collect data on these outcomes.

Teacher knowledge test

In each study year, teachers in the grade exposed to QTEL were asked to complete an assessment of their instructional knowledge. Grade 6 teachers were exposed to QTEL in Year 1 (2007/08), grade 7 teachers in Year 2 (2008/09), and grade 8 teachers in Year 3 (2009/10). In the 12 schools without grade 6, in which grade 7 teachers were exposed to QTEL in Year 1, grade 7 teachers were asked to complete the assessment in Year 1 (2007/08). The teacher knowledge test questioned teachers about their general knowledge of theories and practices and included questions on QTEL principles and practices, which were formed to be understandable and relevant to both intervention and control teachers.

An administrative error precluded linking the teacher-level Year 1 data from the teacher knowledge test to teacher identifiers and grade levels. Therefore, only data from Years 2 and 3 for the teacher knowledge test are presented.

To administer the teacher knowledge tests (and the survey described in the following subsection), the study team requested teacher rosters from all the participating districts at the beginning of each school year. These rosters included email addresses and grade levels/subject areas, identifying the appropriate teachers to survey and contact. The availability and completeness of the rosters, however, varied by district.

Master schedules were also collected at the school level. These documents served as scheduling tools for the study's classroom observation data collection activities and also as a validation measure for teacher eligibility for surveys and other data collection activities.

All teacher knowledge test data were collected electronically through a web-based system, using email addresses from the district rosters, or from school-level contacts, when email addresses were unavailable through the districts. The test began in March of each school year. Since the teacher test administration depended on the availability of the district rosters, distribution of the tests was staggered across districts.

To maximize the response rate to the teacher knowledge test, follow-up emails were sent to teachers every two to three weeks for approximately two months. After relationships between the

study team and school-site contacts were developed, follow-up phone calls and emails requesting more support in encouraging teachers to complete the test were also conducted. In addition, while on site to conduct classroom observations, observers distributed flyers to school contacts and teachers being observed, to further encourage participation in the teacher test. For completing the test, teachers were offered $25 gift cards to local or Internet retail outlets.

The overall response rate for the teacher knowledge tests was 70.6 percent: 67.4 percent for the intervention group and 74.1 percent for the control group. The intervention–control difference was not statistically significant ($p = .081$).

Teacher survey

On the teacher survey, which was fielded in parallel with the teacher knowledge test, teachers in intervention schools were asked to complete the same items as teachers in control schools, except for the wording of the questions relating to professional development experience within the past year. These questions were worded with regard to professional development in general for the control group and were specific to QTEL for the intervention group. The teacher survey included questions on teachers' experience, education, credential and certification status, and grades and courses taught.[19] The overall response rate for the teacher survey was 68.7 percent: 67.6 percent for the intervention group and 70.0 percent for the control group.[20] The difference in response rates was not significant ($p = .425$).

The survey also included attitudinal questions about classroom instructional practices, student learning, and teachers' expectations and attitudes about English language learner students. The outcome measure capturing teacher attitudes toward English language learner students was constructed from eight questions, five designed to elicit teacher agreement with negative views of English language learner students and three designed to elicit teacher agreement with positive views of English language learner students and their communities. Teacher responses were recorded along a four-point Likert scale that ranged from (1) strongly disagree to (4) strongly agree. The composite attitude outcome measure, created by averaging teacher responses after reverse-scoring the negatively oriented items, had a range of 1.00 to 5.00.[21] Reliability for the eight-question teacher attitudes scale ranged from 0.61 in Year 1 to 0.55 in Year 3, with an overall reliability of 0.58, as measured by calculating Cronbach's alpha statistics. Because this attitude measure was created for this study, there is no extant information on its validity or relationship to other measures of teacher attitudes. The low reliability of this measure means that impact results based on this measure must be interpreted with caution.

---

[19] The original plan was to use teacher-level data from the teacher survey as a proxy for baseline data and as covariates in the teacher outcome analyses. Given the incompleteness of the teacher survey and the fact that the data were collected after random assignment, it was decided to rely on school-level baseline teacher characteristics available from the California Department of Education.

[20] While the same teachers were asked to complete both the teacher test and the teacher survey, the same teachers did not necessarily complete both. That is, the sample of teachers who completed the teacher knowledge test is not the same as the sample of teachers who completed the survey.

[21] As a sensitivity analysis, a partial credit model (Masters, 1982) was used to estimate teachers' attitude toward English language learner students. This model is a type of Rasch model and has been commonly used in the measurement filed for the items with ordered response categories. The Winsteps computer program was used in this analysis, and the item mean square fit statistics were used to examine how well the item fitted the data. All items were judged to be reasonably fitting and all of them were used to estimate teachers' overall attitude.

The survey also included some implementation questions, which asked intervention teachers to estimate their degree of participation in QTEL activities. The analysis of these implementation questions is discussed in more detail in chapter 3.

Sheltered Instruction Observation Protocol (SIOP) classroom observations

The SIOP is a classroom observation tool used to measure teacher practice (Echevarria, Vogt, and Short 2004). This instrument was developed by the Center for Research on Education, Diversity & Excellence, a national research center funded by the U.S. Department of Education and is aligned with the SIOP model of teacher professional development. The instrument has 30 items measuring teacher classroom performance on eight constructs found to be important to high-quality sheltered instruction for English language learner students: lesson preparation (six items); building background (three); comprehensible input (three); strategies (three); interaction (four); practice/application (three); effectiveness of lesson delivery (four); and lesson review/assessment (four; Echevarria, Vogt, and Short 2004). Each item was scored with a five-point Likert scale ranging from 0 to 4. Because SIOP is not specifically aligned with the QTEL model of professional development, the content validity of this outcome measure is questionable.

According to the developers of the SIOP (Echeverria, Vogt & Short, 2008), the construct of *lesson preparation* includes the development of content objectives and language objectives, the selection of appropriate content concepts, use of supplementary materials, the design of meaningful activities and adaptation of lesson content *by the teacher*. The SIOP construct of *building background* includes the teacher's performance on linking of key concepts to students' backgrounds, linking past learning and new learning, and developing key vocabulary. The construct of *comprehensible input* includes the teacher's use of appropriate speech, clear explanation of academic tasks and use of a variety of teaching techniques to communicate concepts to students. The SIOP construct of *strategies* includes the cognitive and metacognitive strategies used by learners during the lesson to process information as well as the instructional strategies used by teachers such as scaffolding techniques and questioning techniques. The construct of *interaction* includes teacher-student and student-student interaction. The construct of *practice/application* refers to the opportunities students have to use or apply the concepts and academic language that are the stated objectives of the lesson. *Lesson delivery* is a construct made up of items that rate the teacher's performance in supporting the lesson objectives, promoting student engagement and pacing the lesson. Finally, the SIOP construct of *review/assessment* is made up of items that measure how explicitly the teacher reviewed key vocabulary or concepts, how regularly the teacher provided students with feedback and how thoroughly the teacher assessed the students' learning of the stated objectives.

Each observation was conducted during a regularly scheduled English Language Arts or English Language Development classes. The length of these classes ranged from 50 to 90 minutes depending on the scheduling convention used by each school. Observers attended the entire class period and rated the lesson from beginning to end.

Interrater reliability checks were conducted each year of the study through video recordings of classrooms and through paired field observations. For the video checks, observers viewed and rated three video recordings of each of three classrooms at various times during the four-month classroom observation data collection period, using the SIOP. Each observer's ratings were compared against an anchor rating determined collectively by three experienced observers from

the study team. For the field checks, an anchor observer observed the same class session as the field observer and rated the lesson independently. Interrater reliability for these observations was calculated using simple percentage agreement within one point of the anchor score.[22] The number of interrater reliability checks each year depended on the availability of suitable videos and anchor observers. The overall interrater reliability agreements within one point were 87 percent for Year 1 across 104 observations, 92 percent for Year 2 across 37 observations, and 90 percent for Year 3 across 26 observations.

Table 2.5 Percentages for Interrater Reliability Agreement

	Year 1	Year 2	Year 3
Percent of exact agreement	50.8	63.8	56.7
Percent of agreement within 1 point of anchor score	87.5	92.1	92.5
Total observed classrooms with interreliability checks	104	37	26

Source: Authors' summary of data collection activities.

The percentages of exact agreement between observers were 50.8 percent in Year 1 across 104 observations, 63.8 percent in Year 2 across 8 observations, and 56.7 percent in Year 3 across 8 observations. Interrater reliability for the same class session was calculated using simple percentage agreement within one point of the anchor score. These percentages of interrater reliability agreement were 87.5 percent in Year 1 across 104 observations, 92.1 percent in Year 2 across 8 observations, and 92.5 percent in Year 3 across 8 observations.

Since it was not logistically possible for the study team to observe all eligible teachers in each target year, a convenience sample of classrooms was used in each round of classroom observation data collection. The process for selecting teachers for inclusion in the observation sample involved the following steps:

1. The research team contacted the school/district to obtain the roster of English language arts and English language development teachers in the fall of each study year.

2. Teachers in the target grades were identified by the research team. These teachers included teachers of English language arts in the target grade for each year, plus teachers of English language development who taught students in the target grade (sometimes in multigrade classrooms).

3. The school contact for the research team transmitted the invitation to participate in classroom observations to the identified teachers. This invitation included information on the incentive of a $25 gift card to a local or Internet retail outlet.

---

[22]Establishing a benchmark against which to establish interrater reliability for an instrument such as the SIOP involves a judgment call on the part of the researchers. Using a 30-item protocol, with each item scored on a 5-point scale, the potential score differential is 150. Within a fine-grained scale like this, it is difficult to achieve perfect rater alignment, which is why a one-point variance was considered acceptable in establishing interrater reliability. Using a stricter benchmark would have greatly increased the cost of hiring and training classroom observers for this study.

4. Invited teachers expressed interest either directly to the research team or through the school contact.

5. The research team made the final selection of which teachers to include, based on:

   a. The criteria that the observations at each school site must include at least one English language arts and one English language development classroom.

   b. Logistical and cost considerations, such as scheduling that would allow for observers to complete two to three observations each day.

No additional effort was made to ensure that this sample was representative of the entire population of teachers and classrooms. The research team did not expect a convenience sample to affect the validity or reliability of the measures obtained from these observations. However, it is possible that the selection of this sample resulted in systematic differences between observed and unobserved classrooms and between observed classrooms in intervention schools and those in control schools. It is impossible to assess the extent of these potential biases. A further limitation of the observation protocol is that it is based on just one day and might not capture a typical class session.

The number of attempted classroom observations ranged from three to five per school. The completion rates for the classroom observations were 74.4 percent in the baseline year (2006/07), 100 percent in Year 1, 96.7 percent in Year 2, and 90.7 percent in Year 3.

Before conducting the classroom observations, the observers were not told of the research status of the schools they were visiting. However, the observers could have learned of a teacher's or school's participation in QTEL during the observation (from teacher comments or the presence of QTEL training materials). Thus, the data collectors may not have been blind to the school's research status (intervention group or control group). This potential for observer bias is another limitation of the SIOP data.

Implementation data

To examine implementation, interviews were conducted, participation was documented and collected, and district administrator surveys were administered. To examine the quantity and quality of QTEL implementation, the research team collected participation data and conducted individual and group interviews with QTEL developers and coaches. The participation data consisted of scanned attendance sheets from summer institutes and lesson design meetings, as well as records provided by the coaches for each completed session with a teacher. The scanned attendance sheets included teachers' signatures and school names. Each completed coaching record accounted for one full cycle of coaching (a lesson design meeting, a classroom observation, and a postobservation debriefing).

Interviews with QTEL developers were conducted annually to better understand the implementation context (and how it changed over time) and to document successes, challenges, and variations across districts. Interviews with coaches were conducted twice in Years 1 and 2 and once in Year 3, to get the coaches' perspectives on implementation successes, challenges, and variations at the school level. In all, 26 interviews, each lasting approximately one hour,

were conducted in person or over the phone. All interviews were recorded and transcribed to ensure data accuracy.

A representative from each participating school district was surveyed on professional development needs, professional development offerings to control teachers during the study period, satisfaction with QTEL, and contextual factors affecting the delivery of professional development. The 22-item survey was administered in Years 2 and 3. Response rates were high, with 7 of 8 districts responding in Year 2 and all districts responding in Year 3.

Exploratory outcome measure

The study also included one exploratory outcome measure, the Program Aligned Classroom Observation (PACO) instrument.

Program Aligned Classroom Observation (PACO) instrument

Berkeley Policy Associates developed the PACO instrument specifically for this evaluation. The PACO was developed because while the SIOP was the best available measure for assessing teacher practice in classrooms with English language learners, it does not address all the constructs included in the QTEL program. The PACO instrument was administered in Years 2 and 3. The instrument consisted of 22 items assessed using a five-point Likert scale ranging from 0 to 4. After conducting a literature review and observing the QTEL professional development institutes in Year 1, the study team identified the following three underlying constructs that would serve as observable evidence of the manifestation of the QTEL principles in classrooms: *activity structure*, *interaction patterns*, and *lesson content*.

Activity structure refers to the organization of the activity in which the students are expected to engage. For example, items relating the design, purpose and procedures for completing the lesson as implemented during the observation period are part of the measurement construct of *activity structure*. *Interaction patterns* refer to the ways in which students are observed interacting during the lesson. For example, items relating to the interaction of the teacher with the whole class, small groups and individual students during the observation period are part of the measurement construct of *interaction patterns*. Items measuring student-student interaction are also included within this construct. *Lesson content* refers to the academic content of the lesson delivered during the observation period. For example, items relating to the academic standards addressed by the lesson, the text(s) used, the critical thinking skills required to complete the task, and the academic language skills emphasized are included in the construct of *lesson content* within this instrument. While several of these constructs overlap with those of the SIOP, the focus of the PACO is centered on the students' engagement in activity as the desired manifestation of the teacher's professional practice.

Although the PACO instrument was created with the QTEL goals and principles in mind the items were worded carefully to capture general evidence of quality instructional practices that would be relevant in both intervention and control classrooms. The instrument was administered at the same time and for the same lesson as the SIOP. Therefore, the classrooms in which PACO was conducted were not explicitly chosen to be representative of all English language arts or English language development classrooms or teachers in the school. In Year 2, the number of attempted classroom observations was 75, ranging from one to two classrooms per school. In half the 150 classrooms targeted for observation with SIOP, an independent observer rated the

same lesson on the PACO instrument. In Year 2, 99 percent of the 75 scheduled observations with this instrument were completed, and in Year 3, 91 percent of the 150 scheduled observations were completed.

Interrater reliability of the PACO instrument was also estimated using video observations in Year 2 and video observations plus field checks in Year 3. In each year of administration, interrater reliability checks for each of three video-taped classroom lessons were conducted for each observer over the four-month classroom observation data collection period. Each observer was asked to view and rate a video using the PACO instrument. Each observer's ratings were then compared against an anchor rating determined collectively by two experienced observers on the study team. Interrater reliability for these observations was calculated using simple percentage agreement within one point of the anchor score. The overall interrater reliability agreement was 78.3 percent across 22 observations in Year 2 and 90 percent across 25 observations in Year 3.

A summary of the data collection is presented in table 2.6.

Table 2.6. Data collection activities

Data collection activity	Baseline (2006/07)	Year 1 (2007/08)	Year 2 (2008/09)	Year 3 (2009/10)
Primary confirmatory (student level)				
CST-ELA	X	X	X	X
CELDT	X	X	X	X
Secondary confirmatory (teacher level)				
Teacher instructional knowledge test			X	X
Teacher background and attitude survey		X	X	X
SIOP classroom observations	X	X	X	X
Implementation/treatment contrast				
Teacher survey (implementation questions)		X	X	X
Attendance data		X	X	X
Developer and coach interviews		X	X	X
District administrator survey			X	X
Exploratory				
PACO classroom observations			X	X

Source: Authors' summary of data collection activities.

Sample characteristics and baseline equivalence

To evaluate whether random assignment resulted in research groups that were statistically equivalent at baseline, the researchers compared key background characteristics of the intervention and control groups. These included school characteristics, baseline observations of teachers' instructional practices measured with the SIOP, and student and teacher characteristics. Ideally, these baseline background characteristics would have been compared for the analytic samples of students and teachers featured in the impact analysis. However, no individual-level

data on the characteristics of these sample members at the time of random assignment were available, and thus these comparisons rely on school-level averages for the time period. Data for the baseline comparison were collected from three sources: the California Department of Education, baseline student-level data from participating districts, and baseline classroom observations.

School-level and student-level background variables

No statistically significant differences were found between intervention and control schools in key school-level characteristics (table 2.7).

Table 2.7. Baseline comparison of key school-level characteristics for all students, 2006/07 (percent, unless otherwise noted)

Characteristic	Intervention	Control	Difference	*p*-value
Number of students enrolled	1115.7	1074.8	40.9	.662
Racial/ethnic minority	76.9	77.0	–0.1	.982
Eligible for free or reduced-price lunch	58.0	54.0	4.0	.532
English only	43.7	45.9	–2.2	.669
Initially fluent English proficient	30.0	27.9	2.2	.333
Redesignated fluent English proficient	13.5	12.0	1.5	.524
Limited English proficient	22.9	22.3	0.6	.856
Sample size (*n*)	26	25[a]		

Note: School-level averages were weighted by the number of students enrolled. A two-tailed *t*-test was applied to differences between the intervention and control groups. The significance level for each test was set at .05.
a. One school was newly configured and opened in Year 1. Therefore, it had no baseline data.
Source: Authors' calculations from data obtained from the California Department of Education.

Student achievement background variables

The key samples in this evaluation include all grade 8 students present during the 2009/10 school year and all grade 7 students present during the 2009/10 school year. To assess the baseline equivalence of the schools, aggregate grade 7 and grade 8 test scores on the CST-ELA and the CELDT for the last cohorts of grade 7 and grade 8 students prior to random assignment (2006/07) were compared. Aggregate school-level data were used for these analyses because randomization was at the school level and because individual student baseline achievement data were not available for the students in the impact analysis sample, since these students had not yet enrolled in their middle schools at the time of random assignment. These aggregate school-level baseline student achievement variables were used as control variables in the impact estimation. There were no statistically significant differences between intervention and control schools for student performance on the CST-ELA or CELDT scale scores or for percentage proficient based on all students or on English language learner students only (table 2.8).

Table 2.8. Baseline comparison of grade 7 and grade 8 California Standards Test of English Language Arts (CST-ELA) and California English Language Development Test (CELDT) scores, 2006/07

Test and grade	Intervention	Control	Difference	*p*-value
All students				
CST-ELA mean scale score				
Grade 7	339.5	339.6	0.3	.959
Grade 8	335.5	334.5	1.0	.850
Percent of students proficient and above				
Grade 7	45.1	42.5	2.6	.510
Grade 8	39.7	37.5	2.2	.556
English language learner students				
CST-ELA mean scale score				
Grade 7	291.9	293.2	−1.4	.575
Grade 8	285.2	287.0	−1.8	.451
Percent of students proficient and above				
Grade 7	8.7	6.7	1.9	.073
Grade 8	3.3	3.6	−0.4	.531
CELDT mean score				
Grade 7	538.4	540.1	−1.8	.660
Grade 8	545.7	547.8	−2.1	.651
Percent of students proficient and above				
Grade 7	41.0	44.2	−3.2	.238
Grade 8	40.5	43.5	−3.0	.382
Sample size	26	25[a]		

Note: Calculations used aggregate school-level student test scores. School-level averages were weighted by the number of students tested. A two-tailed *t*-test was applied to differences between the intervention and control groups.

a. One school was newly configured and opened in Year 1. Therefore, it had no baseline data.

Source: Authors' calculations from data obtained from the California Department of Education.

Teacher characteristics

There were no statistically significant differences between the intervention and control schools in key school-level teacher characteristics (table 2.9).

Table 2.9. Baseline comparison of teacher characteristics

Characteristic	Intervention	Control	Difference	*p*-value
Years of experience[a]	13.3	11.9	1.3	.104
Percentage of teachers with master's degree or above[a]	48.2	42.7	5.5	.174
Baseline classroom quality[b]	2.3	2.3	0.0	.958

a. Based on school-level averages for 51 schools (26 intervention schools and 25 control schools). One school was newly configured and opened in Year 1. Therefore, it had no baseline data.
b. Based on school-level averages of Sheltered Instruction Observation Protocol scores for 40 schools (22 intervention schools and 20 control schools)
Source: Authors' analysis of data from the California Department of Education and primary data collected for study.

Teacher knowledge and teacher attitudes were not measured at baseline because the required approval for the data collection was not received from the Office of Management and Budget until after the prior school year had ended.

Data analysis methods

The following sections describe the data analysis methods for implementation data and the impact analysis.

Analysis of implementation data

The implementation data in chapter 3 were analyzed using quantitative and qualitative methods. Interviews with 26 QTEL coaches and developers were analyzed inductively using NVivo8. District administrator survey data were analyzed using descriptive methods in Stata. Treatment contrast information (see table 3.9) was analyzed using a two-tailed, Pearson's chi-square test in SPSS. Teacher participation data were analyzed using descriptive methods in SPSS.

Confirmatory impact analysis

The following sections describe aspects of the impact analysis, including the model, multiple comparisons, approaches to handling missing data, and sensitivity analyses. Unadjusted means for primary student-level outcomes are in appendix D.

Description of model

The study team analyzed the effectiveness of QTEL using hierarchical linear regression models to account for the effect of clustering by school on the variance structure of the data. In each primary student-level impact analysis, school-level and student-level covariates were included to improve the estimates' statistical precision and reduce the likelihood of random sampling

variation affecting the impact estimates.[23] These variables were selected as covariates because, conceptually, they were expected to predict variation in the study's primary measure. Similarly, for the secondary teacher-level analyses, the team included school-level covariates, including teacher characteristics at the school level. A summary of the covariates used in the primary student-level impact analyses is in table 2.10; a summary of the covariates used in the secondary teacher-level impact analyses is in table 2.11.

Table 2.10. Covariates used in the primary student-level analyses

Domain	Baseline measure(s)
School level	
Randomization stratum	School district and, within the largest district, three separate strata by the percentage of students classified as limited English proficient
Aggregate student performance	2006/07 grade 7 or grade 8 CST-ELA and CELDT overall test scores and percentage proficient (as measured for all students or for English language learner students only, depending on the analysis)
School demographics	Percentages of students categorized as limited English proficient and/or eligible for free or reduced-price lunch
Teacher practice	Average baseline classroom observation (SIOP) score for 2006/07 within each school
Teacher education level	Percentage of teachers with master's degree and above from California Department of Education
Teacher experience	Average years of teaching experience at baseline from California Department of Education
Student level	
Student demographics	Gender and primary language

Source: Authors' analysis of individual-level data from participating districts and school-level data from the California Department of Education.

The covariates used in the secondary teacher-level analysis are: teacher demographics at baseline, obtained from the California Department of Education; the school's aggregate baseline SIOP score for 2006/07, obtained from the classroom observations conducted during the baseline year; and the aggregate CST-ELA score in 2006/07 (by grade), obtained from the California Department of Education website.

---

[23] The study team explored the possibility of using teacher-level covariates in student-level impact analyses, but found that the information about class rosters and teacher/student linkages was not sufficiently reliable in all participating school districts for this to be feasible.

Table 2.11. Covariates used in the secondary teacher-level analyses

Domain	Baseline measure
Teacher demographics	Average years of teaching experience at baseline from California Department of Education; percentage of teachers with master's degree and above from California Department of Education
Teacher practice	Average baseline classroom observation (Sheltered Instruction Observation Protocol) score for 2006/07 within each school
Aggregate student performance	Aggregate CST-ELA score in 2006/07
Randomization stratum	School district, and within the largest district, three separate strata by the percentage of students classified as limited English proficient

Source: Authors' analysis of individual-level data from participating districts and school-level data from the California Department of Education.

Primary student impact analysis model

The hierarchical linear regression models used in these analyses have two levels. The student measures have a level for school and a level for student (nested in school).

The hierarchical linear model for the analysis of the student-level primary outcomes is illustrated in equations 1 and 2.

$$Y_{ik} = \alpha_{0k} + \sum^{x}\alpha_x X_{xik} + \varepsilon_{ik} \tag{1}$$

$$\alpha_{0k} = \gamma_0 + \gamma_1 E_k + \sum_{2}^{P}\gamma_p P_{pk} + \sum_{P+1}^{P+10}\gamma_p B_{pk} + \varphi_k \tag{2}$$

In these equations, Y_{ik} is the test score for student i in school k. X_{xik} and P_{pk} are vectors of student-level and school-level background variables. (The latter also include aggregate measures of teacher credentials and experience.) B_{pk} is a vector of ten random assignment block indicators. Equation 1 includes a school effect α_{0k}, the dependent variable in school-level equation 2. That equation has the experimental dummy variable E_k, whose coefficient γ_1 is the main QTEL effect in this system of equations. ε_{ik} and φ_k are random error terms at the two different levels. By estimating these two equations simultaneously using a PROC MIXED procedure in SAS, the researchers ensured that the statistical results were appropriate for the nested nature of the data.

The models in equations 1 and 2 were used for the CST-ELA outcome variables, the CELDT outcome variables, and each analytic sample for the primary research questions (grade 8 students, grade 7 students, English language learner students, and limited English proficient students).

Secondary teacher impact analysis model

The hierarchical linear model for the analysis of the teacher-level secondary outcomes is in equations 3 and 4.

$$Y_{jk} = \alpha_{0k} + \sum_{1}^{x} \beta_x X_{xjk} + \delta_{jk} \tag{3}$$

$$\alpha_{0k} = \gamma_0 + \gamma_1 E_k + \sum_{2}^{P} \gamma_s P_{pk} + \sum_{P+1}^{P+10} \gamma_p B_{pk} + \phi_k \tag{4}$$

In these equations, Y_{jk} is the outcome for teacher j in school k. X_{xjk} and P_{pk} are vectors of teacher and school-level background variables. B_{pk} is a vector of ten random assignment block indicators. Equation 3 includes a school effect α_{0k}, the dependent variable in school-level equation 4. That equation, which is very similar to equation 2, has the experimental dummy variable E_k, whose coefficient γ_1 is the main QTEL effect on the teacher-level outcome.

Multiple comparison procedures

As chapter 1 discussed, the researchers conducted six statistical tests to estimate primary impacts on students and three tests to estimate secondary teacher-level impacts. To reduce the probability of finding impacts due to chance alone, two separate Benjamini-Hochberg adjustments (Benjamini and Hochberg 1995) were applied: one for the six primary student outcomes and another for the three secondary teacher outcomes. These adjustments are included in chapter 4. The Benjamini-Hochberg procedure involves ordering the unadjusted p-values for each outcome from largest to smallest across each level (student and teacher) and multiplying the unadjusted p-value by $N/(N - j + 1)$, where N is the number of outcomes in a level and j is the order of the test based on the unadjusted p-value. Null hypotheses for which the adjusted p-value was less than .05 were rejected.

Approach to missing data

The study team relied on two techniques to minimize the impacts of missing data. These techniques included replications of the secondary impact analyses using different procedures for handling missing data: impact estimates based on cases with complete data on baseline and outcome measures (listwise deletion) and impact estimates with covariates set to zero when missing and with missing value dummy variables included in the model. The findings presented in the body of this report are based on the second method.

The research team did not impute primary student-level outcome measures because the follow-up sample was defined as students tested in spring 2010; therefore, there were no missing outcome data for these students. The study team also did not impute the classroom observation outcomes, because the sample for these observations was based on teacher availability when a school was visited for classroom observation. Thus, there was no subset of classrooms that was explicitly targeted for a visit and for which classroom observation data were missing and could be imputed.

Lastly, in subgroup analyses based on English language learner student status, the analysis team eliminated five observations (less than 1 percent of the student sample) with missing data on the variable defining the subgroup breakdown.

Sensitivity analyses

Sensitivity analyses were conducted to accompany all student-level and teacher-level impact analyses and assess their sensitivity to analytical decisions and contextual factors. See appendix E for results from all the sensitivity analyses.

Sensitivity analyses were conducted to test the sensitivity of the measured impacts to the different methods of dealing with missing data: listwise deletion and setting missing values to 0 and adding a missing-value dummy variable to the impact analyses model for teacher- and student-level covariates.

The intervention intended to expose middle school students in the intervention group to QTEL for their entire middle school career. However, the length of such exposure varied because many students were in 1 of the 12 study schools that did not have a grade 6 and some students moved into their schools in grade 7 or grade 8. To assess whether the study findings were sensitive to the inclusion of these students, the researchers estimated effects for students in three-year middle schools and effects for students who spent three consecutive years in their middle school.

As discussed, the consolidation of three schools into one resulted in one school initially assigned to the control group that was exposed to the intervention. To maintain the integrity of randomization, a third of the students and teachers in the consolidated school were treated as control group members in the impact analyses. As a sensitivity analysis, student and teacher impacts were estimated under the assumption that all the students and teachers in the consolidated school were in the intervention group and were compared with the main impact findings.

For the teacher attitude analysis, a Rasch model was used to test the sensitivity of teacher attitude impacts to the aggregation method used to create the single teacher attitude measure from the teacher survey responses. For the eight teacher attitude items in the teacher survey, teachers indicated their degree of agreement with statements along a four-point Likert scale that ranged from 1, strongly disagree, to 4, strongly agree. For the benchmark model, the teacher attitude outcome measure was constructed by calculating an average of the responses across the eight items, with adjustments made for missing responses (total score divided by total number of items responded to). For this method to be valid, the following assumptions were made: the distances between the Likert-scale points can be treated as equal (or any differences are irrelevant), and the items represent equal measures of the same construct (or any differences in the level of attitude measured are irrelevant).

As a sensitivity analysis, Rasch modeling was used to construct attitude measures for each teacher along a continuous interval scale that represents orientation to English language learner students. While the main impact findings present teacher attitudes using an average to adjust for missing responses, teacher attitude impacts using the Rasch model are presented to test the sensitivity of the impact findings to the aggregation method used.

Finally, because some teachers received two years of QTEL because they taught in two-year middle schools (for grades 7 and 8 only), a sensitivity analysis was conducted to determine the sensitivity of the impacts to the inclusion of these teachers: teacher impacts using a sample that excluded all teachers who received more than one year of QTEL were compared with impacts using a sample that included those teachers.

Exploratory analyses

Exploratory analyses were conducted to contribute to the interpretation of the confirmatory impact analyses and to point to possible directions for future research.

Subgroup analyses

Chapter 4 presents analyses for the full sample of students targeted by QTEL and for English language learner students and limited English proficient students in specific subgroup analyses. Additional subgroup analyses are in chapter 5; these analyses allowed estimation of how QTEL's impacts are moderated by teacher characteristics and of more detailed breakdowns of students' English language learner status. They also allowed exploration of whether different teacher and student subgroups responded differently to being offered QTEL. In these analyses, student achievement on the CST-ELA and the CELDT were the outcomes of interest. The subgroup-focused research questions are discussed in detail below.

1. Do impacts on students' CST-ELA scores vary by student English language learner status?

Separate subgroup analyses for all four official English language learner student subgroups were conducted as exploratory analyses. These subgroups are English only, initially fluent English proficient, redesignated fluent English proficient, and limited English proficient. English-only students are those who speak English at home. Initially fluent English proficient students do not speak English at home but were determined to be fluent in English the first time their English language proficiency was assessed by the school district. Redesignated fluent English proficient students were determined to be limited English proficient at one point but had since been reclassified as speaking English fluently. Limited English proficient students were those whose proficiency was (or continued to be) limited.[24] The statistical significance of the difference in QTEL impacts on the CST-ELA across these four subgroups was assessed using a joint F-test (a Wald test). The analyses were structured as follows:

---

[24] As explained in appendix F on p. 189, the reclassification of some students (17.2 percent) from limited English proficient to redesignated fluent English proficient took place after random assignment and might have been affected by the intervention. As a result, there might have been systematic intervention–control differences in the background characteristics of the limited English proficient and redesignated fluent English proficient subgroups. To the extent that these differences were uncontrolled for in the analyses, they might have biased the impact results. However, considering that the difference in the reclassification rate between the two research groups was not statistically significant ($p = .840$), it is unlikely that such a bias would be considerable.

$$Y_{ik} = \alpha_{EOik} + \alpha_{IFEPik} + \alpha_{RFEPik} + \alpha_{IFEPik} + \sum_{1}^{z}\beta_z Z_{zik} + \delta_{ik} \tag{5}$$

$$\alpha_{EOk} = \gamma_0 + \gamma_{EL}E_k + \sum_{2}^{P}\gamma_p P_{pk} + \sum_{P+1}^{P+10}\gamma_p B_{pk} + \phi_k \tag{6}$$

$$\alpha_{IFEPk} = \gamma_0 + \gamma_{IFEP}E_k + \sum_{2}^{P}\gamma_p P_{pk} + \sum_{P+1}^{P+10}\gamma_p B_{pk} + \phi_k \tag{7}$$

$$\alpha_{RFEPk} = \gamma_0 + \gamma_{RFEP}E_k + \sum_{2}^{P}\gamma_p P_{pk} + \sum_{P+1}^{P+10}\gamma_p B_{pk} + \phi_k \tag{8}$$

$$\alpha_{LEPk} = \gamma_0 + \gamma_{LEP}E_k + \sum_{2}^{P}\gamma_p P_{pk} + \sum_{P+1}^{P+10}\gamma_p B_{pk} + \phi_k \tag{9}$$

In this system of equations, the student-level equation 5 was estimated without any additional intercept terms. This level-1 model included covariates Z_{zik}, which were grand-mean centered. The equation had four school-specific effects—α_{EOik}, α_{IFEPik}, α_{RFEPik}, and α_{LEPik}—which captured school-level variation in student-level achievement outcome Y_{ik} (the CST-ELA for the students' grades) for each English language learner student subgroup. These four variables then became the dependent variables in school-level equations 6 through 9. In these models, P_{pk} were vectors of school-level background variables and B_{pk} were vectors of 10 random assignment block indicators, as introduced in "Data analysis methods." The coefficients γ_{EO}, γ_{IFEP}, γ_{RFEP}, and γ_{LEP} captured QTEL effects for the four subgroups, and the Wald test assessed whether these effects were statistically significantly different from each other. To minimize multiple comparison problems, the research team did not conduct bilateral tests of the significance of impact differences across pairs of these subgroups.

Because these were student-level subgroups, the statistical power of the estimates remained strong (see appendix A). The research team estimated a minimum detectable effect size of 0.24 for a subsample as small as 30 students per school. (The smallest of the four groups—redesignated fluent English proficient—included an average of 35 students per school; the others ranged from 246 to 469.)

2. Did impacts on students' CST-ELA scores vary by teacher characteristics?

The specific questions these exploratory analyses answered were:

a. Does the impact of QTEL on student achievement vary by the level of experience of the teachers in their schools?
b. Does the impact of QTEL on student achievement vary by the extent to which teachers in their schools have an advanced degree (master's or above)?

c. Does the impact of QTEL on student achievement vary by the baseline quality of their school's English language arts and English language development classrooms?

To conduct subgroup analyses based on these characteristics, the research team used school-level aggregates to divide the sample. This was necessary for the teacher characteristics because it was not possible to link individual students to individual teachers in the student-level data. Specifically, the analyses examined how student achievement impacts were moderated by the following three baseline constructs: teacher experience (the average years of teaching experience); teacher education (the percentage of teachers with a master's degree or above); and average total baseline SIOP classroom observation score (a measure of teacher practice).

A challenge associated with analyses like these is that teacher, student, and school characteristics can be strongly correlated with one another, making it difficult to disentangle differences that could be attributed to these characteristics from differences associated with other aggregate school-level characteristics. Thus, before deciding to proceed with these subgroup analyses, a series of school-level regression analyses were conducted in which each potential subgroup variable was regressed on a series of five other baseline variables as follows:

$$X_{pk} = \alpha_0 + \sum_1^5 \beta_p P_{pk} + \delta_{pk} \qquad (10)$$

Each of the six baseline variables P_{pk} in equation 10 is one of the following list: teacher experience; teacher education; the average total baseline SIOP classroom observation score; the percentage of students proficient on the CST-ELA in the 2006/07 (baseline) school year; the percentage of students classified as limited English proficient; or a dummy variable indicating whether the school was "in improvement" in the 2006/07 school year, as defined in the No Child Left Behind Act of 2001.[25] (There are only five baseline variables P_{pk} on the right side of equation 6 and six variables in this list because one of the listed variables took the place of variable X_{pk} on the left side of the equation, while the others were included on the right side.) Before proceeding with any subgroup analyses using a breakdown based on a variable X_{pk}, the researchers first confirmed that the R^2 of the corresponding regression was less than 0.5; that is, half or more of the variance in X_{pk} was not explained by the variables P_{pk}.[26] Based on this assessment, the research team eliminated the following subgroup analyses because school-level regression analyses revealed insufficient independent variation: by the percentage of students proficient on the CST-ELA in the 2006/07 (baseline) school year and by the percentage of students classified as limited English proficient.[27]

---

[25] In accordance with Section 1116 (b)(1) of the No Child Left Behind Act of 2001, schools considered to be "in improvement" include "any elementary school or secondary school served under this part that fails, for 2 consecutive years, to make adequate yearly progress as defined in the State's plan under section 1111(b)(2)."

[26] This 0.5 cutoff value may be considered arbitrary. However, there is no known established precedent for these kinds of analytical decisions, and it appeared that a subgroup breakdown using a baseline variable with less than 50 percent independent variation would not be sufficiently useful to readers.

[27] Although there was sufficient independent variation for the "school improvement" status at baseline, this subgroup analysis was not conducted because only 18 schools (8 control schools) were in school improvement.

The teacher- and school-level subgroup analysis was structured as follows:

$$Y_{ik} = \alpha_{0k} + \sum_{1}^{x} \alpha_x X_{xik} + \varepsilon_{ik} \tag{11}$$

$$\alpha_{0k} = \gamma_0 + \gamma_1 E_k LOSIOP_k + \gamma_2 E_k HISIOP_k + \sum_{3}^{P} \gamma_p P_{pk} + \sum_{P+1}^{P+10} \gamma_p B_{pk} + \phi_k \tag{12}$$

Equation 11 is the same student-level model used throughout the full-sample analyses and was introduced as equation 1. In it, outcome Y_{ik} was regressed on school effects α_{0k} and student baseline variables X_{xik}. The vector of school effects α_{0k} became the dependent variable in equation 12, which included interactions of the intervention dummy E_k and school-level subgroup identifiers $LOSIOP_k$ and $HISIOP_k$ (in this example). These subgroup identifiers were created so that $LOSIOP_k$ was a 0/1 variable that was 1 for schools whose teachers scored low on the baseline classroom observation measure and $HISIOP_k$ was a 0/1 variable that was 1 for schools whose teachers scored high on the baseline classroom observation measure. As discussed, the research team used an *F*-test to assess whether QTEL coefficients γ_1 and γ_2 were statistically significantly different from each other. In equation 12, P_{pk} was a vector of school-level background variables and B_{pk} was a vector of 10 random assignment block indicators.[28]

Statistical power analyses confirmed there was enough statistical power to detect effects for the subgroups for which impacts were estimated. Because the subgroups were not equal in size, there was some variation in statistical power across them. For subgroups of half the schools, the analysis produced a minimum detectable effect size (on student achievement) of 0.29; for subgroups of a third of the schools, it produced a minimum detectable effect size of 0.34.

Additional outcomes

Chapter 5 also presents impact estimates for two additional sets of outcomes as exploratory analyses. These estimates address the following research questions:

3. Does QTEL improve teacher practice, as measured with the PACO instrument?

4. Does QTEL improve teacher practice, as measured by subscales of the SIOP?

As discussed in "Key outcomes and measurement," additional teacher-level outcome data was collected using the PACO instrument. These analyses explore whether QTEL impacted classroom quality. The estimated impacts on the total score on the instrument are in chapter 5, as well as on the 3 subscales. The alphas for the full PACO instrument and the subscales were estimated using Cronbach's alpha[29], which are displayed below:

---

[28] Note that equation 12 does not have teacher- or student-level data, as indicated by the absence of subscripts other than the letters *p* and *k* to indicate variables in a vector and schools, respectively.

[29] The 2-item Language instruction subscale of the PACO was not included in this analysis because it did not meet the criteria of a minimum Cronbach's alpha of .70.

1. Full Program Aligned Classroom Observation instrument (22 items), alpha = 0.92

2. Activity structure subscale (11 items), alpha = 0.89

3. Lesson content subscale (6 items), alpha = 0.84

4. Interaction patterns subscale (3 items), alpha = 0.89

In addition, the research team explored whether QTEL impacted four subscales of the SIOP[30]:

- Lesson Preparation (9 items), alpha = 0.79
  - Includes SIOP items on preparation and building background (Q1-Q9)
- Input and Interactions (7 items), alpha = 0.75
  - Includes SIOP items on comprehensible input and interaction (Q10-Q12, Q16-Q19)
- Lesson Activity (6 items), alpha = 0.85
  - Includes SIOP items on strategies, practice, and application (Q13-Q15, Q20-Q22)
- Lesson Delivery (8 items), alpha = 0.85
  - Includes SIOP items on lesson delivery and evaluation (Q23-Q30)

Dose-response analysis

As chapter 3 described, there was substantial variation in the intensity and fidelity of QTEL implementation across schools and districts. The dose-response analysis in chapter 5 explores how this variation might have affected teacher- and student-level outcomes based on the degree of exposure to English language professional development (in intervention and control schools), as measured with the teacher survey.

The exploratory research questions this analysis addresses are:

1. What is the potential effect on student achievement and other outcomes of extending teacher professional development to an additional 10 percent of teachers in middle schools in a district?

2. What is the potential effect on student achievement and other outcomes of extending intensive teacher professional development (at least 7 days or more) to an additional 10 percent of teachers in middle schools in a district?

Using the teacher survey data, the research team created two sets of variables that measured, at the school level, the percentage of teachers who reported having participated in professional development during the year they were interviewed and the percentage of teachers who reported having participated in professional development for at least seven days during the year they were interviewed. Both intervention and control teachers completed the teacher survey; dose measures were created for teachers from both groups. Two sets of these variables were created: one focused on professional development in English language development standards and the other on professional development in instructional strategies for teachers of English language learner students in secondary schools. Both these constructs directly related to the focus of the QTEL professional development.

---

[30] Although the full SIOP instrument measures eight separate constructs, we only included the four subscales that demonstrated a reliability of .70 or higher as measured by Cronbach's alpha.

The research team then estimated impacts on student and teacher outcomes using these participation variables instead of the experimental intervention dummy. Thus, instead of measuring the direct impact of QTEL on an outcome, these dose-response analyses measured the effect of additional participation and/or intensity of participation on the outcome. The analyses were structured as follows:

$$Y_{ik} = \alpha_{0k} + \sum_{1}^{x} \alpha_x X_{xik} + \varepsilon_{ik} \tag{13}$$

$$\alpha_{0k} = \gamma_0 + \gamma_1 ANYELD_k + \gamma_2 ELD7_k + \gamma_3 ANYSTR + \gamma_4 STR7 + \sum_{5}^{P} \gamma_s P_{sk} + \phi_k \tag{14}$$

Equation 13 is the standard student-level impact model. School-level equation 14 replaced the school-level intervention dummy E_k with four new "treatment" variables—$ANYELD_k$, $ELD7_k$, $ANYSTR_k$, $STR7_k$—which represented the participation variables previously described. All these variables were created using self-reported teacher survey data. Together they captured the effects of professional development provided to the teachers in the study sample, in both the penetration of English language–focused professional development (in intervention and control schools), as measured with related professional development activities in the school, and the intensity of those professional development activities.

Chapter 3. Implementation of QTEL

This chapter details how QTEL was implemented in this study. The following information builds on the introduction to QTEL, its theoretic framework, and its intended effects. To contextualize the results of the impact analysis, descriptive analyses of how implementation varied year to year, of differences in professional development experiences between intervention and control teachers, and of information about teacher participation in QTEL are provided.

Intended implementation

QTEL supports teachers of all content areas, including mathematics, science, and social studies. For this study, however, the developers focused on a single area—English language arts—to ensure the capacity to deliver services to 26 middle schools simultaneously. Serving this many schools was a challenge because QTEL does not have a one-size-fits-all definition of quality across the secondary grade levels and content areas. Nor is it scripted. Instead, the developers, coaches, and presenters of QTEL work with teachers and administrators to tailor its delivery to the individual needs of teachers and the specific implementation context.

QTEL includes three main components: summer institutes (Building the Base), individualized teacher coaching, and lesson design meetings. WestEd's QTEL team included six experienced coaches delivering the professional development. The team staggered implementation to focus coaching on teachers in one grade level per year (grade 6 in 2007/08,[31] grade 7 in 2008/09, and grade 8 in 2009/10). All English language arts and English language development teachers in these grade levels were offered the intervention. QTEL was originally designed so that a team of two to three coaches, each specializing in one or two academic content areas, would work with the same teachers over consecutive years to build on a teacher's understanding of the principles and grow the teacher's expertise in implementing lessons aligned with those principles. Working with the same teachers over three years also allowed the coaches to develop teacher leadership and interdepartmental support for the QTEL principles. However, due to the staggered design, the implementation of this version of QTEL did not typically allow for teachers to work with a coach over multiple years.

To ensure an adequate number of attendees in the summer institutes, these sessions were open to teachers from grades 6 and 7 in Year 1 and grades 7 and 8 in Years 2 and 3. However, the QTEL team had the capacity to provide coaching and in-classroom support to approximately one-third of the teachers in intervention schools in each study year. Therefore, to maximize the potential effects of QTEL and to mimic the potential impacts of schoolwide implementation, the QTEL developers and the research team decided that it would target each middle school grade level for one year.

Approximately seven teachers at each school were offered four to six coaching sessions each year. This allowed the students in intervention schools to potentially move through three consecutive intervention classrooms, experiencing three years of exposure to a teacher with

---

[31] For schools with no grade 6, coaching targeted grade 7 teachers.

access to QTEL in their middle school career[32] (see figures 2.1 through 2.6 in chapter 2). The schedule of implementation of each QTEL component for each year of the study is shown in table 3.1.

Table 3.1. QTEL model intervention components, by study year

Component	Year 1 (2007/08)	Year 2 (2008/09)	Year 3 (2009/10)
Summer institutes	➢ All **grade 6 and grade 7** English language arts and English language development teachers ➢ School site and district administrators ➢ 3 days in June ➢ 4 days in August ➢ QTEL professional developers	➢ All **grade7 and grade 8** English language arts and English language development teachers ➢ School site and district administrators ➢ 3 days in June ➢ 4 days in August ➢ QTEL professional developers	➢ All **grade 7 and grade 8** English language arts and English language development teachers ➢ School site and district administrators ➢ 3 days in June ➢ 4 days in August ➢ QTEL professional developers
Coaching and in-classroom support	➢ **Grade 6** participants (grade **7** in schools with only grades 7 and 8) ➢ Four to six individualized cycles per teacher	➢ **Grade 7** participants ➢ Four to six individualized cycles per teacher	➢ **Grade 8** participants ➢ Four to six individualized cycles per teacher
Lesson design meetings	➢ Four to six after-school study sessions for all English language arts and English language development teachers	➢ Four to six after-school study sessions for all English language arts and English language development teachers	➢ Four to six after-school study sessions for all English language arts and English language development teachers

Source: Authors' summary of QTEL materials from WestEd.

Implementation contexts and experiences

Berkeley Policy Associates collected information on the contexts of implementation across participating school districts through district administrator surveys and 26 in-depth interviews with the QTEL developers and coaches working with intervention teachers. The interviews addressed the developers' and coaches' reflections on implementation of each component of the QTEL model. The interviewers probed for examples of supports and hindrances to implementation as planned for each district. Administrator surveys asked respondents about contextual challenges, such as budget shortfalls and non-QTEL professional development taking place in the districts. Seven administrators (87.5 percent) completed the survey in Year 2, and eight (100 percent) completed it in Year 3.

---

[32] According to the study design, ideally, students in grade 6 would be exposed to teachers experiencing QTEL for the first time in 2007/08, and when those students moved to grade 7, they would again have teachers exposed to QTEL for the first time (grade 7 teachers in 2008/09), and again in grade 8. However, teachers often taught multiple grade levels in one given year. For practical reasons such as this, it was difficult to adhere strictly to the research design. For example, 14 teachers received three years of summer institutes instead of two years; and 24 teachers received more than one year of coaching during the study.

QTEL developers were interviewed annually after the summer institutes. Coaches were interviewed twice during the school year (January and May), except in Year 3, during which each coach was interviewed once, in April 2010. These interviews were either in person or by phone and were recorded and transcribed. The interview transcripts were coded inductively (using NVivo8) for supports and hindrances to implementation.

In addition, attendance data and teacher survey data were used to provide information on implementation described in this chapter. Data collection activities are presented in table 3.2.

Table 3.2. Implementation data collection activities

Data collection activity	Year 1 (2007/08)	Year 2 (2008/09)	Year 3 (2009/10)
Teacher survey (implementation questions)	X	X	X
Attendance data	X	X	X
Interviews with developers	X	X	X
Interviews with coaches	X[a]	X[a]	X
District administrator survey		X	X

a. In Years 1 and 2, interviews were conducted twice per year, once in January and once in May.
Source: Authors' summary of study activities.

The following implementation summary was constructed from analysis of interview data and was triangulated with participation data, such as attendance sheets, coaching logs, and district administrator survey data. Detailed information from administrator surveys is also in this section.

Building the Base summer institutes

QTEL includes a seven-day Building the Base summer institute, divided into two seminars. The ideal grouping for these seminars is 24 teachers from the same district.

The three-day introductory seminar familiarizes teachers with the theoretical base for QTEL and gives them an experiential understanding of its five principles, through participation in the type of activities they will learn to implement throughout QTEL. The following four-day seminar engages teachers in critical analysis of academic task structure and supports them in planning lessons that enact the QTEL principles in their own classrooms.

Contextual variation

Across the three study years, several contextual issues might have influenced the implementation of the summer institutes as designed. In the study design, teachers in each grade level were supposed to have at least one year of Building the Base (grades 6 and 7 in Year 1; grades 7 and 8 in Years 2 and 3). The biggest obstacle in planning the summer institutes was budgetary uncertainty; districts did not know in June which teachers would be employed in which grade level the following September, making it difficult for the developers to identify teachers in the appropriate grade levels for the summer institute each year and difficult for the evaluators to determine participation rates.

The staggered implementation design created logistical problems for the QTEL team, due to staffing uncertainties. To reach the number of schools required by the research design, QTEL

was offered to teachers in one or two grade levels and in two teaching departments (English language arts and English language development). Normally, QTEL is implemented schoolwide, across grade levels and subject areas. Under more typical conditions, all teachers in an intervention school would be exposed to the summer institute each year of implementation. In this study design, however, only grade 7 teachers were exposed to three summers of institutes (grade 6 teachers were eligible for the summer institutes in Year 1 and grade 8 teachers in Years 2 and 3; table 3.3). This limited the depth of coverage and degree of spiraling of content that could be offered in the Years 1 and 2 of the summer institutes.

Contextual issues by year

Other major contextual factors are outlined below by year.

Year 1

Compressed timelines. Site recruitment and random assignment were extended into May 2007. Therefore, districts did not know which teachers would be invited to the summer institute until a few weeks before the first seminar. No accommodations, such as hiring substitute teachers, were made to allow teaching staff on year-round calendars to attend the summer sessions. These factors caused confusion and scheduling conflicts for the teachers participating in QTEL, which required the developers to adapt the schedule for the summer institutes and to offer multiple make-up sessions for those unable to attend the originally scheduled introductory seminar.

Natural disasters. In 2007, California wildfires affected the implementation of summer institutes in two districts. In one of these areas, fires forced many residents to evacuate and delayed the start of the school year. This event interfered with the implementation of the make-up sessions and the second seminar for a large number of teachers in the intervention group. Therefore, implementation of the summer institutes and make-up sessions extended into October 2007.

Limited participation. Two districts joined the study late and had less time to recruit teachers for the summer institutes. As a result, these districts focused teacher recruitment on grade 6 teachers. Although the original implementation plan was to include grade 6 and grade 7 teachers in the Year 1 summer institutes, grade 7 teachers were not included in the Year 1 implementation in these two districts.

Year 2

Teacher incentives. In 2008, one district reported that teachers were reluctant to spend seven days of their summer vacation in professional development without compensation. The district worked out an agreement in which teachers would receive four days of professional development before the end of the 2007/08 school year, and teachers would then spend three days at the end of their summer vacation in the institutes. While this agreement potentially eliminated much of the need for make-up sessions in this district, it also forced the developers to adjust the content of the seminars to accommodate a modified schedule.

Financial uncertainties. As it turned out, make-up sessions were required in September 2008 because a state budget crisis threatened a large number of teachers with layoffs. Teachers did not know whether they had a job with the district the coming school year or which grade they might be teaching. This uncertainty affected participation greatly, as teachers were not willing to attend

professional development under such ambiguous work conditions. Make-up sessions were required for the teachers who had not attended the seminars as planned. Fewer accommodations were required in Year 2 for year-round calendars, as one large district eliminated year-round scheduling and all teachers in that district had the same schedule for summer break.

Grade-level changes. In three districts, layoffs did occur, and teachers were moved to different grade levels. This meant that teachers who had not been exposed to the first Building the Base seminar in June were suddenly eligible for the second seminar in August/September. Because the newly eligible teachers had not attended the first seminar, the content in the second seminar had to be adjusted to include the foundation discussion (normally provided in the introductory seminar), and a make-up session was required to provide these teachers with the content of the second seminar.

Y E A R 3

Financial crises. By the third summer of implementation, the California state budget crisis worsened and three more districts laid off teachers. Additionally, all but two districts reneged on agreements to pay teachers for the time spent in summer professional development, which may have affected teacher participation (see "Teacher participation in QTEL"). Although there were fewer layoffs in Year 3, teachers on temporary contracts who had received the intervention in prior years were typically not rehired. Uncertainty about layoffs continued to make summer planning difficult for developers as well as for teachers who were not sure which grade they would be teaching.

Calendar changes. One district in the sample converted to a year-round schedule in Year 3, shortening the summer break for teachers and making the scheduling of the summer institute problematic. The transition required teachers to attend several non-QTEL-related meetings during the prior school year. In addition, there was uncertainty about teaching assignments for the following year. When the new school year started (only weeks after the last school year ended), teachers were reluctant to give up class time to attend professional development sessions.

Inconsistent sessions. To get critical mass (groups of 24) at the institute sessions—and in a variation from the original implementation design for this study—the developers opened the summer sessions to grade 6 and grade 7 teachers. Some of these teachers had participated in previous institutes; others were new to QTEL. In other cases, the sessions were rescheduled to occur as individual days staggered throughout the school year. This situation of both new and repeating attendees, combined with the adapted scheduling (for example, three and four days, four and three days, seven consecutive days, or individual days staggered) required that developers adapt their presentations and session content to the local situation, resulting in inconsistent coverage across intervention sites. (Participation rates for each QTEL activity are in "Teacher participation in QTEL.")

Individualized coaching

As chapter 1 described, the study design called for the QTEL staff to provide four to six coaching cycles to teachers in the target grade each year (grade 6 teachers in Year 1, grade 7 teachers in Year 2, and grade 8 teachers in Year 3). Each coaching cycle had three sessions: a one-on-one lesson design meeting, an observation of the lesson's implementation, and a

debriefing session. The study design estimated that approximately seven teachers at each school would each participate in four to six coaching sessions each year.

Contextual variation: Year 1

Geographic challenges. Implementation of the coaching component in Year 1 was characterized by the immediate need to establish capacity in the geographic area of the study. The developers accomplished this through a three-pronged approach to staffing: they hired three experienced QTEL coaches who had participated in a prior implementation on the East Coast; they hired a local retired administrator with extensive experience providing professional development to teachers working with English language learner students, and then trained the administrator in the principles and practices of QTEL; and they assigned two highly experienced coaches to travel from WestEd's San Francisco office to deliver coaching services in Southern California. Each staffing choice presented unique advantages and challenges to the delivery of the coaching cycles in the first year.

East Coast QTEL coaches. The three coaches from the East Coast arrived in Southern California when wildfires were devastating large areas of the state, so their first adjustment was to extreme environmental conditions. Their second adjustment was to the California education context (state standards and assessments, local institutional organizations and practices, relevant local and regional political issues, and so on).

New coaches. One coach, who had many years of experience in the California education system and brought relevant, regionally specific institutional memory to the coaching team, left the QTEL team after one year. The schools she served in the first year were taken on by one of the experienced San Francisco-based coaches in the next years.

Travel. The costs of travel from WestEd's San Francisco office to the Southern California schools limited contact between experts and teachers. Coaches often addressed individual teacher needs through phone and email, but they were not able to be as adaptive and responsive as were local coaches.

Contextual variation: Year 2

District changes. During Year 2, one district transitioned to a uniform, districtwide, year-round calendar. Although the new calendar would not take effect until Year 3 of QTEL implementation, planning for this change occurred during Year 2. In this district, transition planning tended to monopolize the attention of both teachers and administrators, who were required to attend many meetings that ultimately interfered with the QTEL coaching cycles. There was also uncertainty about school staffing assignments for the coming year, which caused confusion and anxiety among teachers. During this time, the district was also implementing a new curriculum that required teachers to attend associated trainings. According to the QTEL developers, all these factors combined to distract those teachers' attention from QTEL's goals and to limit their availability for coaching sessions in Year 2.

Budget crisis. The California budget crisis affected all participating districts in Year 2. Districts that had agreed to reimburse teachers for professional development days associated with QTEL rescinded. Teachers in all participating districts received layoff warning notices, which influenced participation in QTEL coaching. In schools where large numbers of teachers were laid

off or transferred in Year 1, the QTEL coach had to build new relationships in those schools from scratch in Year 2.

Administrative support. One district in the sample included school-based literacy coaches for the intervention group in the summer institutes. This district also provided coaching to the grade 6 teachers while the QTEL coach worked with grade 7 teachers. According to interview respondents, this level of administrative support made implementation of the coaching much more pervasive at the intervention schools in this district.

Contextual variation: Year 3

District changes. Although the new curriculum in the district mentioned in the Year 2 discussion was compatible with QTEL principles and tasks, its implementation created many obstacles in planning. For example, teachers in the district were required to work with a district literacy coach, which created scheduling conflicts with QTEL coaching. As a result, the QTEL coach had to adjust the visitation schedule repeatedly throughout the year, creating a scenario of interrupted and incomplete coaching cycles for these teachers.

Budget crisis. In four districts, the budget crisis was particularly severe. In one district, teachers were involved in contract negotiations, and the union required them to stop working at 3:30 p.m. each day. They were not permitted to perform any additional duties after school, including attending QTEL coaching sessions and lesson design meetings. In districts where the summer institutes were delayed (due to uncertainties about teacher assignments), the coaching was delayed as well, creating a compressed timeline for implementation.

Administrative support. According to the QTEL coaches, administrative support at the school level was the biggest influence on their ability to adapt to scheduling constraints and implement the coaching cycles as planned during Year 3. All coaches stated that teachers at school sites with strong administrator buy-in felt more accountable to participate in the QTEL coaching. Strong administrative support encouraged teacher participation in QTEL.

Lesson design meetings

As chapter 1 described, the lesson design meetings supported the implementation of QTEL pedagogical tools and processes. The meetings were monthly collaborative planning sessions or study groups at school sites, facilitated by the QTEL coach. In theory, these sessions would be open to all teachers in each intervention school and would thus promote the development of a schoolwide culture supporting QTEL. The QTEL coach would lead these collaborative sessions and support the application of the QTEL principles to cross-cutting issues affecting multiple teachers and classrooms in the same school.

Contextual variation

In practice, the lesson design meetings varied in their implementation more than any other QTEL component. Throughout all three years of the study, several factors contributed to this variation, including the availability of the coach to provide timely and targeted meetings and the existing school culture and structure.

Coaches' availability. The availability of a coach to provide timely and targeted meetings depended on whether the coach was local or had to travel to the area. The schedules of local

coaches were fairly flexible; they were able to adapt their plans quickly and address relevant topics in a timely manner. The schedules of traveling coaches restricted their ability to respond to local needs with spontaneity. Schools with traveling coaches received no more than four lesson design meetings, because the meetings had to coincide with a traveling coach's visit to the area. However, coaches living locally were able to provide more sustained support to teachers and could even hold additional sessions as needed.

School culture and structure. The organization of the lesson design meetings varied greatly depending on the school culture. In schools where teachers already did a great deal of interdisciplinary teaming and cross-grade planning, lesson design meetings were better attended and enhanced the culture of schoolwide collaboration. In these schools, QTEL lesson design meetings were incorporated into the existing professional learning communities. In other schools, QTEL was seen as separate from the professional learning community, and much of the professional learning time in lesson design meetings was devoted to school-level issues, rather than to instructional planning. In these schools, the meetings were attended solely by teachers receiving coaching that year. Teachers who received coaching the year before could use the lesson design meetings as an opportunity to maintain their existing relationship with the coach, but few teachers who had not received any coaching would attend. According to the QTEL coaches, the expectations set by the site administrators were a key factor in shaping the culture of collaboration at the school and teacher participation rates in lesson design meetings. Strong principal buy-in and visible support for QTEL encouraged teacher enthusiasm for these meetings.

Structural issues also presented challenges to participation in lesson design meetings. The limited availability of compensation for participating teachers influenced participation rates. In one district, a transition to a year-round calendar and a new curriculum presented obstacles to QTEL participation. The reforms in this district created competing demands on teachers' time and greatly limited participation in the lesson design meetings. In another district, union policy during collective bargaining prohibited teachers from participating in lesson design meetings in Year 3.

Teacher mobility

Throughout the study, QTEL implementation had to be modified because approximately 40 percent of the original teacher sample was lost from Year 1 to Year 3 (see table 2.4). Teacher mobility was a factor in implementation because, as new teachers replaced those who left, the content, pacing, and scheduling of QTEL had to be modified. This mobility was caused by teachers getting laid off, moving to nonstudy districts, or leaving their positions for other reasons.

District administrator surveys

Administrators from each study district were surveyed in Year 2 and Year 3 about a variety of district characteristics, including challenges to teaching English language learner students, professional development opportunities, budgetary concerns and other contextual factors, and QTEL implementation. The survey data in this section focus on professional development related to English language learner students, mandated (non-QTEL) professional development,

administrators' descriptions of teacher participation in QTEL, and administrators' concerns about QTEL implementation.

English language learner student–specific professional development

Survey response rates were 88 percent in Year 2 (n = 7 districts responding) and 100 percent in Year 3 (n = 8 districts responding). The types of non-QTEL professional development in which teachers (intervention and control) participated over the three years of the study, according to district administrators, are presented in table 3.3. The most prevalent type was a strategy called Differentiated Instruction, with 66.7 percent of respondents reporting it was offered in their district. Differentiated Instruction is intended to provide students with different ways of accessing academic materials and acquiring skills so that all students can learn effectively regardless of their ability levels. In addition, 60 percent reported offering Specially Designed Academic Instruction in English (SDAIE) training. SDAIE, also referred to as "sheltered instruction," is designed to support students in learning English and various content concurrently through hands-on activities, cooperative learning environments, and interactive strategies. Of the respondents surveyed, 33.3 percent reported offering training in High Point[33], a textbook and curriculum package developed for teachers to support English learners in building language and literacy skills. Finally, 20 percent offered Guided Language Acquisition Design[34], a professional development model that promotes English language acquisition, academic achievement, and cross cultural skills; and 53.3 percent reported offering other professional development. The "other" category includes Cross-Cultural, Language, and Academic Development[35]; Senate Bill 472 language arts[36]; Sheltered Instruction Observation Protocol training[37]; Structured English Immersion[38] instructional design; Thinking Maps[39]; and district English language development curricula.

---

[33] Schifini, A., Short, D., & Villamil Tinajero, J. (2001). *High Point: Success in language, literature, content.* Carmel, CA: Hampton Brown.

[34] Guided Language Acquisition Design (GLAD) provides teachers with instructional strategies, a curriculum model, and demonstration lessons. GLAD emphasizes that both the primary language and the development of English can complement each other. Brechtel, M. (2001). *Bringing it all together : Language and literacy in the multilingual classroom.* Carlsbad, CA: Dominie Press. http://projectglad.com/

[35] Diaz-Rico, L.T., & Weed, K.Z. (2010). *The crosscultural, language, and academic development handbook: A complete K-12 reference guide* (4th ed.). Boston, MA: Allyn & Bacon.

[36] SB 472 provides professional development for implementation of the California state adopted mathematics and reading and language arts instructional materials. SB 472 is a 40-hour institute that targets program organization, knowledge and awareness of research, content standards/frameworks, universal access, assessment, planning and implementation. Senate Bill 472. California Statutes of 2006. Ch. 524. http://www.cde.ca.gov/pd/ca/ma/mard.asp

[37] Echevarria, J., Vogt, M., & Short, D. J. (2004). *Making content comprehensible for English learners: the SIOP model* (2nd ed.). Upper Saddle River, NJ: Pearson Education, Inc.

[38] Structured English Immersion (SEI) dedicates significant amounts of the school day to explicit teaching of English. English is the language of instruction, although modifications are made based on the abilities and proficiency levels of the English learners.

[39] Thinking Maps is a professional development model that is adopted at a district or whole school level. It is designed to train teachers to help students use tools that allow students to graphically organize their thoughts, make connections across disciplines, and track learning over time. Hyerle, David (Ed.), Curtis, S. (Ed.), & Alper, L. (2004). *Student successes with thinking maps: School-based research, results, and models for achievement using visual tools.* Thousand Oaks, CA: Corwin Press.

Table 3.3. Administrators' description of teachers' participation in English language learner student–specific professional development, 2008/09 and 2009/10

Type of professional development	Percentage
Differentiated Instruction	66.7
Specially Designed Academic Instruction in English	60.0
High Point	33.3
Guided Language Acquisition Design	20.0
Other[a]	53.3
Sample size (Years 2 and 3 combined)	15

a. Includes: Cross-Cultural, Language, and Academic Development; English language development curriculum; explicit English language development for elementary and secondary students; Senate Bill 472 language arts; Structured Instruction Observation Protocol; Structured English Immersion instructional design; and Thinking Maps.
Source: Authors' analysis of primary data from administrator surveys.

Mandated professional development

Administrators were asked if their teachers were required to participate in any particular professional development activities during the years surveyed. They were then asked an open-ended question on what types of professional development activities teachers participated in. Fifty-three percent of the respondents reported that their districts mandated that teachers participate in some types of professional development; thus, QTEL was competing for teacher time and attention with other required training activities. The mandated professional development activities specified included: SB 472 Math and Language Arts training for state-approved curricula, English language arts curriculum packages, math and technology training, and other professional development models aimed at successful teaching and learning across curricula (e.g. Thinking Maps).

Teacher participation in QTEL

Surveys asked administrators to describe teacher participation in QTEL in their districts. These surveys were administered in Years 2 and 3. Overall, 20 percent of administrators reported that all or most English language arts/English language development teachers actively and consistently participated in QTEL, 60 percent reported that some English language arts/English language development teachers actively and consistently participated in QTEL, and 20 percent reported that few English language arts/English language development teachers participated actively or consistently in QTEL (table 3.4). No administrators responded that they did not know or were not sure about teacher participation in QTEL.

Table 3.4. Administrators' description of English language arts/English language development teachers' participation in QTEL, 2008/09 and 2009/10

Level of teacher participation in QTEL	Percent
All or most teachers actively and consistently participate in QTEL	20.0
Some teachers actively and consistently participate in QTEL	60.0
Few teachers actively or consistently participate in QTEL	20.0
I do not know or I am not sure	0
Sample size (Years 2 and 3 combined)	15

Source: Authors' analysis of primary data from administrator surveys.

Administrator concerns about QTEL

Administrators were asked about their concerns with QTEL. The surveys included statements about QTEL and a Likert scale response format. Respondents chose a response option from a five-point scale (not at all concerned; not very concerned; somewhat concerned; very concerned; don't know/not applicable). Because of the low number of respondents, the "not at all concerned" and "not very concerned" options were collapsed into one category, and the "somewhat concerned" and "very concerned" options were collapsed into a second. The "don't know/not applicable" data was dropped from the table due to the low frequency of responses.

As shown in table 3.5, 64.3 percent of administrators reported they were somewhat or very concerned about teachers' willingness or commitment to participate. When asked if they were concerned about the helpfulness/usefulness of QTEL, more than 80 percent were not or not very concerned. For resources required to provide substitute teachers so that intervention teachers could participate in summer training, 46.7 percent were somewhat or very concerned. In response to questions about the resources needed to pay teachers to attend training, 53.3 percent were somewhat or very concerned. For teacher time commitments, 66.6 percent of administrators were somewhat or very concerned about the time commitment needed by teachers to participate in QTEL; and 40 percent were somewhat or very concerned about the time commitment needed by teachers to participate in the evaluation data collection activities. Other concerns noted by administrators included the study limitations on where and to whom the intervention would be offered, which constrained district-level professional development; the lack of support from principals in the study; and the time commitment required of district office staff to collect data. Responses about these other concerns were too few to be included in table 3.5.

District administrators generally believed in the usefulness of QTEL but were concerned with practical matters such as time, resources, and funding. This information is consistent with what was gathered through interviews—that budgetary issues affected implementation and participation in QTEL professional development.

Table 3.5. Administrators' concerns with QTEL, 2008/09 and 2009/10

Administrator concerns	Not concerned or not very concerned	Somewhat or very concerned
Teachers' willingness/commitment to participate	35.7	64.3
Helpfulness/usefulness	> 80.0	< 20.0
Resources needed to provide substitutes to allow teachers to participate in the QTEL Building the Base summer training	46.7	46.7
Resources needed to pay teachers while attending training outside of regular school hours	46.7	53.3
Time commitment by teachers to participate in the QTEL training activities (summer institute, coaching, after-school meetings)	33.3	66.6
Time commitment by teachers to participate in the evaluation data collection activities (surveys, site visits, observations)	60.0	40.0
Sample size (Years 2 and 3 combined)	15	15

Note: Response categories were collapsed from a five-point Likert scale into two categories. Due to the low number of respondents, "not at all concerned" and "not very concerned" were collapsed into one category, and "somewhat concerned" and "very concerned" were collapsed into a second. The "don't know/not applicable" data were dropped from the table.
Source: Authors' analysis of primary data from administrator surveys.

Treatment contrast

This study compared QTEL impacts on intervention teacher and student outcomes with those of control schools where teachers received professional development support (with the exception of QTEL), as provided by their respective schools or districts. Intervention teachers in the year-by-year target groups (see table 3.1) were given the opportunity to attend QTEL training activities as prescribed by the staggered design. This section describes the control teachers' experience with professional development, as reported on the teacher surveys.

Experience with professional development

The study team compared the types of professional development experienced by intervention and control teachers to determine whether QTEL provided unique support to teachers of English language learner students. The intervention teachers surveyed reported significantly more exposure to professional development activities aligned with the principles and components of QTEL than did the control teachers surveyed (table 3.6). Survey response rates across all three years were 66.5 percent for the intervention teachers and 69.2 percent for the control teachers, based on teacher survey data pooled across the three years of the study. The distribution of responses between intervention and control teachers was significantly different for English language development standards, content area standards, instructional strategies for secondary English language learner students, and differentiated instruction. There were no statistically significant differences between the intervention and control groups for professional development related to published curricula in language arts or other areas or other professional development.

Table 3.6. Teachers' reported experiences with professional development topics, by intervention- and control-group, 2007/08, 2008/09, and 2009/10

Professional development topic and duration of experience	Intervention		Control		*p*-value
	n	**Percent**	*n*	**Percent**	
English language development standards					.000**
Not at all	34	9.2	75	21.4	
1–3 days	117	31.6	142	40.5	
4–6 days	81	21.9	72	20.5	
7 days or more	138	37.3	62	17.7	
Total sample size	370	100	351	100	
Content area standards					.009**
Not at all	29	7.8	48	13.8	
1–3 days	111	30.0	123	35.2	
4–6 days	97	26.2	80	22.9	
7 days or more	133	35.9	98	28.1	
Total sample size	370	100	349	100	
Support for a published curriculum in language arts/other areas					.201
Not at all	100	27.0	93	26.9	
1–3 days	100	27.0	113	32.7	
4–6 days	89	24.1	83	24.0	
7 days or more	81	21.9	57	16.5	
Total sample size	370	100	346	100	
Instructional strategies for secondary English language learner students					.000**
Not at all	20	5.4	52	14.8	
1–3 days	69	18.7	145	41.2	
4–6 days	99	26.8	80	22.7	
7 days or more	181	49.1	75	21.3	
Total sample size	369	100	352	100	
Differentiated instruction					.000**
Not at all	24	6.5	51	14.6	
1–3 days	120	32.6	155	44.3	
4–6 days	103	28.0	74	21.1	
7 days or more	121	32.9	70	20.0	
Total sample size	368	100	350	100	
Other professional development					.562
Not at all	99	45.0	103	49.5	
1–3 days	40	18.2	42	20.2	
4–6 days	29	13.2	23	11.1	
7 days or more	52	23.6	40	19.2	
Total sample size	220	100	208	100	

**Statistically significant at *p* < .01 using a two-tailed Pearson's chi-square test.
Note: Response rates for teacher surveys across three years of the study were 66.5 percent for intervention teachers and 69.2 percent for control teachers.
Source: Authors' analysis of primary data from teacher surveys administered in 2008, 2009, and 2010.

Professional development for the control condition

Control teachers were asked to identify professional development activities in which they participated without specific reference to QTEL. The survey response rate for control teachers was 75 percent in Year 2 and 68 percent in Year 3. Respondents were allowed to select as many professional development activities as applicable. Differentiated Instruction was the most frequently cited training in both years (33.6 percent in Year 2 and 38.9 percent in Year 3), followed by Specially Designed Academic Instruction in English (SDAIE) (28.3 percent in Year 2 and 31.6 percent in Year 3), High Point (12.7 percent in Year 2 and 14.7 percent in Year 3); and Cross-Cultural, Language, and Academic Development (10.9 percent in Year 2 and 13.7 percent in Year 3; table 3.7). In addition, 22.7 percent of responses in Year 2 and 27.4 percent in Year 3 referred to other activities identified by the control teachers, including Guided Language Acquisition Design, Structured English Immersion, department trainings, self-initiated learning to improve teaching, and collaborative (informal) professional development initiated by colleagues.

Table 3.7. Professional development activities identified by control-group teachers, 2008/09 and 2009/10 (percent)

Professional development activity	Year 2	Year 3
Differentiated Instruction	33.6	38.9
Specially Designed Academic Instruction in English	28.2	31.6
High Point	12.7	14.7
Cross-cultural, Language, and Academic Development	10.9	13.7
Other[a]	22.7	27.4
Sample size	110	95

Note: Response rate = 75 percent in Year 2 and 68 percent in Year 3. Percentages may not sum to 100 percent because respondents were allowed to report participating in more than one professional development activity.
a. Includes Guided Language Acquisition Design, English language development curriculum, explicit English language development for elementary and secondary students, Senate Bill 472 language arts, Structured Instruction Observation Protocol, Structured English Immersion instructional design, and Thinking Maps.
Source: Berkley Policy Associates' analysis of primary data from teacher surveys.

Teacher participation in QTEL

QTEL participation was assessed based on teacher involvement in each component: Building the Base summer institutes, coaching cycles, and lesson design meetings. Participation rates were calculated to reflect the percentages of targeted teachers reached by the intervention.

Building the Base

Participation rates for Building the Base varied across districts and years. These rates were calculated using attendance sheets collected by the QTEL team and submitted to the research team. Because some attendance sheets were missing and individual teachers might have failed to sign in, these rates represent the minimum number of teachers who attended the institutes. The summer institutes were intended for English language arts and English language development

teachers in grades 6 and 7 in Year 1, grades 7 and 8 in Year 2, and grades 7 and 8 in Year 3. Participants were considered to have attended a Building the Base institute if they attended any portion of the multiday summer institutes.[40] In Year 1, 70.2 percent of target teachers attended at least one day of Building the Base; in Year 2, 50.6 percent of target teachers attended; and in Year 3, 37.6 percent of teachers attended (table 3.8). Possible explanations for the decline in participation in Years 2 and 3 include contextual factors, such as district budget crises, structural changes in schools and districts, and staffing uncertainties and layoffs (see "Implementation contexts and experiences").

Table 3.8. Target teachers who attended a Building the Base summer institute, 2007/08, 2008/09, and 2009/10

Year	*n*	Percent
Year 1	181	70.2
Year 2	120	50.6
Year 3	56	37.6

Note: Data for Year 3 exclude District 2.
Source: Authors' analysis of primary data from summer institute attendance sheets submitted.

Coaching

All districts were involved in coaching activities to some degree. To adhere to the study design, the coaching component targeted grade 6 teachers[41] in Year 1, grade 7 teachers in Year 2, and grade 8 teachers in Year 3. QTEL is intended to provide four to six cycles of coaching to each teacher. In Year 1, 8.5 percent of target teachers participated in one cycle of coaching, 10.2 percent participated in two cycles, 10.7 percent participated in three cycles, 26 percent participated in four or more cycles, and 44.6 percent participated in less than one cycle of coaching (table 3.9). This could mean that a teacher received a coaching session, but did not participate in the follow up observation and debrief (a complete coaching cycle). It also could indicate that the eligible teacher did not participate in any coaching sessions. In Year 2, 9.9 percent participated in one cycle, 6.6 percent participated in two, 9.9 percent participated in three, 17.8 percent participated in at least four cycles, and 48 percent participated in less than one full cycle. In Year 3, 11.4 percent participated in one cycle, 3.4 percent participated in two, 18.8 percent participated in three, 12.8 participated in at least four, and 53.4 percent participated in less than one full cycle.

Actual participation rates may be higher than those reflected in table 3.9. The information was derived from records submitted by coaches. Because some coaching records are missing, the data represent minimum participation rates.

---

[40] A limitation of these data is that they were not collected for each day of the summer institutes due to several factors: sign-in sheets for some days were missing, institutes ranged in length, and it was often difficult to glean details from the attendance sheets collected (for example, which day or part of the institute they referred to). Therefore, the research team was unable to determine how many (and which) days a teacher attended, only that the teacher had attended at least one day of a given institute.

[41] Grade 7 teachers were targeted in schools that did not serve grade 6.

Table 3.9. Target teachers who received coaching, 2007/08, 2008/09, and 2009/10

Year	Less than one cycle of coaching		One cycle		Two cycles		Three cycles		At least four cycles	
	n	Percent	*n*	Percent	*n*	Percent	*n*	Percent	*n*	Percent
Year 1	79	44.6	15	8.5	18	10.2	19	10.7	46	26.0
Year 2	85	48.0	15	9.9	10	6.6	15	9.9	27	17.8
Year 3[a]	80	53.7	17	11.4	5	3.4	28	18.8	19	12.8

Note: Due to missing data, these figures represent a minimum estimate of coaching dosage.
a. Data in this row exclude District 2 data; the Year 3 roster of eligible teachers was not submitted by District 2 intervention schools, so these eligible teachers were not counted.
Source: Authors' analysis of primary data from coaching records submitted.

Lesson design meetings

Lesson design meetings were open to all teachers at all intervention schools, regardless of grade level or content area. The percentages of English language arts and English language development teachers in any grade level who participated in the meetings are in table 3.10. As with the other two major components, participation in lesson design meetings varied considerably across districts and years, and some data are missing due to attendance sheets not being submitted. For example, in District 4, lesson design meetings occurred and were observed in Year 1, but because of the lack of documentation, the number of teachers who attended selected lesson design meetings cannot be tabulated. For these reasons, these percentages should be interpreted with caution.

In Year 1, 32.5 percent of target teachers attended at least one lesson design meeting; in Year 2, 24.1 percent of teachers attended at least one; and in Year 3, 20.9 percent of teachers attended at least one.

Table 3.10. English language arts and English language development teachers participating in after-school lesson design meetings, 2007/08, 2008/09, and 2009/10

Year	*n*	Percent
Year 1	110	32.5
Year 2	81	24.1
Year 3	68	20.9

Note: Documentation of lesson design meetings is missing for District 4, Years 1 and 3; Data for Year 3 exclude District 2.
Source: Authors' analysis of primary data from lesson design meeting attendance sheets submitted.

Conclusion

This chapter described the intended QTEL design and its actual implementation during the study. QTEL was implemented in contexts that included natural disasters, a state budget crisis, and, to some extent, shifting school calendars and other reforms. Teacher participation in QTEL varied by district and by year, and participation data were incomplete. Overall, teacher participation in QTEL was low. Yet, of those who responded to the survey, intervention teachers reported receiving significantly more professional development in topics related to QTEL during the study than did control teachers. Impacts related to teacher outcomes are examined in chapter 4,

which presents the results of the impact analyses for students and teachers who participated in this study's QTEL model.

Chapter 4. Impact results

This chapter describes the impacts of QTEL on the English language arts and English language development outcomes of grade 7 and grade 8 students in the study schools. These outcomes are measured with the California Standards Test of English Language Arts (CST-ELA) and the California English Language Development Test (CELDT). The California Standards Test is administered every spring to all California public school students in grades 3–11. The CELDT is administered each October to students determined to be limited English proficient by their districts in the previous school year and to new students whose primary language is not English. The analyses in this chapter include all students who were tested in the study schools, regardless of how long they were enrolled in the school in which they were tested and regardless of whether their teachers participated in QTEL. The only exception is that analyses of the English language development outcomes are limited to students enrolled in the same school the year before and classified that year as limited English proficient.

This chapter also presents estimated impacts on teacher knowledge, teacher attitudes, and teacher practice. Outcome data for estimating these teacher impacts were collected using a teacher knowledge test, a teacher survey, and the Sheltered Instruction Observation Protocol.

The tables in this chapter show the standard error, the p-value, and an adjusted p-value for each impact estimate, to account for multiple comparisons. The standard error is the magnitude of the uncertainty about the true mean of each impact, given the number of schools and students in the analysis. The adjusted p-values take into account the number of statistical tests reported and indicate the chance of obtaining an impact as large as the estimated impact, if there was no true impact. The method used to make this adjustment was the Benjamini and Hochberg (1995) procedure. Impact estimates in these tables are considered statistically significant if the adjusted p-value is less than .05 (benchmark p-value), indicating that, taking into account the number of statistical tests, there would be less than a 5 percent chance of obtaining such an impact if there were no true effect. All estimates in this chapter are based on multilevel models that acknowledge the clustered data structures underlying the impact analyses. The two-level impact models that were used are described in detail in chapter 2.

No analyses in this chapter show significant impacts of QTEL on student or teacher outcomes. The rest of the chapter describes the student-level and teacher-level impact results in detail.

Impacts on student achievement

This section addresses the six primary research questions.

English language arts test scores

The regression-adjusted group means for the 2010 CST-ELA scale scores for the grade 7 and grade 8 students in the outcome sample, by experimental condition, are in table 4.1. These data address two primary research questions on the impact of QTEL on student achievement. The results represent the effects of offering QTEL to all English language arts teachers who taught grade 7 or grade 8 students. In both grades, the intervention students (adjusted p-value =.944)

did not receive scores that were statistically significantly different from those of the control students (adjusted *p*-value = .795). These *p*-values were adjusted across the six student-level research questions using the Benjamini and Hochberg (1995) procedure, to account for multiple comparisons. For grade 7 students, the adjusted mean score was 352.97 for the intervention group and 353.83 for the control group; for grade 8 students, it was 355.24 for the intervention group and 354.78 for the control group. The associated effect sizes (the intervention–control difference divided by the standard deviation of the outcome for the control group) were –0.01 and 0.01, respectively. Based on the 2010 CST-ELA, there was no evidence that QTEL improved student achievement for students in these two grades.

Table 4.1. Impact analysis of student achievement on the California Standards Test of English Language Arts (CST-ELA), grades 7 and 8, tested in spring 2010

	Adjusted means					
Sample	**Intervention (standard deviation)**	**Control[a] (standard deviation)**	**Difference (standard error)**	***p*-value (adjusted *p*-value)[b]**	**Effect size**	**Unweighted student sample size**
Grade 7 students	352.97 (57.01)	353.83 (57.53)	–0.86 (2.70)	.787 (.944)	–0.01	17,837
Grade 8 students	355.24 (59.10)	354.78 (59.45)	0.46 (2.20)	.795 (.795)	0.01	18,180

Note: Data were regression-adjusted using multilevel regression models to account for differences in baseline characteristics and study design characteristics. Effect sizes were calculated by dividing impact estimates by the control-group standard deviation of the outcome variable.
a. Includes one-third of the students of one school that was treated as an intervention school by QTEL. These sample members were assigned to the control group for analytical purposes because their school was consolidated from three smaller schools, one originally assigned to the control group.
b. The Benjamini and Hochberg (1995) method was used to calculate adjusted *p*-values across the six student achievement analyses.
Source: Authors' analysis of student-level data for participating districts.

English language arts test scores for English language learner students

QTEL focuses on the quality of instruction provided to English language learner students. The regression-adjusted group means for the CST-ELA scale scores are displayed, again by condition, in table 4.2, but in this table the sample is limited to students classified limited English proficient or redesignated fluent English proficient when the test was administered (spring 2010). The latter are students once classified as limited English proficient, but whose achievement on the CELDT and other academic criteria redesignated them as fluent English proficient. Together, the limited English proficient and redesignated fluent English proficient students are labeled as English language learner students (limited English proficient and redesignated fluent English proficient).

Table 4.2. Impact analysis of student achievement on the California Standards Test of English Language Arts (CST-ELA), grade 7 and grade 8 limited English proficient and redesignated fluent English proficient students, tested in spring 2010

| | Adjusted means | | | | | |
Sample	Intervention (standard deviation)	Control[a] (standard deviation)	Difference (standard error)	p-value (adjusted p-value)[b]	Effect size	Unweighted student sample size
Grade 7 students classified as limited English proficient and redesignated fluent English proficient in 2008/09	334.32 (52.86)	332.78 (52.24)	1.54 (3.02)	.748 (1.000)	0.03	7,699
Grade 8 students classified as limited English proficient and redesignated fluent English proficient in 2008/09	337.03 (53.92)	336.31 (54.48)	0.72 (2.86)	.666 (1.000)	0.01	8,098

Note: Data were regression-adjusted using multilevel regression models to account for differences in baseline characteristics and study design characteristics. Effect sizes were calculated by dividing impact estimates by the control-group standard deviation of the outcome variable.
a. Includes a third of the students of one school that was treated as an intervention school by QTEL.
b. The Benjamini and Hochberg (1995) method was used to calculate adjusted *p*-values across the six student achievement analyses. Adjusted *p*-values were capped at 1.000.
Source: Authors' analysis of student-level data for participating districts.

The difference in CST-ELA scores between intervention and control English language learner students was 0.03 standard deviation for grade 7 students and 0.01 standard deviation for grade 8 students. For both grades, the scores for intervention English language learner students were not statistically significantly different from those of their control-group counterparts (adjusted $p = 1.000$ for both grades). Thus, there is no evidence that QTEL improved English language arts achievement for English language learner students in either grade.

California English Language Development Test (CELDT) scores

QTEL was hypothesized to improve English language proficiency among English language learner students classified as limited English proficient. Such classification occurs at the beginning of each academic year (in October) and is based on students' performance on the CELDT and other factors, as discussed in chapter 2. Estimated impacts on the CELDT scale scores for grade 6 and grade 7 students in the 2009/10 school year are shown in table 4.3. The students were tested in October 2009 and had been classified as limited English proficient in the prior academic year (2008/09). Thus, the test captures the impact of their exposure to QTEL in grade 6 or grade 7 during the 2008/09 school year, when they qualified for federally funded, specialized English language development instruction, provided by teachers who (in the intervention schools) had been offered the opportunity to participate in QTEL. The scale scores on the test for limited English proficient students in the intervention group were not statistically significantly different from those of their counterparts in the control group (adjusted $p = 3.660$ for grade 6 students and $p = 1.000$ for and grade 7 students). The difference between the intervention group's CELDT scores and those of the control group was 0.05 standard deviation for grade 6 students and 0.03 standard deviation for grade 7 students.

Table 4.3. Impact analysis of student achievement on the California English Language Development Test (CELDT) for grade 6 and grade 7 limited English proficient students identified in the 2008/09 academic year, tested in fall 2009

| Sample | Adjusted means | | Difference (standard error) | *p*-value (adjusted *p*-value)[b] | Effect size | Unweighted student sample size |
	Intervention (standard deviation)	Control[a] (standard deviation)				
Grade 6 limited English proficient students (tested fall Grade 7)	550.90 (59.74)	547.60 (66.07)	3.30 (4.64)	.610 (1.000)	0.05	2,373
Grade 7 limited English proficient students (tested fall Grade 8)	558.62 (70.13)	555.29 (69.69)	2.33 (6.27)	.612 (1.000)	0.03	3,456

Note: Data were regression-adjusted using multilevel regression models to account for differences in baseline characteristics and study design characteristics. Effect sizes were calculated by dividing impact estimates by the control-group standard deviation of the outcome variable.

a. Includes a third of the students of one school that was treated as an intervention school by QTEL. These sample members were assigned to the control group for analytical purposes because their school was consolidated from three smaller schools, one originally assigned to the control group.

b. The Benjamini and Hochberg (1995) method was used to calculate adjusted *p*-values across the six student achievement analyses. Adjusted p-values were capped at 1.000.

Source: Authors' analysis of student-level data for participating districts.

Secondary impacts on teacher-level outcomes

This section addresses the three secondary research questions: on teacher knowledge, teacher attitudes, and teacher practice.

Teacher instructional knowledge

The results of analyses to measure the impact of QTEL on teachers' instructional knowledge, as measured with a 40-item teacher knowledge test, are shown in table 4.4. Outcome data were collected from teachers who taught grade 7 in 2009 and from teachers who taught grade 8 in 2010.[42] The table shows, by experimental condition, the regression-adjusted group mean of the teacher knowledge test scores for all teachers across the two years. The difference in teacher knowledge test scores for teachers in intervention schools (adjusted mean score = 22.60) and teachers in control schools (adjusted mean score = 21.21) approached statistical significance at the pooled analysis across both years (adjusted *p*-value =.051). The difference of 1.39 translated into an effect size of 0.33. The *p*-values for the three teacher-level outcomes were adjusted for multiple comparisons using the Benjamini and Hochberg (1995) procedure.

---

[42] Teacher instructional knowledge data were also collected in 2008 but, as a result of an administrative error, the research team was not able to link the teacher-level Year 1 data from the teacher knowledge test to teacher identifiers and grade levels. Therefore, only data from Years 2 and 3 for the teacher knowledge test are presented.

Table 4.4. Impact analysis of teacher outcome measures, for grade 6 teachers in spring 2008, grade 7 teachers in 2009, and grade 8 teachers in 2010

Impact measure and sample	Adjusted means		Difference (standard error)	p-value (adjusted p-value)[b]	Effect size	Unweighted teacher sample size
	Intervention (standard deviation)	Control[a] (standard deviation)				
Total teacher knowledge score (all surveyed teachers, Years 2–3)	22.60 (5.20)	21.21 (4.27)	1.39 (0.58)	.017 (.051)	0.33	404
Average teacher attitude score (all surveyed teachers, Years 1–3)	3.14 (0.37)	3.08 (0.38)	0.06 (0.05)	.218 (.436)	0.16	623
SIOP average score (all observed teachers, Years 1–3)	2.45 (0.68)	2.48 (0.65)	–0.04 (0.10)	.711 (.711)	–0.06	527

Note: Data were regression-adjusted using multilevel regression models to account for differences in baseline characteristics and study design characteristics. Effect sizes were calculated by dividing impact estimates by the control-group standard deviation of the outcome variable. Standard errors for the multiyear sample were adjusted for multiple teacher responses across survey years using a robust cluster variance estimator.
a. Includes a third of the teachers of one school that was treated as an intervention school by QTEL. These sample members were assigned to the control group for analytical purposes because their school was consolidated from three smaller schools, one originally assigned to the control group.
b. The Benjamini and Hochberg (1995) method was used to calculate adjusted *p*-values across the three teacher outcome analyses.
Source: Authors' analysis of teacher-level data collected for study.

Teacher attitudes toward English language learner students

The results in table 4.4 also reflect the estimated impacts of QTEL on teachers' attitudes toward English language learner students, as measured by the teacher survey introduced in chapter 2. The overall outcome measure for teacher attitudes was calculated as the average teacher agreement with eight statements about English language learner students along a four-point Likert scale from (1) strongly disagree to (4) strongly agree. Teacher attitude outcome data were collected from teachers who taught grade 6 or grade 7 alone or with other grades in spring 2008,[43] from teachers who taught grade 7 in spring 2009, and from teachers who taught grade 8 in spring 2010.

There were no statistically significant differences between the scores of intervention teachers and those of control teachers across all the years. The intervention teachers had an adjusted mean score of 3.14; the control teachers, 3.08 (adjusted *p* = .436). The difference of 0.06 translated to an effect size of 0.16, which was not statistically significant.

---

[43] Target teachers in 2008 included teachers who taught grade 6 and other grades at schools that included grade 6 (grades 6–8 middle schools) and teachers who taught grade 7 and other grades at schools that did not include grade 6 (grades 7–8 middle schools).

Teacher practice

Observation outcome data were collected from teachers who taught grade 6 or grade 7 alone or with other grades in spring 2008,[44] from teachers who taught grade 7 in spring 2009, and from teachers who taught grade 8 in spring 2010. There was no impact of QTEL on the Sheltered Instruction Observation Protocol (SIOP) average score across all the years combined (adjusted $p = .711$). The difference of –0.04 translated to an effect size of –0.06, which was not statistically significant.

Results of sensitivity analyses

A number of sensitivity analyses were conducted to determine how much the impact estimates in tables 4.1–4.4 depended on the assumptions made. As detailed in appendix E, in most cases the direction, magnitude, and statistical significance of the estimated intervention effects did not change considerably whether: listwise deletion was used for missing baseline, SIOP, and teacher knowledge test data; the 12 schools with no grade 6 were excluded from all student-level analyses; students were only included if they had been in the school in which they were tested in 2010 for their entire middle school career; the consolidation of three schools (two intervention and one control) during Year 2 was treated as one intervention school; a Rasch model was used to construct the teacher attitudes outcome; or teachers who received more than one year of the intervention were dropped from the analyses.

---

[44] Target teachers in 2008 included teachers who taught grade 6 and other grades at schools that included grade 6 (grades 6–8 middle schools) and teachers who taught grade 7 and other grades at schools that did not include grade 6 (grades 7–8 middle schools).

Chapter 5. Exploratory analyses

This chapter presents the results of the exploratory analyses conducted to supplement the impact estimates. These exploratory analyses were intended to capture variation in QTEL effects across different groups of students, teachers, and schools and to describe how key outcome measures varied across different levels of exposure to the kinds of QTEL professional development. The analyses include several outcome measures, referenced in "Exploratory analysis," that were not included in the previous impact analyses.

The purpose of these analyses is not to change or directly contribute to the overall conclusions about the effectiveness of QTEL. Those conclusions are based on the confirmatory impact analyses. Instead, the findings in this chapter will inform the development of new hypotheses about QTEL and similar teacher professional development efforts and might lead to additional confirmatory research in future evaluation efforts.

Subgroup analyses by English language learner student status

Chapter 4 presented estimated impacts of QTEL on the California Standards Test of English Language Arts (CST-ELA) for students in grades 7 and 8 for all students taking the test and for English language learner students (defined as either redesignated fluent English proficient or limited English proficient) taking the test. The latter group is a subsample of the full sample of test-taking students. This analysis further disaggregates these impacts by students' English language learner status.

The regression-adjusted group means for the 2010 CST-ELA scale score for the grade 7 and grade 8 students in four different subgroups of students in the outcome sample are displayed, by experimental condition, in tables 5.1–5.4. These four subgroups of students are English only, initially fluent English proficient, redesignated fluent English proficient, and limited English proficient. English-only students were those who reported speaking only English at home. Initially fluent English proficient students reported not speaking English at home but were determined to be fluent in English the first time their English language proficiency was tested by the school district. Redesignated fluent English proficient students were determined to be limited English proficient at one point but had since been reclassified as speaking English fluently. Limited English proficient students were those whose English proficiency was (or continued to be) limited. There are no statistically significant impacts of QTEL on the English language arts achievement of any of these four groups in either grade 7 or grade 8. However, the differences among the estimated impacts across these four groups were statistically significant when tested as a group using a joint F-test ($p=.001$). Such a test does not reveal which individual impacts differ significantly from one another but it supports the conclusion that the estimated impacts varied across this subgroup dimension.

Table 5.1. Exploratory analysis of student achievement on the California Standards Test of English Language Arts (CST-ELA), grade 7 and grade 8 English-only students, spring 2010

| | Adjusted means | | | | | |
Impact measure and sample	Intervention (standard deviation)	Control[a] (standard deviation)	Difference (standard error)	*p*-value	Effect size	Unweighted student sample size
CST-ELA for grade 7 English-only students	365.46 (56.09)	368.69 (57.17)	–3.23 (4.28)	.562	–0.06	8,602
CST-ELA for grade 8 English-only students	368.42 (59.33)	367.60 (59.22)	0.83 (3.37)	.534	0.01	8,559

Note: Data were regression-adjusted using multilevel regression models to account for differences in baseline characteristics and study design characteristics. Effect sizes were calculated by dividing impact estimates by the control-group standard deviation of the outcome variable.

a. Includes one-third of the students of one school that was treated as an intervention school by QTEL. These sample members were assigned to the control group for analytical purposes because their school was consolidated from three smaller schools, one originally assigned to the control group.

Source: Authors' analysis of student-level data from participating districts.

Table 5.2. Exploratory analysis of student achievement on the California Standards Test of English Language Arts (CST-ELA), grade 7 and grade 8 initially fluent English proficient students, spring 2010

| | Adjusted means | | | | | |
Impact measure and sample	Intervention (standard deviation)	Control[a] (standard deviation)	Difference (standard error)	*p*-value	Effect size	Unweighted student sample size
CST-ELA for grade 7 initially fluent English proficient students	378.59 (51.17)	374.01 (53.18)	4.57 (3.61)	.316	0.09	1,534
CST-ELA for grade 8 initially fluent English proficient students	376.31 (56.44)	382.95 (57.50)	–6.64 (4.21)	.126	-0.12	1,520

Note: Data were regression-adjusted using multilevel regression models to account for differences in baseline characteristics and study design characteristics. Effect sizes were calculated by dividing impact estimates by the control-group standard deviation of the outcome variable.

a. Includes one-third of the students of one school that was treated as an intervention school by QTEL. These sample members were assigned to the control group for analytical purposes because their school was consolidated from three smaller schools, one originally assigned to the control group.

Source: Authors' analysis of student-level data from participating districts.

Table 5.3. Exploratory analysis of student achievement on the California Standards Test of English Language Arts (CST-ELA), grade 7 and grade 8 redesignated fluent English proficient students, spring 2010

| Impact measure and sample | Adjusted means | | Difference (standard error) | *p*-value | Effect size | Unweighted student sample size |
	Intervention (standard deviation)	Control[a] (standard deviation)				
CST-ELA for grade 7 redesignated fluent English proficient students	362.43 (41.68)	361.53 (40.26)	0.90 (2.03)	.672	0.02	4,541
CST-ELA for grade 8 redesignated English proficient students	364.17 (44.41)	364.83 (44.79)	–0.66 (2.11)	.828	-0.01	4,905

Note: Data were regression adjusted using multilevel regression models to account for differences in baseline characteristics and study design characteristics. Effect sizes were calculated by dividing impact estimates by the control-group standard deviation of the outcome variable. An *F*-test was applied to differences between the impacts for the subgroups. Statistical significance levels are indicated as ** =1 percent and *= 5 percent.
a. Includes one-third of the students of one school that was treated as an intervention school by QTEL. These sample members were assigned to the control group for analytical purposes because their school was consolidated from three smaller schools, one originally assigned to the control group.
Source: Authors' analysis of student-level data from participating districts.

Table 5.4. Exploratory analysis of student achievement on the California Standards Test of English Language Arts (CST-ELA), grade 7 and grade 8 limited English proficient students, spring 2010

| Impact measure and sample | Adjusted means | | Difference (standard error) | *p*-value | Effect size | Unweighted student sample size |
	Intervention (standard deviation)	Control[a] (standard deviation)				
CST-ELA for grade 7 limited English proficient students	293.94 (38.30)	291.41 (38.78)	2.53 (2.76)	.547	0.07	3,158
CST-ELA for grade 8 limited English proficient students	295.43 (37.13)	292.45 (37.28)	3.00 (2.36)	.193	0.08	3,193

Note: Data were regression-adjusted using multilevel regression models to account for differences in baseline characteristics and study design characteristics. Effect sizes were calculated by dividing impact estimates by the control-group standard deviation of the outcome variable.
a. Includes one-third of the students of one school that was treated as an intervention school by QTEL. These sample members were assigned to the control group for analytical purposes because their school was consolidated from three smaller schools, one originally assigned to the control group.
Source: Authors' analysis of student-level data from participating districts.

Subgroup analyses by teacher background characteristics

The full-sample impact estimates implicitly assumed that the impact of QTEL on students and teachers would be consistent across different subgroups of teachers. However, it is possible that teachers with more or less experience or with higher or lower levels of education may have experienced QTEL differently, causing its impacts on their students to vary as well. As discussed in chapter 2, the research team did not have access to reliable student/teacher linkage data in all

the school districts. To examine whether the impacts varied by teacher background, the researchers created subgroup breakdowns at the school level, designating schools with higher and lower proportions of teachers across all subjects who had advanced degrees (master's degrees or doctorates) and schools with higher and lower proportions of experienced teachers across all subjects (more or less than an average of 12.8 years of teaching experience).[45] Tables 5.5–5.8 present regression-adjusted group means for the 2010 CST-ELA scale scores for the grade 7 and grade 8 students in the outcome sample in four different subgroups based on the school-level average of the background characteristics of the teachers in their schools at baseline: those whose schools had more or less than 43.5 percent of teachers with an advanced degree and those whose schools had more or less than an average of 12.8 years of teaching experience. Both these cutpoints represented the average value of the aggregate teacher characteristics.

QTEL was estimated to have increased the test scores of grade 8 students in schools where more than 43.5 percent of the teachers have an advanced degree (the top half of the distribution of schools on this teacher characteristic; table 5.5). The difference of 10.40 points translated to an effect size of 0.17 of a standard deviation. This impact estimate was statistically significant ($p = .027$) and statistically significantly different from the estimated QTEL impact on the test scores of grade 8 students in schools with fewer highly educated teachers (effect size = 0.01, $p = .167$). No other impact estimates in tables 5.1–5.4 were statistically significant. However, the difference between the point estimates of the impacts on grade 7 English language arts scores disaggregated by teacher experience was statistically significant, as shown in table 5.8 ($p = .001$).

---

[45] The cutoff of 12.8 years for teacher experience is based on the sample data and not on a scientific review of expected variation in teacher effectiveness by years of experience. Such a review was outside the scope of these analyses and would have resulted in a range of possible cutoff values, all of which would have translated to different subgroup breakdowns. The subgroup analysis here explores variation in QTEL effectiveness around the school-level distribution of teacher experience as it manifests itself in this particular sample.

Table 5.5. Impact analysis of student achievement on the California Standards Test of English Language Arts (CST-ELA), grade 8 in spring 2010, by level of postgraduate attainment of schools' teachers at school level

| Impact measure and sample | Adjusted means | | Difference (standard error) | *p*-value (for difference between impacts) | Effect size | Unweighted student sample size |
	Intervention (standard deviation)	Control[a] (standard deviation)				
43.5 percent or more teachers with a postgraduate degree						
CST-ELA for grade 8 students	363.09 (59.10)	352.69 (61.96)	10.40* (3.42)	.027 (.009)**	0.17	7,865
Less than 43.5 percent of teachers with a postgraduate degree						
CST-ELA for grade 8 students	352.33 (58.90)	351.95 (58.01)	0.38 (3.45)	.167 (.009)**	0.01	10,049

Note: Data were regression adjusted using multilevel regression models to account for differences in baseline characteristics and study design characteristics. Effect sizes were calculated by dividing impact estimates by the control-group standard deviation of the outcome variable. A *t*-test was applied to the impact estimates to test their statistical significance. Statistically significant impact estimates are flagged with their level of statistical significance as follows: ** = 1 percent and * = 5 percent. An *F*-test was applied to test the significance of differences between the impacts for the two subgroups shown. Statistical significance levels for this test, whose p-values are shown in parentheses, are indicated as ** = 1 percent and
* = 5 percent.
a. Includes one-third of the students of one school that was treated as an intervention school by QTEL. These sample members were assigned to the control group for analytical purposes because their school was consolidated from three smaller schools, one originally assigned to the control group.
Source: Authors' analysis of student-level data from participating districts.

Table 5.6. Impact analysis of student achievement on the California Standards Test of English Language Arts (CST-ELA), grade 7 in spring 2010, by level of postgraduate attainment of schools' teachers at school level

| Impact measure and sample | Adjusted means | | Difference (standard error) | *p*-value (for difference between impacts) | Effect size | Unweighted student sample size |
	Intervention (standard deviation)	Control[a] (standard deviation)				
43.5 percent or more teachers with a postgraduate degree						
CST-ELA for grade 7 students	357.17 (58.31)	353.46 (61.24)	3.71 (2.61)	0.227 (0.560)	0.06	7,708
Less than 43.5 percent of teachers with a postgraduate degree						
CST-ELA for grade 7 students	351.23 (55.49)	352.46 (55.33)	3.26 (–3.22)	0.655 (0.560)	–0.02	9,881

Note: Data were regression-adjusted using multilevel regression models to account for differences in baseline characteristics and study design characteristics. Effect sizes were calculated by dividing impact estimates by the control-group standard deviation of the outcome variable. An *F*-test was applied to test the significance of differences between the impacts for the two subgroups shown. Statistical significance levels for this test, whose p-values are shown in parentheses, are indicated as ** = 1 percent and
 * = 5 percent.
a. Includes one-third of the students of one school that was treated as an intervention school by QTEL. These sample members were assigned to the control group for analytical purposes because their school was consolidated from three smaller schools, one originally assigned to the control group.
Source: Authors' analysis of student-level data from participating districts.

Table 5.7. Impact analysis of student achievement on the California Standards Test of English Language Arts (CST-ELA), grade 8 in spring 2010, based on teachers' years of experience at school level

Impact measure and sample	Adjusted means		Difference (standard error)	*p*-value (for difference between impacts)	Effect size	Unweighted student sample size
	Intervention (standard deviation)	Control[a] (standard deviation)				
More than 12.8 years of experience (average)						
CST-ELA for grade 8 students	361.39 (59.97)	356.84 (62.69)	4.55 (3.55)	.420 (.341)	0.07	8,645
Less than 12.8 years of experience (average)						
CST-ELA for grade 8 students	352.68 (56.98)	350.15 (57.24)	2.52 (4.15)	.520 (.341)	0.04	9,269

Note: Data were regression-adjusted using multilevel regression models to account for differences in baseline characteristics and study design characteristics. Effect sizes were calculated by dividing impact estimates by the control-group standard deviation of the outcome variable. An *F*-test was applied to test the significance of differences between the impacts for the two subgroups shown. Statistical significance levels for this test, whose p-values are shown in parentheses, are indicated as ** = 1 percent and * = 5 percent.

a. Includes one-third of the students of one school that was treated as an intervention school by QTEL. These sample members were assigned to the control group for analytical purposes because their school was consolidated from three smaller schools, one originally assigned to the control group.

Source: Authors' analysis of student-level data from participating districts.

Table 5.8. Impact analysis of student achievement on the California Standards Test of English Language Arts (CST-ELA), grade 7 in spring 2010, based on teachers' level of experience at school level

Impact measure and sample	Adjusted means		Difference (standard error)	*p*-value (for difference between impacts)	Effect size	Unweighted student sample size
	Intervention (standard deviation)	Control[a] (standard deviation)				
More than 12.8 years of experience (average)						
CST-ELA for grade 7 students	359.69 (58.29)	356.24 (60.93)	3.45 (3.54)	.315 (.001)**	0.06	8,702
Less than 12.8 years of experience (average)						
CST-ELA for grade 7 students	347.32 (54.03)	350.44 (54.78)	–3.12 (3.77)	.245 (.001)**	–0.06	8,887

Note: Data were regression-adjusted using multilevel regression models to account for differences in baseline characteristics and study design characteristics. Effect sizes were calculated by dividing impact estimates by the control-group standard deviation of the outcome variable. An *F*-test was applied to test the significance of differences between the impacts for the two subgroups shown. Statistical significance levels for this test, whose p-values are shown in parentheses, are indicated as ** = 1 percent and * = 5 percent.

a. Includes one-third of the students of one school that was treated as an intervention school by QTEL. These sample members were assigned to the control group for analytical purposes because their school was consolidated from three smaller schools, one originally assigned to the control group.

Source: Authors' analysis of student-level data from participating districts.

Subgroup impacts by baseline classroom quality

Another dimension along which the impact of QTEL might have varied is the extent to which classrooms in participating schools already had a high degree of quality at the time of random assignment. As chapter 4 discussed, QTEL was hypothesized to improve teacher practice and classroom quality—potential improvements documented with the Sheltered Instruction Observation Protocol, administered in all study schools at baseline and in Years 1, 2, and 3. For the exploratory analysis in tables 5.9 and 5.10, the researchers divided the sample of study schools into two subsamples based on the school-level average baseline total Sheltered Instruction Observation Protocol score (2.15). Such a full-sample mean provides an intuitive natural cutpoint for analyses like these, especially without other natural cutpoints in the data. (The one group of schools has scores higher than the sample average; the other has scores lower than the sample average.) The tables show, by experimental condition, the regression-adjusted group means for the 2010 CST-ELA scale scores for the grade 7 and grade 8 students in the outcome sample. These impact estimates are presented separately for each subgroup of schools, based on their baseline classroom quality as measured with the Sheltered Instruction Observation Protocol. The estimated impacts of QTEL on the English language arts achievement of these two subgroups did not differ significantly and none of the four impact estimates was statistically significant or exceeded 0.1 standard deviation in magnitude.

Table 5.9. Impact analysis of student achievement on the California Standards Test of English Language Arts (CST-ELA), grade 8 in spring 2010, based on baseline classroom Sheltered Instruction Observation Protocol score at school level

| | Adjusted means | | | | | |
Impact measure and sample	**Intervention (standard deviation)**	**Control[a] (standard deviation)**	**Difference (standard error)**	**p-value (for difference between impacts)**	**Effect size**	**Unweighted student sample size**
2.15 or greater (average baseline SIOP score)						
CST-ELA for grade 8 students	363.21 (58.96)	365.61 (58.98)	–2.41 (2.74)	.109 (.950)	–0.04	8,523
Less than 2.15 (average baseline SIOP score)						
CST-ELA for grade 8 students	349.70 (59.82)	349.63 (59.71)	0.07 (4.59)	.107 (.950)	0.00	6,283

SIOP is Sheltered Instruction Observation Protocol.
Note: Data were regression-adjusted using multilevel regression models to account for differences in baseline characteristics and study design characteristics. Effect sizes were calculated by dividing impact estimates by the control-group standard deviation of the outcome variable. An *F*-test was applied to test the significance of differences between the impacts for the two subgroups shown. Statistical significance levels for this test, whose p-values are shown in parentheses, are indicated as ** = 1 percent and * = 5 percent.
a. Includes one-third of the students of one school that was treated as an intervention school by QTEL. These sample members were assigned to the control group for analytical purposes because their school was consolidated from three smaller schools, one originally assigned to the control group.
Source: Authors' analysis of student-level data from participating districts

Table 5.10. Impact analysis of student achievement on the California Standards Test of English Language Arts (CST-ELA), based on classroom Sheltered Instruction Observation Protocol score, grade 7, spring 2010

Impact measure and sample	Adjusted means		Difference (standard error)	*p*-value (for difference between impacts)	Effect size	Unweighted student sample size
	Intervention (standard deviation)	Control[a] (standard deviation)				
2.15 or greater (average baseline SIOP score)						
CST-ELA for grade 7 students	363.43 (56.38)	362.50 (55.97)	0.93 (3.68)	.643 (.079)	0.02	8,252
Less than 2.15 (average baseline SIOP score)						
CST-ELA for grade 7students	349.49 (58.91)	345.84 (59.84)	3.65 (4.80)	.724 (.079)	0.06	6,186

SIOP is Sheltered Instruction Observation Protocol.

Note: Data were regression-adjusted using multilevel regression models to account for differences in baseline characteristics and study design characteristics. Effect sizes were calculated by dividing impact estimates by the control-group standard deviation of the outcome variable. An *F*-test was applied to test the significance of differences between the impacts for the two subgroups shown. Statistical significance levels for this test, whose p-values are shown in parentheses, are indicated as ** = 1 percent and * = 5 percent.

a. Includes one-third of the students of one school that was treated as an intervention school by QTEL. These sample members were assigned to the control group for analytical purposes because their school was consolidated from their smaller schools, one originally assigned to the control group.

Source: Authors' analysis of student-level data from participating districts

QTEL effects on different measures of classroom quality

The only measure of classroom quality included in chapter 4 was the total score on the Sheltered Instruction Observation Protocol (SIOP). In this section, this measure of classroom quality is supplemented with eight additional measures, four of which are subscales of the 30-item SIOP, capturing lesson preparation (9 items), input and interactions (7 items), lesson activity (6 items), and lesson delivery (8 items). QTEL might have impacted one or more subscales without affecting the others, and positive effects on classroom quality in one dimension might have been offset by negative effects in another. The regression-adjusted group mean scale scores on the SIOP for all observed teachers across the three years for the full 30-item scale and for the four subscales are shown, by experimental condition, in table 5.11. QTEL did not have any statistically significant impacts on any of the SIOP subscales. Some reported effects were positive and others were negative, but none were statistically significant.

Table 5.11. Impact analysis of teacher practice: Sheltered Instruction Observation Protocol (SIOP) average subscale scores, spring 2008, 2009, and 2010

| | Adjusted means | | | | | |
Impact measure	Intervention (standard deviation)	Control[a] (standard deviation)	Difference (standard error)	*p*-value	Effect size	Unweighted teacher sample size
Average total score	2.45 (0.68)	2.48 (0.65)	–0.04	.711	–0.06	527
Lesson preparation score	2.38 (0.73)	2.46 (0.67)	–0.08 (0.09)	.368	–0.12	527
Input and interaction score	2.72 (0.65)	2.67 (0.63)	0.05 (0.09)	.582	0.08	527
Lesson activity score	2.44 (0.81)	2.41 (0.78)	0.03 (0.11)	.748	0.04	527
Lesson delivery score	2.28 (0.78)	2.41 (0.75)	–0.12 (0.11)	.246	–0.16	527

Note: Data were regression-adjusted using multilevel regression models to account for differences in baseline characteristics and study design characteristics. Effect sizes were calculated by dividing impact estimates by the control-group standard deviation of the outcome variable. Standard errors for the multiyear sample were adjusted for multiple teacher responses across survey years using a robust cluster variance estimator.

a. Includes one-third of the teachers of one school that was treated as an intervention school by QTEL. These sample members were assigned to the control group for analytical purposes because their school was consolidated from their smaller schools, one originally assigned to the control group.

Source: Authors' analysis of primary observation data collected for study.

The research team created an additional measure of classroom quality more closely aligned with QTEL. This instrument, Program Aligned Classroom Observation (PACO), measures classroom quality along activity structure (11 items), lesson content (6 items), and student–student interaction (3 items). Regression-adjusted group mean scale scores for the full instrument average scores and for the three average subscale scores are presented, by experimental condition, in table 5.12. Among these four estimates, only the impact on student/student interaction subscale was statistically significant. This intervention–control difference translated to an effect size of 0.46 (*p* = .003).

Table 5.12. Impact analysis of teacher practice: Program Aligned Classroom Observation (PACO) average score and average subscale scores, spring 2009 and 2010

| | Adjusted means | | | | |
Impact measure	Intervention (standard deviation)	Control[a] (standard deviation)	Difference (standard error)	*p*-value	Effect size
Overall average score	2.19	2.15	0.04	.744	0.05
	(0.70)	(0.73)	(0.13)		
Activity structure subscale score	2.19	2.25	–0.06	.674	–0.08
	(0.72)	(0.77)	(0.13)		
Lesson content subscale score	2.35	2.28	0.07	.640	0.08
	(0.83)	(0.87)	(0.15)		
Interaction patterns subscale score	2.07	1.50	0.58*	.003	0.46
	(1.08)	(1.25)	(0.19)		
Unweighted teacher sample size (206)					

Note: Data were regression-adjusted using multilevel regression models to account for differences in baseline characteristics and study design characteristics. Effect sizes were calculated by dividing impact estimates by the control-group standard deviation of the outcome variable. Standard errors for the multiyear sample were adjusted for multiple teacher responses across survey years using a robust cluster variance estimator. Statistical significance levels are indicated as ** = 1 percent and * = 5 percent.
a. Includes a third of the teachers of one school that was treated as an intervention school by QTEL. These sample members were assigned to the control group for analytical purposes because their school was consolidated from their smaller schools, one originally assigned to the control group
Source: Authors' analysis of primary observation data collected for study.

Analyses of the association between teacher participation in professional development and student outcomes

The last section of this chapter describes how measures of teacher attitudes, teacher knowledge, classroom quality, and student achievement varied with different levels of teacher participation in professional development, as measured with teacher surveys administered during the three study years. This exploratory analysis is nonexperimental and not protected from potential selection bias by the evaluation's experimental study design. Teachers in intervention and control schools were asked, in each of the three follow-up surveys administered in 2007/08, 2008/09, and 2009/10, whether they received professional development in English language development standards and in instructional strategies for secondary-school English language learner students. Teachers were asked to indicate whether they received it the previous year and whether they received it for at least seven days or more. Combining these two variables, it is possible to capture exposure to professional development in these two areas relevant to the QTEL experience, as well as the intensity of such exposure.

English language arts and English language development teachers in intervention schools, due to their access to QTEL, were expected to have experienced intensive exposure to both these types of professional development. Through other programs and services, teachers in the control schools may have been similarly exposed.

The estimated effect on student achievement, as measured by the 2010 grade 7 and grade 8 CST-ELA, of each of the following four professional development variables: any participation in English language development standards; seven days or more participation in English language development standards; any participation in instructional strategies for English language learner students; and seven days or more participation in instructional strategies for English language learner students, is shown in table 5.13. This analysis combined the responses of teachers across intervention and control schools and did not control for the random assignment status or for teachers' exposure to QTEL. More teacher exposure to professional development in English language development standards and in instructional strategies for secondary English language learner students was not associated with different levels of student achievement, regardless of whether such exposure was to professional development in general or involved more intensive participation, and regardless of the focus of the professional development activities reported by teachers.

Table 5.13. Estimated effect of increasing teacher participation in instructional strategies and professional development in English language development standards by 10 percentage points: analysis of student achievement on the California Standards Test of English Language Arts (CST-ELA)

	Any participation			Participation for more than seven days		
Impact measure	Estimate	Standard error	*p*-value	Estimate	Standard error	*p*-value
Estimated effect of increasing teacher participation in instructional strategies by 10 percentage points						
CST-ELA for grade 7 students	–75.16	13.24	.575	171.31	10.64	.118
CST-ELA for grade 8 students	20.98	10.90	.849	–31.92	8.26	.702
Estimated effect of increasing teacher participation in professional development in English language development standards by 10 percentage points						
CST-ELA for grade 7 students	20.72	20.27	.919	–188.01	10.23	.076
CST-ELA for grade 8 students	–167.42	16.74	.325	104.54	8.16	.210

Note: Data were regression-adjusted using multilevel regression models to account for differences in baseline characteristics and study design characteristics. The control group in these analyses includes one-third of the students of one school that was treated as an intervention school by QTEL. These sample members were assigned to the control group for analytical purposes because their school was consolidated from three smaller schools, one originally assigned to the control group.
Source: Authors' analysis of student-level data for participating districts and primary data collected for study.

Chapter 6. Summary of findings

This study was the first randomized controlled trial to assess the impact of QTEL on middle school student achievement and teacher knowledge, attitudes, and instructional practices. The study was rigorous, sufficiently powered and designed as a cluster randomized effectiveness trial. As a result, the study generated statistically unbiased estimates of the effects of QTEL, implemented in naturalistic conditions,[46] on student achievement in English language arts and in English language development (for English language learner students). The study was implemented with an intent-to-treat design, which measures the effects of offering QTEL rather than the effects of actually participating in it. QTEL participation varied greatly across districts and years, with low participation possibly explaining the lack of significant outcomes observed. This chapter summarizes the findings, study limitations, implications of the study, and areas for future research.

Implementation study findings

For this study, QTEL varied from the original model. Typically, QTEL coaches work directly with teachers of all content areas in a school. Because of the number of sites required by the research design, the QTEL developers focused on a single academic department—English language arts—to ensure the capacity to deliver services to 26 middle schools simultaneously. In addition, QTEL coaches usually work with the same teachers over multiple years. But in this study, they generally worked with several teachers in a targeted grade level for one year, then moved on to a new group of teachers in the same schools in each subsequent year.

Participation rates varied across the different activities and years of the study. In Year 1, 70.2 percent of target teachers attended a Building the Base institute; in Year 2, 50.6 percent attended this institute; and in Year 3, 37.6 percent attended. In Year 1, 26 percent of target teachers received at least four cycles of coaching; in Year 2, 17.8 percent received at least four cycles, and in Year 3; 12.8 percent received at least four cycles. While lesson design meetings were open to all teachers, in Year 1, 32.5 percent of the English language arts and English language development teachers (in any grade level) attended them. In Year 2, 24.1 percent attended; in Year 3, 20.9 percent attended.

Primary outcomes: effects of QTEL on student achievement and language development

The main finding is that offering QTEL did not cause a statistically significant overall effect on grade 7 or grade 8 achievement in English language arts or English language development. The magnitude of the effects on English language arts achievement was –0.01 of a standard deviation for grade 7 and 0.01 of a standard deviation for grade 8 students. The magnitude on English language development was 0.03 and 0.01 of a standard deviation for grade 7 and grade 8 students, respectively. The average student test scores did not differ significantly between

---

[46] Naturalistic conditions are those under which QTEL was implemented in the field. In contrast to strictly controlled laboratory research conditions, under naturalistic conditions researchers have virtually no control over the research setting after random assignment.

intervention and control schools. Stated differently, the grade 7 and grade 8 students in classrooms of teachers offered QTEL as a partial substitute for their regular professional development performed no different on the English Language Arts subtest of the California Standards Test (CST-ELA) or the California English Language Development Test (CELDT) than did students in control classrooms. Sensitivity analysis showed that this conclusion did not change when the analysis focused on students who attended schools with grade 6, grade 7, and grade 8, or when the analysis focused on students who spent three consecutive years in one middle school. The conclusion also did not change when missing data were analyzed using listwise deletion.

Secondary outcomes: effects of QTEL on teacher knowledge, attitudes, and practice

The secondary finding is that offering QTEL did not produce a statistically significant overall effect on teacher attitudes toward working with English language learner students, teacher knowledge, or teacher practice. The magnitude of the effects on these outcomes was 0.16 for teacher attitudes, 0.33 for teacher knowledge, and –0.06 for teacher practice. There was no statistically significant variability across schools by study group. Stated differently, the teachers offered QTEL as a partial substitute for their regular professional development performed no different than did teachers in control classrooms on the attitudinal survey items related to working with English language learner students, on an assessment of their knowledge of pedagogical practice, or in systematic observations of their classroom practice. The conclusion for teacher attitudes and for teacher practice did not change when missing data were analyzed using listwise deletion, but it did for teacher knowledge (see table E4 in appendix E).[47] The conclusion for teacher knowledge also changed when an alternative method for handling the consolidation of three study schools was used (see table E13 in appendix E) and when scores for teachers who received QTEL for more than one year were excluded from the analysis (see table E15 in appendix E).

Exploratory analysis outcomes

To probe further for the possibilities of QTEL's effects on students and teachers, several exploratory analyses were conducted. Subgroup analyses examined whether outcomes for particular subgroups of students and teachers were significantly affected by having QTEL offered at their schools.

The first exploratory analysis entailed examining outcomes by student subgroups. The 2010 CST-ELA scale scores for grade 7 and grade 8 students were analyzed for four subgroups: English only, initially fluent English proficient, redesignated fluent English proficient, and limited English proficient. No statistically significant impacts of QTEL on the English language arts achievement of any of these four groups in either grade 7 or grade 8 were found, with effect sizes ranging from –0.17 to 0.02.

---

[47] The change in results was because the sensitivity analysis did not correct for multiple comparisons. When multiple comparisons adjustments are included, the missing data sensitivity analysis for teacher knowledge matched the benchmark analysis in chapter 5.

The next analysis examined impacts by teacher education background and experience. To examine whether the impacts varied by teacher background, the researchers created subgroup breakdowns at the school level, designating schools with higher and lower proportions of teachers who had advanced degrees (master's degrees or doctorates) and schools with higher and lower proportions of experienced teachers (more or less than 12.8 years of teaching). The 2010 CST-ELA scale scores for the grade 7 and grade 8 students in the outcome sample were examined. QTEL was estimated to have increased the test scores of grade 8 students in schools where teachers were more likely to have an advanced degree. The estimated impact for these students' test scores was an increase of 0.17 of a standard deviation. This impact estimate was statistically significant ($p = .027$) and statistically significantly different from the estimated QTEL impact on the test scores of grade 8 students in schools with less highly educated teachers (effect size = 0.01, $p = .167$). No other impact estimates was statistically significant or differed significantly across the subgroups.

The next analysis focused on outcomes by classroom quality, as measured by the Sheltered Instruction Observation Protocol (SIOP). The sample was divided into two subgroups of quality. The estimated impacts of QTEL on the English language arts achievement of these two subgroups did not differ significantly, and no impact estimates were statistically significant or exceeded 0.1 standard deviation in magnitude.

The next exploratory analysis examined outcomes on different subscales of the 30-item SIOP, capturing lesson preparation (9 items), input and interactions (7 items), lesson activity (6 items), and lesson delivery (8 items). QTEL did not have any statistically significant impacts on any of the SIOP subscales. Effect sizes ranged from –0.14 to 0.14.

The research team created an additional measure of classroom quality more closely aligned with QTEL. This instrument, Program Aligned Classroom Observation (PACO), measures classroom quality along activity structure (11 items), discussion content (6 items), and student–student interaction (3 items). One set of analyses examined these dimensions as well as the overall score. Only the impact on student–student interaction was statistically significant. This intervention–control difference translated to an effect size of 0.46 ($p = .003$).

The last exploratory analysis focused on how measures of student achievement varied with different levels of teacher participation in professional development, as measured with teacher surveys. This exploratory analysis is nonexperimental and not protected from potential selection bias by the evaluation's experimental study design. Exposure to professional development in English language development standards and in instructional strategies for secondary English language learner students was significantly higher for teachers in intervention schools than for those in control schools. In addition, intervention teachers were more likely to report participating in at least seven days of such professional development in both these subject areas. However, more teacher exposure to professional development in English language development standards and in instructional strategies for secondary English language learner students was not associated with different levels of student achievement, regardless of whether such exposure was to professional development in general or involved more intensive participation (more than seven days) and regardless of the focus of the professional development activities reported by teachers.

Limitations

The internal validity of the findings is limited by the following:

1. Students and teachers left the schools between random assignment and when outcome data were collected. For example, as chapter 2 discussed, 29.5 percent of grade 6 students in 2007/08 were no longer in the grade 8 impact sample in 2009/10. For English language learner students, these grade 6–8 attrition rates were different between intervention schools (41.5 percent) and control schools (28.9 percent). If there are systematic differences between the expected outcomes of students and teachers who leave the intervention and control schools, such differences would bias the resulting impact estimates.

2. Three schools (two intervention and one control) were consolidated. To maintain the integrity of random assignment, one third of the teachers and students in the consolidated school were randomly selected and considered control group members in the outcome analysis, even though they were treated as intervention group members by QTEL after the consolidation took place. As a result, any estimated QTEL impact on this school would likely be attenuated (biased toward zero), causing a small bias in the overall impact estimates across the entire study sample.

3. Teachers in the control group continued to have access to their regular professional development activities as provided and prescribed by their school or district. The data on the control group are limited to survey responses from administrators and teacher self-reports. As a result, control teachers might have had undetected crossover exposure to QTEL or content similar.

4. Classroom observations were conducted using a convenience sample, meaning that the classrooms observed were not necessarily representative of their schools and grades and there might have been systematic differences between observed classrooms in intervention schools and those in control schools, which could bias the findings based on these observations.

5. Nonresponse on teacher surveys might have caused nonresponse bias in estimates based on data from those surveys. Teachers who responded in intervention schools might have systematically differed from teachers who responded in control schools.

The external validity of the findings is limited by the following:

1. The sample of school districts and schools where the study was carried out was not a random sample of schools in the U.S., California, or Southern California. There is no way to know whether the results in this report generalize beyond the study sample.

2. Participating schools and districts volunteered for a study in which schools were randomized to receive QTEL or to not receive it. This means that the results may not be representative of schools who are fully committed to QTEL, in which case they may have avoided randomization and contracted for QTEL (or similar services) directly.

3. Participation in QTEL was not universal. Most teachers eligible to participate did not receive all services as intended. Thus, the findings do not generalize to a setting where all participants receive all intended services.

4. The fidelity of implementation was limited. The delivery of some services was compromised by staffing and logistical issues, and tracking of service receipt by staff was limited. As a result, the findings do not generalize to a setting with complete implementation fidelity.

5. While the classroom observation instruments demonstrated acceptable levels of internal consistency[48] and interrater reliability, there is insufficient data available to establish external validity of either the SIOP or the PACO instrument. Prior research establishing the external validity of the SIOP was not available.

Four data quality issues limit the reliability of the findings:

1. Data on coaching and professional development attendance were incomplete and poorly documented. As a result, the description of implementation might be unreliable.

2. The main student outcome measures captured only part of the anticipated impact on student outcomes, because standardized test scores do not measure the full range of skills and competencies required for success in grade-level content area courses.

3. Baseline data on students and teachers were limited. Students entered grade 6 after random assignment. And no elementary school data were available to establish their baseline equivalence, both in grade 6 and after attrition in grades 7 and 8. Teacher baseline data were limited to classroom observations, which were not linked to individual teachers and could be used only to establish baseline equivalence at the school level.

4. Because teacher rosters were either unavailable or unreliable, individual students could not be linked to individual teachers. This prevented analyses of the direct relationship between a teacher's receiving QTEL services and that teacher's students' outcomes.

Implications and future research

Educating English language learner students continues to challenge all educators, and there is little agreement on the most effective approach, especially because all students, including English language learner students, learn differently. This study shows the difficulty of studying a nonscripted intervention tailored to teachers' and students' needs, built on a solid foundation of learning theories. It is a challenge to capture the quality and potentially positive effects of an intervention like QTEL using a design that necessitates altering the intervention.

Anecdotally (according to surveys), teachers who participated in QTEL meetings and coaching found this support helpful; however, this does not demonstrate whether QTEL was significantly

---

[48] We considered a Cronbach's alpha of 0.70 as the minimum acceptable for internal consistency for the classroom observation measures. The 30-item SIOP demonstrated a Cronbach's alpha of 0.94. The 22-item PACO demonstrated a Cronbach's alpha of 0.92.

more helpful than other approaches. The exploratory analysis suggests that QTEL might have differential effects on the students of teachers with different levels of education. It also suggests there might be a positive relationship between QTEL and the amount of student–student interaction within classrooms, as measured by an observation instrument aligned with the intervention. These findings call for more research on how QTEL interacts with teachers' prior learning and how it affects teacher practice. Such research will require improved measures of teacher knowledge and teacher practice.

To understand QTEL more fully, in-depth studies must capture teacher and student responses to the intervention on an ongoing basis and generate analyses of it over time as it is intended to be implemented: as a schoolwide intervention in which coaches work intensively with the same teachers over multiple years. Longitudinal studies might provide information on these potential effects.

Appendix A. Statistical power analysis

This study design requires that any impact estimates have sufficient statistical precision so that impacts that are practically meaningful also are statistically significant. Middle school education interventions usually have impacts between 0.30 to 0.60 standard deviation (Hill et al. 2008), so the current study was designed at the outset to detect impacts at least of that magnitude on student outcomes in key student subgroups. On the advice of the Institute of Education Sciences and its technical advisors, a more conservative full sample minimum detectable effect size of 0.2 was adopted during the recruitment phase of the study. (The minimum detectable effect size, or MDES, is the smallest true effect that has at least an 80 percent probability of being detected).

To calculate student-level MDES, empirical data reported by Hill et al. (2008) were used to determine appropriate values for the intraclass correlation and explanatory power (R^2) of the study's impact regressions. An intraclass correlation of 0.15 and an R^2 of 0.65 were used as the basis for the power analyses, which was conducted using Optimal Design software. An alpha of 0.05 was used and 80 percent power. The relationship between the number of schools and the minimum detectable effect size under those assumptions for an anticipated average sample of 350 students per grade per school is shown in figure A1. To achieve the desired MDES of 0.20, a target sample size of 50 schools was needed. As discussed in the report, we recruited 54 schools so that the study would not become statistically underpowered if one or more schools would drop out. After two schools did drop out, our final school-level sample size was 52, which would have given us slightly more statistical power than we needed for an MDES of 0.2. .

Figure A1. Relationship between number of schools and minimum detectable effect size, full sample of 350 students per grade per school
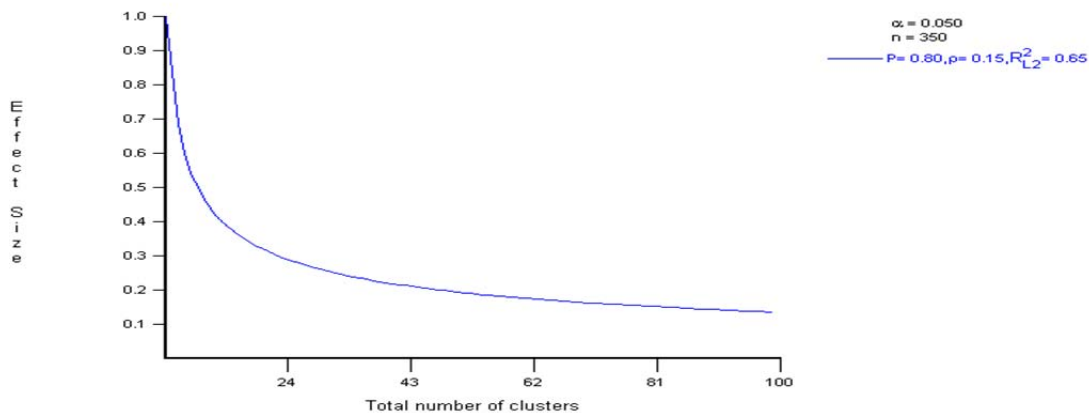
Source: Authors' calculation.

For English language learner students, who were estimated to comprise approximately 25 percent of the student population, figure A2 shows the corresponding statistical power figure. The a priori minimum detectable effect size for this student subsample (maintaining all other assumptions as previously described) was estimated to be 0.21.

Figure A2. Relationship between number of schools and minimum detectable effect size, full sample of 88 English language learner students per grade per school
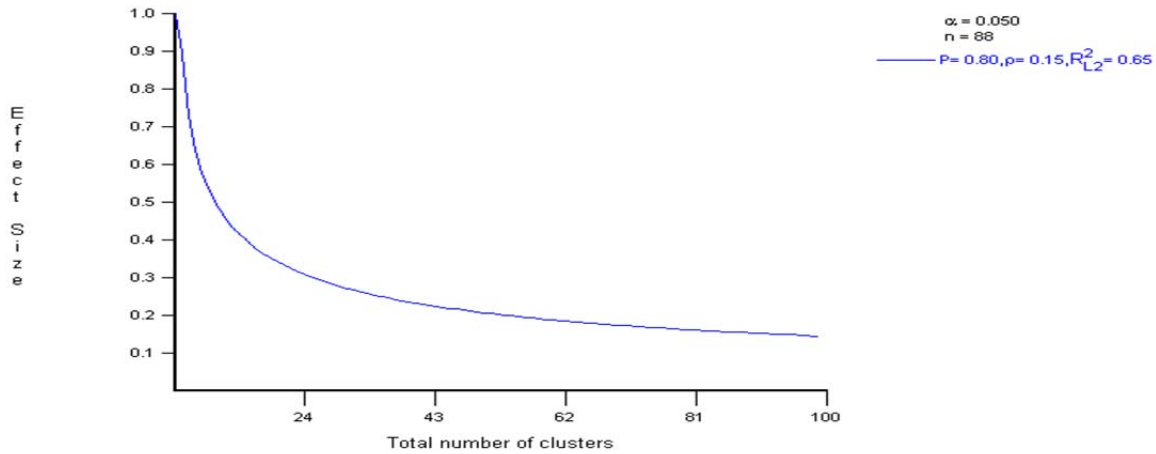
Source: Authors' calculation.

For calculating teacher-level minimum detectable effect sizes, no reliable data were available covering the explanatory power of teacher-level covariates in impact regressions. Nor were they available for teacher- and classroom-level intraclass correlations. Thus, in estimating teacher-level minimum detectable effect sizes, intraclass correlation coefficients of 0.05 and 0.20 were used to bracket a likely empirical value, as was an R^2 of 0, which is as conservative as it can be. This resulted in teacher-level minimum detectable effect sizes that ranged from 0.37 to 0.47 for a sample of 50 percent of all available teachers, using these assumptions. Outcomes measured via classroom observations had expected minimum detectable effect sizes ranging from 0.49 to 0.55, assuming an average of three successful classroom observations per school.

Because none of the primary impact estimates presented in this report approached either our minimum detectable effect sizes or statistical significance, we did not estimate or report actual realized statistical power for this study.

Appendix B. Random assignment

SAS code for first round of random assignment

Below is the SAS code for the first round of QTEL random assignment. This round of randomization included 48 schools. Before the schools were randomized, they were stratified by school district to produce a balanced sample of schools in each research group.

```
optionsnocenternofmterrls=256 ps=1000 ;
title 'random assignment for QTEL March 30 2007' ;
libnameqtel 'c:\data' ;

proc sort data=qtel.qtelschools ; by district ; run ;

data qtel.qtelschools2 ;
setqtel.qtelschools ;

 /* rename two Middle Schools with the same name to avoid confusion */

if district = 'District 01' and school = 'School 04' then
        school = 'School 04 District 01' ;

if district = 'District 07' and school = 'School 45' then
        school = 'School 45 District 07' ;

label school = 'name of school' ;
label district = 'name of district' ;

random1 = ranuni(856921) ; /* These seeds were randomly generated in Excel */
random2 = ranuni(29088) ;

label random1 = 'first random number' ;
label random2 = 'second random number' ;

if district ^= ' ' ; /* drop excessive rows in spreadsheet */

run ;

/* create a schoolcount variable */
proc sort ; by district ; run ;
proc means noprint ;
var random1 ;
by district ;
output out=districtcounts N=schoolcount mean=districtrandom ;
run ;


data qtel.qtelschools3 ;
merge qtel.qtelschools2 districtcounts ;
by district ;
run ;

proc print ;
title2 'print of schools including district Ns and mean random number by district' ;
run ;

/* now determine which odd-numbered districts get one extra program school */

dataodddistricts ;
setdistrictcounts ;
mergdum = 1 ; /*dummy for table merges */
where mod(schoolcount,2) ^= 0 ; /* drop districts with even numbers of schools for now */
run ;

proc print ;
```

```
title2 'print of districts with odd numbers of schools' ;
run ;
/* now create a rank variable to rank districts based on districtrandom */
proc rank out=districtrank ;
vardistrictrandom ;
ranksdistrictrandomr ;
run ;

datadistrictrank ;
setdistrictrank ;
mergdum = 1 ; /* dummy for table merges */
labeldistrictrandomr = 'rank of odd districts by random1' ;
run ;

/* now count the number of districts */
proc means noprint ;
vardistrictrandom ;
output out=districtn N=districtn ;
run ;
/* add the mergdum variable and label districtn */
datadistrictn ;
setdistrictn ;
mergdum = 1 ;
labeldistrictn = 'number of districts with odd number of schools' ;
run ;

/* merge everything together and create preferred dummy*/
data odddistricts2 ;
mergeodddistrictsdistrictrank ;
by district ;
run ;
data odddistricts3 ;
merge odddistricts2 districtn ;
bymergdum ;
preferred = (districtrandomr le (districtn/2)) ; /* in top half */
 /* NOTE: This ONLY works when there is an even number of districts with odd numbers of schools.
In
        other words, the total number of schools must be even across all districts! */

label preferred = 'School is in district which has odd number of schools and has more program
than control schools' ;
run ;

/* merge preferred dummy onto the full sample file */
data qtel.qtelschools4 ;
merge qtel.qtelschools3 odddistricts3 (keep = district preferred districtrandom) ;
by district ;
if preferred =. then preferred = 0 ; /* make preferred 0 for schools in districts with even #s of
schools */
drop _TYPE_ _FREQ_ ; /* unnecessary system vars */
run ;

/* now rank the schools by random2 (within their districts) */
proc rank out=schoolrank ;
var random2 ;
ranksschoolrank ;
by district ;
run ;

dataqtel.RAoutcome ;
setschoolrank ;

labelschoolrank = 'rank of school within district based on random2' ;
 /* now pick the top half of schools for districts with even numbers of schools and the top half
+ 1 for districts
with odd numbers of schools */
 QTEL = ((schoolrank - preferred) le (schoolcount/2)) ;
label QTEL = 'Assigned to the QTEL program (1=In QTEL, 0=Control)' ;
run ;

proc freq ; tables QTEL ;
```

```
title2 'This number must be 24' ;
run ;
proc freq ;
tables QTEL * district ;
title2 'Must be 50/50 or within 1 school' ;
run ;
proc print ;
var district school QTEL ;
title2 'Random Assignment outcome' ;
run ;
```

Results of random assignment

Table B1 presents the results of the random assignment, including the number of strata developed.

Table B1. Random assignment results

District	Number of schools randomly assigned	Number of strata developed for random assignment	Number of schools	
			Intervention	**Control**
District 1	4	1	2	2
District 2	2	1	1	1
District 3	4	1	2	2
District 4	3	1	2	1
District 5	5	1	2	3
District 6	19	3	9	10
District 7	11	1	6	5
District 8	4	1	2	2
District 9	2	1	1	1
Total	54	11	27	27

Note: After random assignment, District 9 opted out of the study, resulting in the loss of its two schools.
Source: Authors' analysis.

Appendix C. Data collection instruments

QTEL 2010 Teacher Instructional Knowledge Test

CONSENT TO PARTICIPATE IN RESEARCH (QUESTIONAIRES)
On The Quality Teaching for English Learners (QTEL) Program
For Teachers

Your school is participating in a research project conducted by Berkeley Policy Associates (BPA), in partnership with WestEd, with funding from the U.S. Department of Education. This study is directed by Dr. Johannes Bos and Dr. Raquel Sanchez of BPA. The purpose is to learn whether the training provided by the Quality Teaching for English Learners (QTEL) program enhances the quality of instruction and the achievement of English Learners. This form provides you with information about the study below.

Please read the information below and ask questions about anything you do not understand before deciding whether or not to participate.

QTEL was developed by WestEd's Teacher Professional Development Program in 1999. Since that time it has grown to be a highly respected provider of teacher professional development in the country. QTEL includes group training sessions and on-site coaching for each teacher in the treatment group. The QTEL curriculum covers topics in sociocultural notions of teaching and learning; development of teacher expertise; scaffolding of teaching and learning for adolescent English Learners; and constructing meaning through sustained and deep interactions.

If you agree to participate in this study, you will be asked to complete an online questionnaire about your professional preparation and teaching experience. Your responses to this questionnaire will be anonymous. At a later date, you will be asked to complete a second questionnaire about your knowledge about working with English learners.

In thanks for participation in the study, you will receive a gift card upon completion of each questionnaire.

There are no significant risks to the school or the individual teacher as a result of participating in this study. Your participation in the research project may benefit your school, staff, and students by helping to improve the quality of instruction. The results of the research may influence policy and promote public investments in quality instruction for English Learners and teacher professional development.

Any information that is obtained in connection with this study and that can be identified with you, your school, or the students in your classes will remain confidential and will be disclosed only with your permission or as required by law. Observation or suspicion of abuse or neglect is reportable by law to the appropriate authorities. All individual and school-level information will be recorded with an identification number, and names will be kept in a separate location. Results will be analyzed and reported only in averages for groups of students and groups of schools; no individuals, individual schools, or districts will be identified by name.

The privacy of the information collected about you and your school will be protected by keeping all paper data in locked files at the offices of West Ed or Berkeley Policy Associates. All computer records will be kept in password-protected, secure storage under the direct control of the researchers.

Your participation in the program and in the research study is completely voluntary. If you volunteer to be in this study, you may withdraw at any time without consequences of any kind. If you have any questions or concerns about the research, please feel free to contact Raquel Sanchez, Project Director, Berkeley Policy Associates, 440 Grand Avenue, Suite 500, Oakland, CA, 94610, (510) 465-7884. In addition, if you have questions about your rights as a research participant, or if you have complaints, concerns, or questions about the research, please contact Fannie Tseng, Human Subjects Protection Administrator, Berkeley Policy Associates, (510) 465-7884 or fannie@bpacal.com; or Independent Review Consulting at 800-472-3241 or subject@irb-irc.com.

After reading the consent form above, please indicate your choice below:

❍ I consent
❍ I do not consent

I. Background Information

1. What school district do you currently teach in?

_____

2. What school do you currently teach in?

_____

3. What grade(s) do you currently teach? (Check all that apply)

❑ 6th
❑ 7th
❑ 8th

4. Which courses do you currently teach? (Check all that apply)

❑ English Language Arts (ELA)
❑ English as a Second Language (ESL)
❑ Other (please specify):

_____

Instructions: Please choose the best response from those provided for each item. Do not spend too much time on any one item.

II. Scaffolding

5. Schema building refers to _____.

❍ Constructing concept maps for textual themes
❍ Using thematic units to develop language skills
❍ Monitoring your own thinking process
❍ Developing complex connections between and across ideas
❍ Using graphic organizers to develop meanings for vocabulary items

6. Text re-presentation is _____.

❍ Simplifying language for English Language Learners
❍ Inviting students to process language by using a different genre
❍ Comparing language structures for the purpose of clarification
❍ Paraphrasing ideas from a text
❍ Showing students pictures/video/music to better explain a concept/subject

7. The purpose of scaffolding is to _____.

❍ Enable a student to complete a task that he/she would otherwise not be able to do alone
❍ Teach language to English Language Learners in stages
❍ Guide an activity that is carried out as the result of processing language
❍ Activate a students' prior knowledge about the content of a text
❍ Help a student to whatever extent necessary so that the task is completed

8. What kind of scaffolding does the Think-Pair-Share provide?

❍ Schema building and story telling
❍ Bridging and schema building
❍ Role-play and reflection
❍ Modeling and metacognitive development
❍ Dialogue writing and bridging

9. Bridging refers to _____.

❍ Starting with familiar concepts and connecting them to new ideas
❍ Inviting students to process language by using a different genre
❍ Language processes that are based on repeated written use and habit formation
❍ Monitoring your own thinking process
❍ The ability to internalize new vocabulary

10. Which of the following is the most effective type of instructional scaffolding for ELLs?

❍ Repetition
❍ Reconstruction
❍ Comprehensible input
❍ Bridging
❍ Verbalization

11. Which of the following is an example of modeling?

❍ Showing video clips to emphasize a particular concept
❍ Having students orally repeat a set of vocabulary words until pronunciation approximates that of a native speaker
❍ Having students create a poster that summarizes a story or passage
❍ Demonstrating a science experiment
❍ Using prior knowledge and experience to connect ideas

12. The instructional use of 'anticipatory guides' is effective because they _____.

❍ Identify information that students will need to know
❍ Provide a model of what a student should learn
❍ Elaborate and extend the text
❍ Allow comparisons with what a student knows in his first language
❍ Activate students' prior knowledge

13. An example of a technique that encourages metacognitive development is _____.

❍ Creating a Mind Mirror
❍ Retelling a story to someone
❍ A teacher using explicit language
❍ A teacher thinking aloud about how to solve a problem
❍ A student taking notes from a story while reading

14. Extended anticipatory guides promote metacognitive development by _____.

❍ Having students use a graphic organizer
❍ Requiring that students have an opinion
❍ Asking students to support their thinking with evidence from a text
❍ Asking students to read difficult texts
❍ Allowing students to take some time to jot down their ideas before speaking

III. Literacy Development

15. Academic language refers to _____.

❍ Language that is used by a teacher or instructor
❍ Abstract concepts
❍ Written language like that used in college texts
❍ Language used in formal contexts for academic subjects
❍ Non-comprehensible input

16. Which of the following should occur first in the lesson planning process?

❍ Creating different activities for students at different levels of proficiency
❍ Grouping students according to proficiency level
❍ Inviting students to read the text one time on their own first
❍ Identifying how you will prepare the learners prior to the reading
❍ Identifying which vocabulary might prove difficult and providing definitions

17. The main goal for inviting students to interact with the text is…(Check one only)

❍ Requiring that more advanced students provide support to struggling peers
❍ Inviting students to practice their English reading skills
❍ Exposing students to new literature
❍ Having students analyze text for meaning
❍ Enabling students to appropriate new vocabulary

18. The main factor that makes academic text difficult for ELLs in secondary classrooms is _____.

❍ Disciplinary discourse
❍ Word length
❍ Sentence length
❍Unfamiliar vocabulary
❍ Elaboration

19. Which of the following is a key feature of quality teaching for English language learners?

❍ Exploration of ideas at increasing levels of depth
❍ Teaching word meanings in small groups
❍ Direct instruction of vocabulary
❍ Lessons that are structured on Initiation-Response-Feedback
❍ Instructional language that is simplified to match students' linguistic proficiency

20. Which of the following is NOT a characteristic of effective classrooms for intermediate to advanced ELL students?

○ Structured opportunities for student participation
○ Multiple entries for participation
○ Simplified linguistic input
○ Access to rigorous curriculum
○ Appropriate use of teacher "wait time"

21. Which of the following strategies is most useful in helping ELLs interpret new vocabulary?

○ Having students repeat the new word 3 times in a row
○ Helping students learn how to use a dictionary effectively
○ Pre-teaching the unfamiliar words in a new text
○ Having students write a sentence for each new word
○ Helping students recognize central and peripheral information

22. A successful English language learner _____.

○ Surrounds herself with native speakers of English
○ Enjoys grammar and vocabulary study
○ Can tolerate ambiguity of meaning
○ Tries to use English more than his/her L1
○ Is able to learn new vocabulary words daily

23. The best way to teach vocabulary to ELL students is to _____.

○ Invite students to keep a personal glossary of vocabulary words
○ Teach simple vocabulary first, then move toward more difficult vocabulary
○ Embed new vocabulary words in meaningful chunks of text
○ Pre-teach vocabulary through lists and examples
○ Test vocabulary regularly (about once a week)

IV. Sociocultural Theories of Learning

24. The instructional strategy most clearly aligned with a sociocultural view of learning is _____.

○ Scaffolding
○ Direct instruction
○ Differentiated instruction
○ Cross-age tutoring
○ Sheltered instruction

25. In contrast to Vygotsky's Zone of Proximal Development, the *Expanded* Zone of Proximal Development includes interaction with _____ and _____.

❍ Native speakers and challenging texts
❍ More capable peers and challenging texts
❍ Equal peers and less capable peers
❍ Adults and native speakers
❍ Equal peers and native speakers

26. _____ is the main vehicle of thought.

❍ Reasoning
❍ Vocabulary
❍ Language
❍ Culture
❍ Intelligence

27. The primary process by which learning takes place is _____.

❍ Interaction
❍ Internalization
❍ Assimilation
❍ Repetition
❍ Memorization

28. In the apprenticeship model, the learner moves from _____ to _____.

❍Direct instruction to facilitation
❍ Peripheral participation to appropriation
❍ Imitation to invention
❍ Marginal appropriation to full appropriation
❍ Simple ideas to complex ideas

29. The Zone of Proximal Development is the _____.

❍ Level at which the material is too challenging for a student to comprehend
❍ Level at which students plateau and struggle to further their development
❍ Level at which a student is able to work independently of help from their teacher or peers
❍ Difference between the level at which a learner can complete a task independently and the level at which she can complete it with support
❍ Frustration level

30. Metacognitive development refers to _____.

❍ Learning that occurs as a result of imitation, practice, and reinforcement
❍ When a teacher's lesson plan focuses on developing students' analytical skills
❍ The ways in which students examine and guide their thinking or cognitive processes
❍ The result of a teacher's use of schema building within their lesson plans
❍ When students' development occurs outside of structured learning activities

31. The constructivist view of learning is one in which _____.

❍ Students are believed to be active participants in the learning process
❍ Students only learn from tightly constructed and compartmentalized lesson plans
❍ Learners build from an understanding of basic concepts toward more abstract reasoning
❍ Similarities between the first and second languages allow the learner to acquire second language structures with ease
❍ Learning occurs as a result of imitation, practice, and reinforcement

32. _____ creates a conceptual framework in which a meaningful context is maintained for several days or weeks.

❍ Culturally responsive teaching
❍ Thematic instruction
❍ Direct instruction
❍ Cooperative group work
❍ Information processing model

V. Second Language Acquisition

33. The best way to organize instruction for English Language Learners is _____.

❍ To use simple sentences and below grade - level texts
❍ To ensure that students reach a certain level of English proficiency before teaching grade level content
❍ To provide a specialized all-day program until ELLs reach oral fluency in English
❍ Use grade level curricula with appropriate support and scaffolding
❍ Teach in stages, beginning with simple vocabulary and then moving to more complex vocabulary

34. An indication that an ELL student is learning can be seen by _____.

❍ An increase in participation over time
❍ An ability to hold elaborate conversations with few errors
❍ An improvement in standardized test scores
❍ Fewer errors in written products
❍ Consistent improvement in pronunciation

35. Communicative Competence involves _____.

❍ The ability to communicate on a basic level
❍ Mastery of meaning within social and cultural contexts
❍ The level at which students are ready to transition into a mainstream classroom
❍ Language that is to be used in a classroom setting
❍ Code switching

36. Additive Bilingualism is _____.

❍ Developing the student's primary language while he or she acquires a second language
❍ The ability to engage in problem-solving, deduction, and complex memory tasks
❍ Having equal proficiency in two languages across a range of contexts
❍ The act of acquiring a third or fourth language
❍ Replacing the primary language with a new language

37. Cummins' Cognitive Academic Language Proficiency (CALP) is _____.

❍ The ability to engage in problem-solving, deduction, and complex memory tasks
❍ The level at which students are at a proficient enough level to be introduced into a mainstream classroom
❍ When the first language is partially or completely lost as a second language is acquired
❍ The ability to use language in all its forms as a tool for thinking and communicating effectively
❍ The language required to succeed in higher order, literacy-related tasks of the classroom

38. Metalinguistic knowledge involves the ability to _____.

❍ Comprehend multiple languages
❍ Find hidden meanings in the text
❍ Talk about language forms and functions
❍ Connect new texts with prior knowledge
❍ Translate texts accurately

39. Instructional conversation is an effective means for engaging ELLs in classroom discourse because _____.

❍ It provides different opportunities for modeling and feedback
❍ It enables language learners to memorize correct forms
❍ It allows for student and teacher to follow a prepared script
❍ Students are able to avoid working independently, which could result in more mistakes
❍ It prevents students from repeating each other's errors

40. Cummins' Common Underlying Proficiency is _____.

○ Made up of basic interpersonal communication skills (BICS)
○ When teachers encourage all students to memorize key vocabulary prior to the reading of text
○ Developing networks among clusters of meaning that are interconnected
○ The foundational linguistic knowledge and skills on which a learner can draw to learn a new language
○ A single underlying abstract structure of all languages that children must acquire in early childhood

41. Krashen's Comprehensible Input is _____.

○ The recommendation that teachers use language just beyond students' current ability level
○ A metaphor for the interaction of emotional factors with other factors that serve to make a learner more or less open to second-language input
○ Translation into the primary language to ensure that the students will grasp key concepts
○ The order in which certain features of a language are acquired
○ Simplification of language input to the students' current ability level

42. Transfer refers to _____.

○ The replacement of the primary language with an acquired language
○ The omission of elements of a sentence
○ Overgeneralization of learned grammar rules
○ The continued use of a student's first language after they have gained competency in the newly learned language
○ A learner's use of patterns of the first language in second language sentences

43. Developmental errors are _____.

○ The omission of elements of a sentence
○ An error that reflects the learner's gradual discovery of the second language system
○ A learner's use of patterns of the first language in second language sentences
○ Overgeneralization of learned grammar rules
○ Errors learned or picked up from interaction with peers with limited fluency

44. Overgeneralization errors are the result of _____.

○ A learner's use of patterns of the first language in second language sentences
○ The omission of elements of a sentence
○ The learner's gradual discovery of the second language system
○ Trying to apply a linguistic rule in a context where it does not belong
○ Illogical reasoning

QTEL 2010 teacher survey, intervention group

CONSENT TO PARTICIPATE IN RESEARCH (QUESTIONAIRES)
On The Quality Teaching for English Learners (QTEL) Program
For Teachers

Your school is participating in a research project conducted by Berkeley Policy Associates (BPA), in partnership with WestEd, with funding from the U.S. Department of Education. This study is directed by Dr. Johannes Bos and Dr. Raquel Sanchez of BPA. The purpose is to learn whether the training provided by the Quality Teaching for English Learners (QTEL) program enhances the quality of instruction and the achievement of English Learners. This form provides you with information about the study below.

Please read the information below and ask questions about anything you do not understand before deciding whether or not to participate.

QTEL was developed by WestEd's Teacher Professional Development Program in 1999. Since that time it has grown to be a highly respected provider of teacher professional development in the country. QTEL includes group training sessions and on-site coaching for each teacher in the treatment group. The QTEL curriculum covers topics in sociocultural notions of teaching and learning; development of teacher expertise; scaffolding of teaching and learning for adolescent English Learners; and constructing meaning through sustained and deep interactions.

If you agree to participate in this study, you will be asked to complete an online questionnaire about your professional preparation and teaching experience. Your responses to this questionnaire will be anonymous. At a later date, you will be asked to complete a second questionnaire about your knowledge about working with English learners.

In thanks for participation in the study, you will receive a gift card upon completion of each questionnaire.

There are no significant risks to the school or the individual teacher as a result of participating in this study. Your participation in the research project may benefit your school, staff, and students by helping to improve the quality of instruction. The results of the research may influence policy and promote public investments in quality instruction for English Learners and teacher professional development.

Any information that is obtained in connection with this study and that can be identified with you, your school, or the students in your classes will remain confidential and will be disclosed only with your permission or as required by law. Observation or suspicion of abuse or neglect is reportable by law to the appropriate authorities. All individual and school-level information will be recorded with an identification number, and names will be kept in a separate location. Results will be analyzed and reported only in averages for groups of students and groups of schools; no individuals, individual schools, or districts will be identified by name.

The privacy of the information collected about you and your school will be protected by keeping all paper data in locked files at the offices of West Ed or Berkeley Policy Associates. All

computer records will be kept in password-protected, secure storage under the direct control of the researchers.

Your participation in the program and in the research study is completely voluntary. If you volunteer to be in this study, you may withdraw at any time without consequences of any kind. If you have any questions or concerns about the research, please feel free to contact Raquel Sanchez, Project Director, Berkeley Policy Associates, 440 Grand Avenue, Suite 500, Oakland, CA, 94610, (510) 465-7884. In addition, if you have questions about your rights as a research participant, or if you have complaints, concerns, or questions about the research, please contact Fannie Tseng, Human Subjects Protection Administrator, Berkeley Policy Associates, (510) 465-7884 or fannie@bpacal.com; or Independent Review Consulting at 800-472-3241 or subject@irb-irc.com.

After reading the consent form above, please indicate your choice below:

❍ I consent
❍ I do not consent

I. BACKGROUND INFORMATION

1. What school district do you currently teach in?

_____

2. What school do you currently teach in?

_____

II. CERTIFICATION & EXPERIENCE

3. What grade(s) do you currently teach? (Check all that apply)

❑ 6th
❑ 7th
❑ 8th

4. Which courses do you currently teach? (Check all that apply)

❑ English Language Arts (ELA)
❑ English as a Second Language (ESL)
❑ Other (please specify):

_____

5. Which of the following teaching credentials or professional licenses do you hold? (Check all that apply)

❑ California Teaching Credential (Preliminary or Clear)
❑ Special Endorsement, Special Education
❑ Special Endorsement, CLAD
❑ Special Endorsement, B-CLAD
❑ Other Special Endorsement
❑ Administrator
❑ Substitute/Provisional License
❑ National Board Certification
❑ Teaching License from state other than California
❑ Other (please specify):

_____

6. What is your highest level of education?

❍ Bachelor's degree
❍ Master's degree
❍ Doctorate or professional degree (Ph.D., Ed.D., M.D., J.D., D.D.S., etc.)

7. Counting this year as one year, how many years have you been teaching at any school in any district in the U.S.? Include all teaching experience except student teaching and substitute teaching.

❍ 1-3 years
❍ 4-6 years
❍ 7-9 years
❍ More than 10 years

8. Counting this year as one year, how many years have you been teaching in the county in which you currently teach? Include all teaching experience except student teaching.

❍ 1-3 years
❍ 4-6 years
❍ 7-9 years
❍ More than 10 years

III. INSTRUCTIONAL PRACTICES

9. Please indicate how often you use the following instructional techniques. (Choose one per row)

	Never / almost never	1–2 times per week	1–2 times per day	More than 2 times a day	Don't know / not sure
Explicitly teach academic language particular to your content area	○	○	○	○	○
Use multiple techniques to make concepts and tasks clear (e.g., visuals, manipulatives, realia, modeling)	○	○	○	○	○
Make lower level materials available for students with lower English proficiency	○	○	○	○	○
Provide opportunities for all students to use higher-order thinking skills (e.g., problem solving, predicting, organizing, evaluating, self-monitoring)	○	○	○	○	○
Simplify input to make it more comprehensible to English learners	○	○	○	○	○
Adjust expectations for students' whose limited English proficiency prevents them from meeting state or district standards	○	○	○	○	○
Create heterogeneous groups with regard to English language proficiency levels	○	○	○	○	○
Perform regular comprehension checks (e.g., requests for clarification, repetition, on-going assessment of students' performance)	○	○	○	○	○
Create homogeneous groups with regard to English language proficiency level	○	○	○	○	○

10. How frequently do you ask students in your class(es)—including ELLs—to engage in the following activities? (Check one per row)

	Never / almost never	1–2 times per week	1–2 times per day	More than 2 times a day	Don't know / not sure
Develop oral or written summaries of a reading	○	○	○	○	○
Evaluate their own work	○	○	○	○	○
Complete workbook or textbook exercises in class	○	○	○	○	○
Evaluate a piece of work completed by another student	○	○	○	○	○
Memorize vocabulary, facts, rules or procedures	○	○	○	○	○
Engage in discussions about a reading	○	○	○	○	○
Listen to a lecture and take notes	○	○	○	○	○
Work in small groups of two or more students	○	○	○	○	○
Use data and text references to support their ideas	○	○	○	○	○

11. Please indicate the extent to which you agree or disagree with the following statements about student learning. (Check one per row)

	Strongly disagree	Disagree	Agree	Strongly agree	Don't know / not sure
Some students I teach are not capable of learning the content I am supposed to teach them	○	○	○	○	○
It is important for ELL students to develop and maintain their primary language	○	○	○	○	○
Teachers should incorporate the cultures of their students into instructional activities	○	○	○	○	○
The use of primary language in the classroom slows down English language learning	○	○	○	○	○
ELL students require a disproportionate amount of a teacher's time compared to non-ELL students	○	○	○	○	○
I feel I have the professional preparation necessary to meet the needs of ELL students	○	○	○	○	○
The presence of ELL students in mainstream classrooms has a negative impact on the achievement of other students	○	○	○	○	○
The appropriate way to deal with an ELL's lack of comprehension is to use simplified language	○	○	○	○	○
Teachers should modify their instruction to meet the needs of ELL students	○	○	○	○	○

IV. TEACHING CONTEXT

12. To what extent is each of the following a challenge for your school? (Check one per row):

	Serious challenge	Moderate challenge	Minor Challenge	Not at all a challenge	Don't know / not sure
Shortage of qualified teachers	○	○	○	○	○
Shortage of bilingual teachers	○	○	○	○	○
Time for teachers to collaborate	○	○	○	○	○
A high proportion of ELLs	○	○	○	○	○
Student behavior/discipline	○	○	○	○	○
A lack of community or parent support	○	○	○	○	○
A lack of student motivation	○	○	○	○	○
A lack of appropriate materials for ELLs	○	○	○	○	○
Collegiality among faculty	○	○	○	○	○

13. Please indicate to what extent in the last two years you have participated in professional development focused on the following topics. (Check one per row):

	Not at all	1–3 days	4–6 days	7 days or more
English language development standards	○	○	○	○
Content area standards	○	○	○	○
Support for a published curriculum in language arts or other content areas	○	○	○	○
Instructional strategies for secondary ELLs	○	○	○	○
Differentiation of instruction/ Differentiated Instruction	○	○	○	○
_____	○	○	○	○

14. To what extent do the following factors hamper your ability to teach to the standards? (Check one per row):

	A great deal	Somewhat	Very little	Not at all	Don't know / not sure
The number of ELLs in my class	○	○	○	○	○
The ability level of my students	○	○	○	○	○
Level of parent or community support	○	○	○	○	○
Limited knowledge for working with ELLs	○	○	○	○	○
The range of students' needs in my class	○	○	○	○	○
Level of support from principals/administrators	○	○	○	○	○
Limited knowledge of my content area	○	○	○	○	○
Level of support from other teachers	○	○	○	○	○
Inadequate materials and resources	○	○	○	○	○

V. PROFESSIONAL DEVELOPMENT SUPPORT

15. Please indicate the extent to which you found the QTEL summer professional development (Building the Base) useful. (Check one only)

○ Not useful at all
○ Not very useful
○ Somewhat useful
○ Very useful
○ I did not attend

16. Have you received any coaching from a QTEL coach?

○ Yes
○ No (Please Skip to Question 20)

17. My coaching sessions have consisted of (Check all that apply):

❏ One-on-one lesson planning meetings
❏ Observation of the lesson's implementation
❏ Discussion of the classroom observation with my coach
❏ Reviewing and discussing samples of student work
❏ Other (please specify):

_____

18. Please indicate the extent to which you agree or disagree with each of the following statements. (Check one per row):

	Strongly disagree	Disagree	Agree	Strongly agree	Don't know / not sure
My QTEL coach encouraged me to network with other teachers at my school	○	○	○	○	○
My QTEL coach encouraged me to collaborate with other teachers and work on lesson plans together	○	○	○	○	○
My QTEL coach emphasized the QTEL principles that were learned in the Building the Base workshops	○	○	○	○	○
My QTEL coach helped me to develop lesson plans	○	○	○	○	○
After each QTEL coaching session, I was asked to reflect on the session and evaluate how I felt about the coaching	○	○	○	○	○
My QTEL coach helped me to understand the purpose behind student tasks	○	○	○	○	○
My QTEL coach was available to me outside the normal session time, e.g., by email, by phone	○	○	○	○	○
Coaching sessions were well thought out and organized	○	○	○	○	○
My QTEL coach was knowledgeable with respect to the QTEL principles	○	○	○	○	○
My QTEL coach was knowledgeable with respect to implementing a variety of student tasks	○	○	○	○	○
My QTEL coach was knowledgeable with respect to implementing a variety of scaffolding techniques	○	○	○	○	○
I will be able to continue using QTEL principles after the coaching sessions have ended	○	○	○	○	○
The coaching was invaluable to my understanding of the QTEL method	○	○	○	○	○

19. Please indicate the extent to which you feel the following QTEL coaching activities were useful. (Check one per row):

	Not useful at all	Not very useful	Somewhat useful	Very useful	Don't know / not sure
Gathering information about my teaching context	○	○	○	○	○
Identifying learning goals for my target lessons	○	○	○	○	○
Identifying activities that will support the learning goals	○	○	○	○	○
Discussing grade appropriate academic content	○	○	○	○	○
Discussing language learning issues	○	○	○	○	○
The questions my coach asked to help me deepen the content of my lesson	○	○	○	○	○
Discussing how my lesson supports all students' achievement	○	○	○	○	○
Coach's notes on the content and structure of my target lessons during the pre-conference	○	○	○	○	○
Coach's record of the types of interactions that predominate in my classroom	○	○	○	○	○
Coach's record of the content of the interactions in my classroom	○	○	○	○	○
Coach's suggestions for tasks that engage students in productive activity	○	○	○	○	○
Coach's notes on the alignment of the observed lesson with my pre-conference plan	○	○	○	○	○
Discussing the observed lesson	○	○	○	○	○
The way my coach presented specific evidence from my classroom	○	○	○	○	○
The questions my coach asked to help me reflect on specific aspects of my teaching	○	○	○	○	○
Revising/identifying the learning goals for my future lessons	○	○	○	○	○

20. Please indicate the extent to which you found the lesson design meetings/collaborative planning sessions/study groups with other teachers at your school useful. (Check one only)

❍ Not useful at all
❍ Not very useful
❍ Somewhat useful
❍ Very useful
❍ I did not attend
❍ Collaborative planning sessions have not been scheduled at my school

21. Please indicate the extent to which participating in QTEL affected the following. (Check one per row):

	Not effective at all	Not very effective	Somewhat effective	Very effective	Don't know / not sure
Your knowledge regarding the needs of ELLs	❍	❍	❍	❍	❍
Your confidence in implementing QTEL tasks in the classroom	❍	❍	❍	❍	❍
Your awareness of the cultures of ELLs	❍	❍	❍	❍	❍
Your knowledge regarding the purposes of scaffolding	❍	❍	❍	❍	❍
Your use of scaffolding techniques to move students to higher levels of understanding	❍	❍	❍	❍	❍
Your ability to construct tasks that allow for multiple points of entry	❍	❍	❍	❍	❍
Your use of homogeneous and heterogeneous groupings in order to facilitate understanding	❍	❍	❍	❍	❍
Your use of academic, rather than simplified, language	❍	❍	❍	❍	❍

QTEL 2010 teacher survey, control group

QTEL 2010 Teacher Background and Implementation Survey

CONSENT TO PARTICIPATE IN RESEARCH (QUESTIONAIRES)
On The Quality Teaching for English Learners (QTEL) Program
For Teachers

Your school is participating in a research project conducted by Berkeley Policy Associates (BPA), in partnership with WestEd, with funding from the U.S. Department of Education. This study is directed by Dr. Johannes Bos and Dr. Raquel Sanchez of BPA. The purpose is to learn whether the training provided by the Quality Teaching for English Learners (QTEL) program enhances the quality of instruction and the achievement of English Learners. This form provides you with information about the study below.

Please read the information below and ask questions about anything you do not understand before deciding whether or not to participate.

QTEL was developed by WestEd's Teacher Professional Development Program in 1999. Since that time it has grown to be a highly respected provider of teacher professional development in the country. QTEL includes group training sessions and on-site coaching for each teacher in the treatment group. The QTEL curriculum covers topics in sociocultural notions of teaching and learning; development of teacher expertise; scaffolding of teaching and learning for adolescent English Learners; and constructing meaning through sustained and deep interactions.

If you agree to participate in this study, you will be asked to complete an online questionnaire about your professional preparation and teaching experience. Your responses to this questionnaire will be anonymous. At a later date, you will be asked to complete a second questionnaire about your knowledge about working with English learners.

In thanks for participation in the study, you will receive a gift card upon completion of each questionnaire.

There are no significant risks to the school or the individual teacher as a result of participating in this study. Your participation in the research project may benefit your school, staff, and students by helping to improve the quality of instruction. The results of the research may influence policy and promote public investments in quality instruction for English Learners and teacher professional development.

Any information that is obtained in connection with this study and that can be identified with you, your school, or the students in your classes will remain confidential and will be disclosed only with your permission or as required by law. Observation or suspicion of abuse or neglect is reportable by law to the appropriate authorities. All individual and school-level information will be recorded with an identification number, and names will be kept in a separate location. Results will be analyzed and reported only in averages for groups of students and groups of schools; no individuals, individual schools, or districts will be identified by name.

The privacy of the information collected about you and your school will be protected by keeping all paper data in locked files at the offices of West Ed or Berkeley Policy Associates. All computer records will be kept in password-protected, secure storage under the direct control of the researchers.

Your participation in the program and in the research study is completely voluntary. If you volunteer to be in this study, you may withdraw at any time without consequences of any kind. If you have any questions or concerns about the research, please feel free to contact Raquel Sanchez, Project Director, Berkeley Policy Associates, 440 Grand Avenue, Suite 500, Oakland, CA, 94610, (510) 465-7884. In addition, if you have questions about your rights as a research participant, or if you have complaints, concerns, or questions about the research, please contact Fannie Tseng, Human Subjects Protection Administrator, Berkeley Policy Associates, (510) 465-7884 or fannie@bpacal.com; or Independent Review Consulting at 800-472-3241 or subject@irb-irc.com.

After reading the consent form above, please indicate your choice below:

❍ I consent
❍ I do not consent

According to the Paperwork Reduction Act of 1995, no persons are required to respond to a data collection activity unless it displays a valid OMB control number. The valid OMB control number for this information collection is 1850-0842. The time required to complete this session should be about 60 minutes, including the time to review instructions. If you have any comments about the accuracy of the time estimate(s) or suggestions for improving this session, please write to: U.S. Department of Education, Washington, D.C. 20202-4651. If you have comments or concerns about your individual responses, write directly to: Rafael Valdivieso, U.S. Department of Education, 555 New Jersey Avenue, NW, Room 506E, Washington, D.C. 20208.

Per The Education Sciences Reform Act of 2002, Title I, Part E, Section 183, responses to this data collection will be used only for statistical purposes. The reports prepared for this study will summarize findings across the sample and will not associate responses with a specific district or individual. We will not provide information that identifies you or your district to anyone outside the study team, except as required by law.

I. BACKGROUND INFORMATION

1. What school district do you currently teach in?

_____

2. What school do you currently teach in?

_____

II. CERTIFICATION & EXPERIENCE

3. What grade(s) do you currently teach? (Check all that apply)

❑ 6th
❑ 7th
❑ 8th

4. Which courses do you currently teach? (Check all that apply)

❑ English Language Arts (ELA)
❑ English as a Second Language (ESL)
❑Other (please specify):

_____

5. Which of the following teaching credentials or professional licenses do you hold? (Check all that apply)

❑ California Teaching Credential (Preliminary or Clear)
❑ Special Endorsement, Special Education
❑ Special Endorsement, CLAD
❑ Special Endorsement, B-CLAD
❑ Other Special Endorsement
❑ Administrator
❑ Substitute/Provisional License
❑ National Board Certification
❑ Teaching License from state other than California
❑ Other (please specify):

_____

6. What is your highest level of education?

❍ Bachelor's degree
❍ Master's degree
❍ Doctorate or professional degree (Ph.D., Ed.D., M.D., J.D., D.D.S., etc.)

7. Counting this year as one year, how many years have you been teaching at any school in any district in the U.S.? Include all teaching experience except student teaching and substitute teaching.

❍ 1–3 years
❍ 4–6 years
❍ 7–9 years
❍ More than 10 years

8. Counting this year as one year, how many years have you been teaching in the county in which you currently teach? Include all teaching experience except student teaching.

❍ 1–3 years
❍ 4–6 years
❍ 7–9 years
❍ More than 10 years

III. INSTRUCTIONAL PRACTICES

9. Please indicate how often you use the following instructional techniques. (Choose one per row)

	Never / almost never	1–2 times per week	1–2 times per day	More than 2 times a day	Don't know / not sure
Explicitly teach academic language particular to your content area	○	○	○	○	○
Use multiple techniques to make concepts and tasks clear (e.g., visuals, manipulatives, realia, modeling)	○	○	○	○	○
Make lower level materials available for students with lower English proficiency	○	○	○	○	○
Provide opportunities for all students to use higher-order thinking skills (e.g., problem solving, predicting, organizing, evaluating, self-monitoring)	○	○	○	○	○
Simplify input to make it more comprehensible to English learners	○	○	○	○	○
Adjust expectations for students' whose limited English proficiency prevents them from meeting state or district standards	○	○	○	○	○
Create heterogeneous groups with regard to English language proficiency levels	○	○	○	○	○
Perform regular comprehension checks (e.g., requests for clarification, repetition, on-going assessment of students' performance)	○	○	○	○	○
Create homogeneous groups with regard to English language proficiency level	○	○	○	○	○

10. How frequently do you ask students in your class(es)—including ELLs—to engage in the following activities? (Check one per row)

	Never / almost never	1–2 times per week	1–2 times per day	More than 2 times a day	Don't know / not sure
Develop oral or written summaries of a reading	○	○	○	○	○
Evaluate their own work	○	○	○	○	○
Complete workbook or textbook exercises in class	○	○	○	○	○
Evaluate a piece of work completed by another student	○	○	○	○	○
Memorize vocabulary, facts, rules or procedures	○	○	○	○	○
Engage in discussions about a reading	○	○	○	○	○
Listen to a lecture and take notes	○	○	○	○	○
Work in small groups of two or more students	○	○	○	○	○
Use data and text references to support their ideas	○	○	○	○	○

11. Please indicate the extent to which you agree or disagree with the following statements about student learning. (Check one per row)

	Strongly disagree	Disagree	Agree	Strongly agree	Don't know / not sure
Some students I teach are not capable of learning the content I am supposed to teach them	○	○	○	○	○
It is important for ELL students to develop and maintain their primary language	○	○	○	○	○
Teachers should incorporate the cultures of their students into instructional activities	○	○	○	○	○
The use of primary language in the classroom slows down English language learning	○	○	○	○	○
ELL students require a disproportionate amount of a teacher's time compared to non-ELL students	○	○	○	○	○
I feel I have the professional preparation necessary to meet the needs of ELL students	○	○	○	○	○
The presence of ELL students in mainstream classrooms has a negative impact on the achievement of other students	○	○	○	○	○
The appropriate way to deal with an ELL's lack of comprehension is to use simplified language	○	○	○	○	○
Teachers should modify their instruction to meet the needs of ELL students	○	○	○	○	○

IV. TEACHING CONTEXT

12. To what extent is each of the following a challenge for your school? (Check one per row):

	Serious challenge	Moderate challenge	Minor Challenge	Not at all a challenge	Don't know / not sure
Shortage of qualified teachers	O	O	O	O	O
Shortage of bilingual teachers	O	O	O	O	O
Time for teachers to collaborate	O	O	O	O	O
A high proportion of ELLs	O	O	O	O	O
Student behavior/discipline	O	O	O	O	O
A lack of community or parent support	O	O	O	O	O
A lack of student motivation	O	O	O	O	O
A lack of appropriate materials for ELLs	O	O	O	O	O
Collegiality among faculty	O	O	O	O	O

13. Please indicate to what extent in the last two years you have participated in professional development focused on the following topics. (Check one per row):

	Not at all	1–3 days	4–6 days	7 days or more
English language development standards	O	O	O	O
Content area standards	O	O	O	O
Support for a published curriculum in language arts or other content areas	O	O	O	O
Instructional strategies for secondary ELLs	O	O	O	O
Differentiation of instruction/ Differentiated Instruction	O	O	O	O
_____	O	O	O	O

14. To what extent do the following factors hamper your ability to teach to the standards? (Check one per row):

	A great deal	Somewhat	Very little	Not at all	Don't know / not sure
The number of ELLs in my class	○	○	○	○	○
The ability level of my students	○	○	○	○	○
Level of parent or community support	○	○	○	○	○
Limited knowledge for working with ELLs	○	○	○	○	○
The range of students' needs in my class	○	○	○	○	○
Level of support from principals/administrators	○	○	○	○	○
Limited knowledge of my content area	○	○	○	○	○
Level of support from other teachers	○	○	○	○	○
Inadequate materials and resources	○	○	○	○	○

V. PROFESSIONAL DEVELOPMENT SUPPORT

15. What EL-specific professional development activities did you participate in <u>during the 2008–2009</u> school year? (Please check all that apply).

❑ SDAIE (Specially Designed Academic Instruction in English)
❑ CLAD (Cross-cultural, Language, and Academic Development)
❑ GLAD (Guided Language Acquisition Design)
❑ High Point Training
❑ Differentiated Instruction Training
❑ Other (please specify):

_____

16. Did any of these professional development activities take place during the summer?

○ Yes
○ No

17. Please indicate the extent to which you found the summer professional development provided by your school or district useful. (Please choose the N/A option if you did not participate in summer professional development sessions.)

○ Not useful at all
○ Not very useful
○ Somewhat useful
○ Very useful
○ I did not attend
○ N/A: My school or district does not offer any summer professional development sessions

18. Does your school or district's professional development program offer coaching sessions?

❍ Yes
❍ No (Please Skip to Question 22)

19. My coaching sessions have consisted of (Check all that apply):

❑ One-on-one lesson planning meetings
❑ Observation of the lesson's implementation
❑ Discussion of the classroom observation with my coach
❑ Reviewing and discussing samples of student work
❑Other (please specify):

_____

20. Please indicate the extent to which you agree or disagree with each of the following statements. (Check one per row):

	Strongly disagree	Disagree	Agree	Strongly agree	Don't know / not sure
Coaching sessions were well thought out and organized	❍	❍	❍	❍	❍
My coach emphasized the principles I learned in professional development workshops	❍	❍	❍	❍	❍
My coach helped me to develop lesson plans	❍	❍	❍	❍	❍
My coach was knowledgeable about the best practices around working with ELLs	❍	❍	❍	❍	❍
My coach was knowledgeable about to implementing a variety of scaffolding techniques	❍	❍	❍	❍	❍
My coach was knowledgeable about to implementing a variety of student tasks	❍	❍	❍	❍	❍
My coach encouraged me to network with other teachers at my school	❍	❍	❍	❍	❍
My coach encouraged me to collaborate with other teachers to work on lesson plans	❍	❍	❍	❍	❍
My coach was available to me outside the normal session time, e.g., by email, by phone	❍	❍	❍	❍	❍
After each coaching session, I was asked to reflect on the session and evaluate how I felt about the coaching	❍	❍	❍	❍	❍
I will continue using the principles I learned from my professional development program after the coaching sessions have ended	❍	❍	❍	❍	❍

21. Please indicate the extent to which you feel the following coaching activities were useful. (Check one per row):

	Not useful at all	Not very useful	Somewhat useful	Very useful	Don't know / not sure
Gathering information about my teaching context	○	○	○	○	○
Identifying learning goals for my target lessons	○	○	○	○	○
Identifying activities that will support the learning goals	○	○	○	○	○
Discussing grade appropriate academic content	○	○	○	○	○
Discussing language learning issues	○	○	○	○	○
The questions my coach asked to help me deepen the content of my lesson	○	○	○	○	○
Discussing how my lesson supports all students' achievement	○	○	○	○	○
Coach's notes on my lesson plans during the pre-conference	○	○	○	○	○
Coach's record of the types of interactions that predominate in my classroom	○	○	○	○	○
Coach's record of the content of the interactions in my classroom	○	○	○	○	○
Coach's suggestions for tasks that engage students in productive activity	○	○	○	○	○
Coach's notes on the alignment of the observed lesson with my pre-conference plan	○	○	○	○	○
Discussing the observed lesson	○	○	○	○	○
The way my coach presented specific evidence from my classroom	○	○	○	○	○
The questions my coach asked to help me reflect on specific aspects of my teaching	○	○	○	○	○
Revising/identifying the learning goals for my future lessons	○	○	○	○	○

22. Do you have opportunities to collaborate with other teachers (in common lesson planning sessions or study groups, for example) at your school?

❍ Yes
❍ No (Please Skip to Question 24)

23. Please indicate the extent to which you find collaborative planning sessions/study groups with other teachers at your school useful. (Check one only)

❍ Not useful at all
❍ Not very useful
❍ Somewhat useful
❍ Very useful
❍ I did not attend
❍ There are no teacher collaboration activities at my school

24. Please indicate the extent to which participating in your school/district's professional development program(s) affected the following. (Check one per row):

	Not effective at all	Not very effective	Somewhat effective	Very effective	Don't know / not sure
Your knowledge regarding the needs of ELLs	❍	❍	❍	❍	❍
Your confidence in implementing tasks in the classroom that support ELLs in acquiring academic knowledge	❍	❍	❍	❍	❍
Your awareness of the cultures of ELLs	❍	❍	❍	❍	❍
Your knowledge regarding the purposes of scaffolding	❍	❍	❍	❍	❍
Your use of scaffolding techniques to move students to higher levels of understanding	❍	❍	❍	❍	❍
Your ability to construct tasks that allow for multiple points of entry	❍	❍	❍	❍	❍
Your use of homogeneous and heterogeneous groupings in order to facilitate understanding	❍	❍	❍	❍	❍
Your use of academic, rather than simplified, language	❍	❍	❍	❍	❍

District Administrator Survey 2009–2010

Quality Teaching for English Learners (QTEL) Survey

Dear District Administrator,

Berkeley Policy Associates (BPA) is a major subcontractor to Regional Education Laboratory, West (REL West), which has been contracted by the U.S. Department of Education to evaluate WestEd's Quality Teaching for English Learners (QTEL) professional development program.

As part of this study, we need to collect information from you about your district, the challenges facing ELA and ELD teachers with respect to working with English learners, and the professional development opportunities available to them. Additionally, we require your thoughts and perspectives on QTEL based on your experiences in the schools within your district where the QTEL program has been implemented.

Your input is a very important part of this study. Our report aims to be of value to education leaders across the country. We know how precious your time is and we greatly appreciate your participation.

Participation is voluntary. Your responses are protected from disclosure by federal statute (PL 107-279 Title I, Part C, Sec. 183). All responses that relate to or describe identifiable characteristics of individuals may be used only for statistical purposes and may not be disclosed, or used, in identifiable form for any other purpose, unless otherwise compelled by law.

Data from this survey will be combined to produce statistical reports. No individual data that links your name, school, or district with your responses will be included in the statistical reports.

A. Background Information

1) Name of the school district:

_____

2) Position/Title:

_____

3) Counting 2008-09 as one year, how many years have you been working at this school district in your current position?

❍ 1 year
❍ 2 years
❍ 3 - 5 years
❍ 6 - 9 years
❍ 10 or more years

Instructions: Please answer all questions and choose the best response for each item. Do not spend too much time on any one item.

You may stop at any point in the survey and return to complete the survey at a later time. You may do this by closing the survey window when you want to stop, and using the same survey link at a later time to pick up where you left off. As long as you are working from the same computer, you will be able to return to where you left off by clicking 'Resume' on the page displayed.

B. Challenges facing ELA/ELD teachers who work with English Learners (ELs)

4) What do you see as the main challenges facing ELA/ELD teachers who teach ELs in this school district? (For each row, please check the box that best describes your observation.)

	Serious Challenge	Moderate Challenge	Minor Challenge	Not a Challenge
High proportion of ELs in the classes they teach	❍	❍	❍	❍
Lack of a teaching credential	❍	❍	❍	❍
Lack of bilingual certification	❍	❍	❍	❍
Lack of opportunities for professional development specific to working with ELs	❍	❍	❍	❍
Have not participated in professional development specific to working with ELs	❍	❍	❍	❍
No time for teachers to collaborate with each other	❍	❍	❍	❍
Student behavior/discipline	❍	❍	❍	❍
Lack of student motivation	❍	❍	❍	❍
Lack of community or parent support	❍	❍	❍	❍
Lack of appropriate instructional materials for ELs	❍	❍	❍	❍
Other (Please specify in the 'Additional comments' box below)	❍	❍	❍	❍

5) What do you see as the main challenges to the implementation of EL-focused professional development activities for ELA/ELD teachers in your school district? (For each row, please check the box that best describes your observation.)

	Serious Challenge	Moderate Challenge	Minor Challenge	Not a Challenge
Finding time to implement professional development	❍	❍	❍	❍
Finding appropriate professional development activities for teaching ELs	❍	❍	❍	❍
Low interest from teachers/obtaining teacher buy-in	❍	❍	❍	❍
Availability of substitutes while teachers are in professional development	❍	❍	❍	❍
Compensation for substitutes while teachers are in professional development	❍	❍	❍	❍
Compensation for teachers for participating in professional development in the summer	❍	❍	❍	❍
District support for teacher professional development	❍	❍	❍	❍
Other (Please specify in the 'Additional comments' box below)	❍	❍	❍	❍

C. EL-focused teacher professional development

6) For the 2008-09 school year, were teachers in this school district required to participate in professional development activities mandated by the district?

❍ Yes
❍ No

7) Please list all district-mandated professional development activities in which teachers were required to participate.

_____
_____
_____

8) To the best of your knowledge, what other professional development opportunities did teachers in this school district participate in during the 2008–09 school year?

Please list all professional development opportunities, not mandated by the district, in which teachers participated.

_____
_____
_____

9) To the best of your knowledge, what EL-specific professional development activities did ELA/ELD teachers in your school district participate in during the 2008–2009 school year? (Please check all that apply)

❑ SDAIE (Specially Designed Academic Instruction in English)
❑ CLAD (Cross-cultural, Language, and Academic Development)
❑ GLAD (Guided Language Acquisition Design)
❑ High Point Training
❑ Differentiated Instruction Training
❑ Not applicable
❑ Other (please specify)

If you selected other, please specify
_____

10) What type of professional development programs do you plan on implementing for your ELA/ELD teachers during the 2009–10 school year? (Please check all that apply)

❑ SDAIE (Specially Designed Academic Instruction in English)
❑ CLAD (Cross-cultural, Language, and Academic Development)
❑ GLAD (Guided Language Acquisition Design)
❑ High Point Training
❑ Differentiated Instruction Training
❑ No Response
❑ Other (please specify)

If you selected other, please specify
_____

D. External contextual factors affecting the school district

11) Have any teaching positions been eliminated at this school district for the 2009–10 school year as result of the state budget cuts?

❍ Yes
❍ No

12) How many teaching positions have been eliminated?

 _____

13) Have any administrative or resource positions (EL or literacy coaches, for example) been eliminated at this school district for the 2009–2010 school year as result of the state budget cuts?

❍ Yes
❍ No

14) Please specify which positions were eliminated and how many.

_____
_____
_____

15) Please list and explain other external contextual factors that have significantly impacted the school district during the 2008–2009 school year (e.g. administration changes at the school or district level, local issues, etc.)

_____
_____
_____

E. Implementation of QTEL (Quality Teaching for English Learners) at this school district

16) Some of the ELA/ELD teachers in your school district are participating in the QTEL professional development program. How would you describe the participation of those teachers? Please select the response that best describes your observations.

❍ All or most ELA/ELD teachers actively and consistently participate in QTEL
❍ Some ELA/ELD teachers actively and consistently participate in QTEL
❍ Few ELA/ELD teachers participate actively or consistently in QTEL
❍ I do not know/ I am not sure

17) How would you describe your support for QTEL?

❍ I fully support QTEL and encourage all ELA/ELD teachers at my school to participate
❍ I fully support QTEL, but I take a neutral stance on the teachers' own decision about participating in the program
❍ I support QTEL, but with a little reservation
❍ I support QTEL, but with considerable reservation
❍ I do not support QTEL
❍ I do not know/ I am not sure

18) Do you have any concerns about the following issues regarding QTEL? (Please check the box for each row that best describes your level of concern.)

	Don't Know / Not Applicable	Not At All Concerned	Not Very Concerned	Somewhat Concerned	Very Concerned
Teachers' willingness/commitment to participate	❍	❍	❍	❍	❍
Helpfulness/usefulness of the program	❍	❍	❍	❍	❍
Resources needed to provide substitutes to allow the teachers to participate in the QTEL Building the Base summer training	❍	❍	❍	❍	❍
Resources needed to pay teachers while attending training outside of regular school hours	❍	❍	❍	❍	❍
Time commitment by teachers to participate in the QTEL training activities (e.g., summer institute, coaching, afterschool meetings)	❍	❍	❍	❍	❍
Time commitment by teachers to participate in the evaluation data collection activities (e.g., surveys, site visits, observations)	❍	❍	❍	❍	❍
Other positions (Please specify in the 'Additional comments' box below)	❍	❍	❍	❍	❍

19) Have you visited or attended any of the following QTEL activities? (Please check all that apply.)

❑ Summer Institute (Building the Base Training)
❑ Teacher Collaboration/Lesson Design Meeting
❑ Coaching Session between individual teacher and QTEL coach
❑ I have not attended any QTEL activities

20) Do you think there has been an improvement in teachers' overall classroom performance as a result of their participation in QTEL?

❍ Yes
❍ No
❍ I do not know/ I am not sure

21) If you have any comments about any aspects of the QTEL program or its evaluation activities, please provide them below.

_____
_____
_____

QTEL 2010 modified Sheltered Instruction Observation Protocol[49]

SIOP PLUS Classroom Observation Protocol (Modified SIOP)

Date:	Observer:
School/District:	Teacher:
Grade(s):	Class: Lesson Topic:
(Enter Military Time) Start time: _____: _____ End time: _____: _____	*(Circle one)* Lesson Type: Multi-day Single-day Don't know
❑ Solo Observation ❑ Paired Observation *Paired with_____*	Number of Students:

Directions: Using the rubrics on the following pages, circle the number that best reflects what you observe in the lesson. You may give a score from 0–4 (or NA on selected items). Cite under "Comments" specific examples of the behaviors/evidence observed.

---

[49] Permission to include this instrument in the report is pending approval from the publisher (Pearson Higher Education).

1. Preparation: Content Objectives				
4 Language arts **content objectives** for current lesson explicitly defined orally and in writing for students	**3**	**2** Language arts **content objectives** for current lesson implied for students	**1**	**0** No **language arts content objectives** for students

Notes:
Content objectives refer to WHAT students will learn in the current lesson. The content objectives should state simply what students should know and be able to do at the end of the day's lesson. The teacher should make these objectives explicit to students, orally and in writing.

Comments: *(Please write the content objectives for the observed lesson here)*

2. Preparation: Language Objectives				
4 **Language objectives** for current lesson explicitly defined orally and in writing for students	**3**	**2** **Language objectives** for current lesson implied for students	**1**	**0** No **language objectives** for students

Notes:
The language objectives should state HOW students will use language, during or at the end of the current lesson, to demonstrate what they have learned. The language objectives should be stated explicitly to students, orally and in writing.

Comments: *(Please write the content objectives for the observed lesson here)*

3. Preparation: Lesson Concepts				
4 **Lesson concepts** are aligned with grade level language arts standards	**3**	**2** **Lesson concepts** are somewhat aligned with grade level language arts standards	**1**	**0** **Lesson concepts** are not aligned with grade level language arts standards.

Comments: *(List the relevant concepts here)*

4. Preparation: Supplementary Material(s)

4	3	2	1	0
Teacher's use of **supplementary material(s)** directly supports student understanding of lesson objectives.		Teacher's use of **supplementary material(s)** indirectly supports student understanding of lesson objectives.		Teacher does not use **supplementary material(s)**.

Notes:
*Examples of supplementary materials include real life objects, pictures, visuals (maps, posters, overhead transparencies, etc.), multimedia (CDs, DVDs, videotapes etc), internet-based materials, related literature, Hi-lo readers, sentence strips, graphic organizers, and so on. While scoring this item, **focus is not on the variety of materials but on the way the teacher uses them to support students' understanding of lesson objectives**.*

Comments: (List the supplementary materials here)

5. Preparation: Adapted Text

4	3	2	1	0	NA
Teacher has provided **adapted text** to all levels of student proficiency.		Teacher has provided **adapted text** to more than one level of student proficiency.		Teacher has not provided **adapted text** to the students	

Notes:

*Adapted materials include leveled study guides, highlighted text, taped text, adapted or abridged text that accommodates the students' reading levels, marginal notes and using graphic organizers, the use of native language texts, etc. While scoring this item, **focus on the way teachers modifies or uses adapted text to make the language/ content of the lesson more accessible to students**. A score of **NA** should be given in cases where **no** adaptation is required to make the lesson materials accessible to **all** students.*

Comments: (List the adapted materials here)

6. Preparation: Authentic Language Production

4	3	2	1	0
Lesson activities provide ample opportunities for **authentic language production**.		Lesson activities provide little opportunity for **authentic language production**.		Lesson activities do not provide any opportunity for **authentic language production**.

Notes: The goal of authentic language production is one of constructing and expressing meaning. While scoring this item, focus on whether the lesson supports students creating meaning and expressing ideas through the oral and written language they produce. Formal aspects of language (grammar, pronunciation, etc) are important to the extent they help students communicate their ideas more effectively (For example, editing is important in writing lessons because it helps students articulate their ideas better).

Comments: *(List the relevant activities here)*

---

7. Building Background: Concepts Explicitly Linked to Student Experiences Outside of School

4	3	2	1	0	NA
Concepts explicitly linked to students׳ background experiences outside school/ schooling contexts		**Concepts loosely linked** to students׳ background experiences outside school/ schooling contexts		**Concepts not linked** to students׳ background experiences outside school/ schooling contexts	

Notes:

The lesson should engage the students' interests, and the teacher should establish a connection between the main concepts of the lesson/ activity and the students' lives outside the classroom, so that students can find these concepts meaningful outside the lesson itself.

Comments: *(List concepts here)*

8. Building Background: Connections Explicitly Made to Prior Learning				
4	**3**	**2**	**1**	**0**
Strong/explicit connections made between previous lessons/prior school learning and current lesson		**Weak/implicit connections** made between previous lessons/prior school learning and current lesson		**No connections made** between previous lessons/prior school learning and current lesson
Comments: *(List connections here)*				

9. Building Background: Key Vocabulary				
4	**3**	**2**	**1**	**0**
9. **Key vocabulary** is introduced in a systematic manner and the connections to the current lesson are clear.		**Key vocabulary** introduced but it is not done so in a systematic manner (or) the connections to the current lesson are unclear		**Key vocabulary** not introduced
Notes: Introducing key vocabulary in a systematic manner includes introducing the words, in context and in writing, to the students; providing definitions and examples; repeating them during the lesson to increase familiarity; and highlighting them for students to see. *For a #1 rating, the key vocabulary is introduced, but it is done so in an unsystematic manner* **or** *the connections to the current lesson are unclear.*				
Comments: *(List vocabulary here)*				

10. Comprehensible Input: Teacher's Speech

4	3	2	1	0
Teacher's speech generally appropriate for students' English proficiency levels		**Teacher's speech** sometimes inappropriate for students' English proficiency levels		**Teacher's speech** generally inappropriate for students' English proficiency levels

Notes:

If there are multiple student levels the teacher needs to accommodate all levels (adjusts speech)
Examples:

- ▪ ***Beginning students****: (e.g., slower rate, enunciation, and simple sentence structure for beginners). Appropriate volume and tone (mature not elementary, appropriate for age level, balancing between language level and developmental level)*
- ▪ ***Intermediate/Advanced****: Amplify –exposing students to more complex language but taking care to reinforce meaning. Providing definitions, "Amplify don't simplify". Natural rate of speech. Adjust speech if teachers notice students need clarification. Paraphrasing and repetition. Complex sentences and embedded definitions.*
- ▪ *Advanced/GATE**: Challenging vocabulary. Rich examples of language used by the teacher. Complex sentence structure. Natural rate of speech. Providing word definitions if necessary.**

Comments: *(Describe teacher's speech style here)*

11. Comprehensible Input: Explanation of Academic Tasks

4	3	2	1	0
The teacher provides clear and explicit **explanation** of academic tasks (defined and written). Students appear to understand tasks.		Teacher's **explanation** of the academic tasks is either implicit or unclear, but students appear to understand tasks/instruction.		Teacher provides no **explanation** of the academic tasks <u>and</u> students appear unclear on tasks/instruction.

Notes:
#1 If explanation is unclear <u>and</u> students are confused.
#3 Some explanation of academic tasks and students seem to know what to do.

#4: Instructions are given in at least two modalities (oral, written, visuals, kinesthetic.)

Comments: *(List instructions here)*

12. Comprehensible Input: A Variety of Teaching Techniques

4	3	2	1	0
A variety of teaching techniques used to make content concepts clear		Some teaching **techniques** used to make content concepts clear		No teaching **techniques used** to make content concepts clear

Notes:
*While scoring this item, focus on the teacher, and consider the techniques s/he is using to support instruction. What is she doing? Is s/he introducing new texts to the students by **previewing the material**? Is s/he explaining a task by **modeling** it or using **think-alouds**? Does she provide **repeated exposure** to the main concepts of the day's lesson? Does she **repeat** instructions and **review** main concepts? The phrases in bolded text are some examples of techniques used by the teacher.*

Comments: *(List techniques here)*

13. Strategies: Learning Strategies

4	3	2	1	0
Ample opportunities provided for students to use **learning strategies** (metacognitive, cognitive, social affective strategies)		Some opportunities provided for students to use **learning strategies**		No opportunity for students to use **learning strategies**

Notes:
Here the focus is on the students, and the strategies they initiate and use to support their learning of lesson concepts.
Learning strategies include
- ***cognitive/ task-based strategies*** *such as using background knowledge, making personal connections, predicting, making inferences, finding patterns, using graphic organizers, taking notes, summarizing, etc.*
- ***metacognitive strategies****, which help students to learn how to learn, such as organizing/ planning, managing the task, monitoring thinking, self-regulating learning, self-questioning, clarifying purpose of the activity/ lesson, recalling, and taking corrective action and*
- ***social affective strategies****, which include interacting with other students, asking peers questions or for clarifications, and working in pairs or cooperative learning groups.*

For a #4 rating, there should be ample opportunities provided for students to use all three learning strategies.

Comments: *(List learning strategies here)*

14. Strategies: Scaffolding Techniques

4	3	2	1	0
Teacher effectively uses **scaffolding techniques** to support students in understanding lesson content or activity.		Teacher uses **scaffolding techniques**, but they are not effective at supporting students in understanding the lesson content or activity.		Teacher does not use **scaffolding techniques,** or Uses **scaffolding techniques** that are inappropriate for supporting students in understanding lesson content and activity.

Notes:
Verbal scaffolding*: Teachers use prompting, questioning and elaboration of student responses to facilitate student progress toward higher levels of language proficiency and thinking. Examples: Paraphrasing a student's response, using 'think alouds' to model effective thinking strategies and monitor understanding, and reinforcing contextual meanings of key vocabulary.*
Procedural scaffolding: *Teacher organizes the lesson content and structure activities in ways that support language learners in completing complex academic tasks. Examples: Instructional routine that includes explicit teaching, modeling opportunities for practice and independent application,*
one-on-one teaching, coaching, modeling, small group instruction, partner practice, and heterogeneous grouping.
Instructional scaffolding*. Teachers use and create materials that guide students through the process of completing and academic task. Graphic organizers and anticipatory guides are a common tool used for providing instructional scaffolding.*

Comments: *(Describe scaffolding here)*

15. Strategies: Promoting High Order Thinking Skills				
4	**3**	**2**	**1**	**0**
Activity or tasks **promote high order thinking skills**.		Activity or tasks **promote mid level thinking skills**.		Activity or tasks **promote low level thinking skills**.
Notes: *While scoring this item, **focus on what the students are required to do** and the thinking skills they are using in the process of completing the lesson activity. Note: All three high level thinking skills (evaluation, synthesis, and analysis) do not need to be observed for a "high" thinking skill level scoring.*				
High	• *Evaluation:* *Determining value and providing a rationale for the response* • *Synthesis:* *Creating something "new" from the parts* • *Analysis:* *Breaking the concept into component parts*			
Mid	• *Application*: *Demonstrating knowledge by applying concepts to one's own life*			
Low	• *Comprehension*: *Basic understanding of concept (e.g., providing definitions)* • *Knowledge:* *Simple recitation of information.*			
Comments: (List thinking skills here)				

16. Interaction: Opportunities for Interaction				
4	**3**	**2**	**1**	**0**
Frequent opportunities for **interaction** and discussion between teacher / student and among students, which encourage elaborated responses about lesson concepts		**Interaction** somewhat teacher-dominated with some opportunities for students to talk about or question lesson concepts		**Interaction** primarily teacher-dominated with no opportunities for students to discuss lesson concepts
Notes: *The rating should be based on observer's assessment of entire class period. In the comments section, please provide examples of the various interactions that occurred during the lesson.* *Examples: Promoting student discussion, encouraging elaborated responses from students, giving students time to express their thoughts and feelings, providing opportunities for students to interact with one another to complete a task through jigsaw readings, think-pair-share, debates, dialog journals, experiments, etc*				
Comments: (List types of interactions here)				

17. Interaction: Flexible and Strategic Grouping				
4	**3**	**2**	**1**	**0**
A variety of **flexible and strategic grouping configurations** are evident throughout the lesson (whole class, small groups, peer assisted, independent). At each stage of the lesson, grouping configuration supports lesson objectives (both language and content).		At least two **grouping configurations** are evident during the lesson. Grouping may support only one of the lesson objectives (either language or content objectives).		No more than one **grouping configuration** is evident during the lesson and/or grouping configurations do not support language and/or content objectives.

Notes:
Flexible groups are temporary groups assigned for a specific task. Though groups may sometimes be based on skill level, they are designed to provide cooperative peer support to students in the completion of a specific step in a larger assignment. These are not ability groups designed for the teacher to provide levelized small group instruction.
For a rating of 3 the groupings should be strategic or flexible.

Comments:*(List configurations here)*

18. Interaction: Wait Time				
4	**3**	**2**	**1**	**0**
Sufficient **wait time for student responses** consistently provided		Sufficient **wait time for student responses** occasionally provided		Sufficient **wait time for student responses** not provided

Notes:

If there are multiple student proficiency levels the teacher needs to accommodate all levels (by adjusting wait time)

Wait time refers to the length of time that teachers wait for students to respond before interrupting, answering a question themselves, or calling on someone else to participate. Wait time should be provided to all students. The length of wait time, however, varies by proficiency and/or skill level.

Teachers should allow <u>beginning and intermediate</u> students additional (ample-could add 3 second time) time to formulate their answers or generate questions. The teacher should pause after a question is asked to enable students to process the question before responding. Some students may need time to translate their first language into English, others may need time to find the appropriate words for their response. <u>Advanced</u> students should be provided with wait time, but the length of wait time can vary from 1–3 seconds, depending on the processing time needed by the student.

Comments: *(Give examples of wait time here)*

19. Interaction: Clarify Key Concepts in L1

4	3	2	1	0	NA
Ample opportunities for students to **clarify key concepts in L1** as needed with aide, peer, or L1 text		Some opportunity for students to **clarify key concepts in L1**		No opportunity for students to **clarify key concepts in L1**	**NA**

Notes:
This is an item of concern only for classes where the students are at the beginning/ intermediate levels of language development. In classes with beginners there should be evidence that students receive assistance from teacher, aide, peer, or primary language materials, like bilingual dictionaries. This item should be scored N/A for advanced students, who indicate proficiency in their use of English.

Comments:

20. Practice/Application: Hands-on Materials

4	3	2	1	0	NA
Students effectively use **hands-on materials and/or manipulatives** to practice using content knowledge and improve understanding and meet lesson objectives.		Students use of **hands-on materials and/or manipulatives** does not provide meaningful practice using content knowledge or meet lesson objectives.		No **hands-on materials and/or manipulatives** for students to practice using content knowledge or improve understanding.	**NA**

Notes:
*In language arts classes it can be difficult and even a distraction to make the lesson hands-on. Writing in particular is an application activity that does not typically require manipulatives. **If there is a meaningful application activity that meets the lesson objective but is not hands on, then it should be scored as N/A**. If the activity does not meet the lesson objective and is not hands on, then it should be scored as O.*

Comments: (List hands-on materials/manipulatives here)

21. Practice/Application: Apply Content and Language Knowledge				
4	**3**	**2**	**1**	**0**
Activities provide ample opportunity for students to **apply lesson content and language knowledge** in the classroom.		Activities provide limited opportunity for students to **apply lesson content and language knowledge** in the classroom.		No activities provided for students to apply **lesson content or language knowledge** in the classroom

Notes: In order to obtain a rating of "4", the activities may not necessarily involve all four language skills (reading, writing, speaking, listening) – this is addressed in Item #22. However in order to be ample, the activities must address both content and language objectives.

Comments: *(Describe opportunities here)*

22. Practice/Application: Language Skills				
4	**3**	**2**	**1**	**0**
Activities provide opportunities to practice all **language skills** (reading, writing, listening and speaking) for the majority of the students.		Activities provide opportunities to practice some **language skills** for the majority of the students.		Activities do not provide opportunities to practice **language skills** other than passive listening.

Notes: For a rating of a 3, there must be opportunities for students to use three language skills (reading, writing, and speaking).

Comments: *(List opportunities here)*

23. Lesson Delivery: Language Arts Content Objectives

4	3	2	1	0
Current **language arts content objectives** clearly supported by lesson delivery		Current **language arts content objectives** somewhat supported by lesson delivery		Current **language arts content objectives** not supported by lesson delivery

Notes:
If item #1 of this instrument was scored as 0, then this item should also be scored 0. If stated objective(s) was (were) not fully met, then this item should be scored no higher than 2. If the content objectives were implied in item #1, this item should be scored no higher than 3.

Comments:

24. Lesson Delivery: Language Objectives

4	3	2	1	0
Current **language objectives** clearly supported by lesson delivery		Current **language objectives** somewhat supported by lesson delivery		Current **language objectives** not supported by lesson delivery

Notes:
If item #2 of this instrument was scored as 0, then this item should also be scored 0. If stated objective(s) was(were) not fully met, then this item should be scored no higher than 2. If the content objectives were implied in item #2, this item should be scored no higher than 3.

Comments:

25. Lesson Delivery: Paying Attention and On Task

4	3	2	1	0
Most students are **paying attention and on task** 90% to 100% of the period.		Most students are **paying attention and on task** approximately 70% of the period.		Most students are **paying attention and on task** less than 50% of the period.

Comments:

26. Lesson Delivery: Pacing

4	3	2	1	0
Pacing of the lesson appropriate to the students' language proficiency levels		**Pacing** generally appropriate to the students' language proficiency levels, but at times too fast or too slow		**Pacing** generally inappropriate to the students' proficiency levels

Comments:

---

27. Evaluation: Review of Vocabulary

4	3	2	1	0
Comprehensive **review or assessment of key vocabulary** relevant to current lesson		Uneven **review or assessment of key vocabulary** for current lesson		No **review or assessment of key vocabulary** for current lesson

Notes:
Review does not need to occur at the end of the lesson. If the teacher does not provide an in depth review or review all of the vocabulary words, then this item should be scored no higher than 2.

Comments: (Describe review/assessment here)

---

28. Evaluation: Review of Content Concepts

4	3	2	1	0
Comprehensive **review of key language arts content concepts**		Uneven **review of key language arts content concepts**		No **review of key language arts content concepts**

Comments:

29. Evaluation: Feedback

4	3	2	1	0
Teacher provides regular constructive **feedback** to students on their output (e.g., clarifies misunderstandings, specific to language, content, or work product).		Teacher provides limited or formulaic **feedback** to students on their output (e.g., evaluative remarks, "good job", specific constructive comments to few students).		No **feedback** provided to students on their output

Notes:
Examples of constructive feedback include (1) Teacher periodically reviews language, content, and vocabulary for the current lesson; (2) Teacher monitors group activity, where s/he walks around the class, listening to groups, making comments that guide their discussion, and asking questions to assess student understanding; (3) Teachers questions and requests clarification from individuals, small groups and whole class to help assess student knowledge and understanding; (4) Teacher paraphrases and validates students' answers in small group- and whole class- discussions; and (5) Other feedback that corrects misconceptions and misunderstandings given orally or in writing by teacher or peers Facial expressions, body language, nodding, smile, pat on shoulder should be included, but a lesson that includes only these expressions of feedback should not be rated very highly.

Comments: *(List examples of feedback here)*

30. Evaluation: Assessments of Student Comprehension

4	3	2	1	0
30. The teacher conducts **assessments of student comprehension and learning** of virtually all students for all lesson objectives.		The teacher conducts **assessments of student comprehension and learning** of some students and/or some lesson objectives.		**No assessment of student comprehension and learning** of lesson objectives

Notes:
Rating of 3: The teacher conducts assessments of student comprehension and learning of some students and <u>all</u> lesson objectives.

Comments:

QTEL 2010 Program Aligned Classroom Observation Instrument

Date:	Observer:
School/District:	Teacher:
Grade(s):	Class: Lesson Topic:
Start time: End time: *(military time)*	Lesson Type: Multi-day Single-day Don't know *(circle one)*
Solo Observation Paired Observation *Paired with_____*	Number of Students:

Directions: *Using the rubrics on the following pages, circle the number that best reflects what you observe in the lesson. You may give a score from 0-4.*
Cite under "Comments" specific examples of the behaviors/evidence observed. Make your comments as clear and complete as possible.

		Highly Evident		Somewhat Evident		Not Evident
1		**4** Activity structure guides students to support their thinking with evidence. Delivery successfully leads students to construct **well reasoned explanations**.	**3**	**2** Activity structure has potential to lead students to construct **reasoned explanations**. Delivery may undermine that potential, or student explanations may not provide sound support for ideas.	**1**	**0** Students have little or no opportunity to construct **reasoned explanations**.
	Note to the observer: Focus on the support students must provide for their ideas. Are they asked to justify or explain their choices or conclusions? Are they required to explain their responses? If so, how well do they support their reasoning?					
	Comments:					
2		**4** Activity(ies) is(are) **purposeful** (text to text, text to world, or text to self connections are evident).	**3**	**2** Activity(ies) may be somewhat **formulaic and disconnected**.	**1**	**0** This lesson consists of only one activity, or activities are a series of disconnected exercises/procedural tasks.
	Note to the observer: Purposeful activities have some meaning or utility beyond the classroom. They have some importance to the student in addition to a grade or a test. *If the activity is formulaic AND purposeful, this item should be scored no lower than a 3.* *If exercises are well integrated, this item should be scored no lower than a 1.*					
	Comments:					

	Highly Evident		Somewhat Evident		Not Evident
3	**4** Activity structure fosters **intellectual engagement from all students at all skill levels**.	**3**	**2** Activity allows for **intellectual engagement at more than one skill level**.	**1**	**0** Activity provides **no opportunity for intellectual engagement**.
	Note to the observer:				
	Comments:				
4	**4** **All students are challenged** by the activity regardless of proficiency level. There is no visible evidence of frustration/confusion.	**3**	**2** **Some students may not be sufficiently challenged**. Others may be frustrated/confused.	**1**	**0** **Students may be 'on task', but are not invested** in the outcome of the activity. No visible evidence of challenge or frustration/confusion.
	Note to the observer: The score of three may be given when there is no evidence of frustration but the observer is not convinced that there is sufficient challenge for all students. *If the activity is too challenging for most students and that seems to be the source of 'off task' behaviors, the score of 1 is appropriate.*				
	Comments:				

	Highly Evident		**Somewhat Evident**		**Not Evident**
5	**4** Activity structure creates **multiple entry points** that promote **student participation at various language proficiency levels**. Each student participates in a meaningful way, at a level appropriate to his/her language proficiency level.	**3**	**2** The activity structure provides **only one entry point**. Some students are engaged in the activity and play a meaningful role. Individual students may not have the skills to participate fully. Others may not be able to adequately utilize or expand their skills.	**1**	**0** Activity structure provides **only one entry point**. The activity divides the class by language proficiency and students are engaged in distinct activities based on their language level.

Note to the observer: Multiple entry points are evident when students can apprentice into a task. More experienced peers play a central role and less experienced peers play a more peripheral role. These roles may be informal, and are temporary and flexible. The teacher may provide additional scaffolding on the same task to students working at lower skill levels and less scaffolding to more advanced students.

Comments:

	4		**2**		**0**
6	**4** Activity(ies) result(s) in high quality class discussion (whole class, pair or small group). Discussion is **focused on central ideas of the discipline and/or on relevant intellectual issues.**	**3**	**2** Discussion **touches on the ideas of the discipline and/or on relevant intellectual issues.** Discussion may lack depth.	**1**	**0** Discussion **does not touch on central ideas of the discipline or on relevant intellectual issues**, or they may be no opportunity for authentic class discussion.

Note to the observer: Central ideas of English language arts include specific literary devices, recurring themes, characteristics of genres, etc. Check the grade level standards for other examples and appropriate level of depth.

Comments:

	Highly Evident		**Somewhat Evident**		**Not Evident**
7	**4** Activity structure guides **groups/pairs to collaborate** on the creation, interpretation or reflection on a challenging text. All students participate in the activity.	3	**2** Activity structure creates opportunity for **groups/pairs to collaborate** on the creation, interpretation or reflection on a text. Some students may not participate.	1	**0** Activity structure provides no opportunity for **groups/pairs to collaborate** on the creation, interpretation nor reflection on a text. Or if there is such opportunity, few or no students participate.
	Note to the observer: Texts include any media or material that contains relevant content to be interpreted (e.g., visuals, video, websites, etc.) If there is no text (such as when students are writing an essay without referring to any source material) this item should be scored as 0. Please explain your score in the comments section.				
	Comments:				
8	**4** Activity(ies) promote(s) **individual engagement** in the creation, interpretation or reflection on a challenging text. All students participate in the activity.	3	**2** Activity(ies) provide(s) the opportunity for **individual engagement** in the creation, interpretation or reflection on a text. Some students may not participate.	1	**0** Activity(ies) provide(s) no opportunity for **individual engagement** in the creation, interpretation nor reflection on a text. Or if there is such opportunity, few or no students participate.
	Note to the observer: Texts include any media or material that contains relevant content to be interpreted (e.g., visuals, video, websites, etc.) If there is no text (such as when students are writing an essay without referring to any source material) this item should be scored as 0. Please explain your score in the comments section.				
	Comments:				

	Highly Evident			Somewhat Evident			Not Evident
9	**4** **Whole class discussion builds on the ideas of many participants** (teacher and students) to promote student understanding. Most students are engaged listeners.	**3**		**2** **Whole class discussion builds on the ideas of some participants** (teacher and a few students). Teacher may miss opportunities to incorporate student ideas. Some students may not be fully engaged.	**1**		**0** **Whole class discussion is dominated by the teacher** (monologue rather than dialogue). Students have no opportunity to participate in the construction of knowledge. Q/A, "Ping-pong" or IRF (Interaction-Response-Feedback) pattern of interaction may predominate.
	Note to the observer: High quality discussion involves the contributions of many participants.						
	Comments:						
10	**4** Meaningful, on-task, student-student interactions are extended and **routinely sustained over several turns.**	**3**		**2** **Some on-task student-student interactions are sustained over several turns**, while some are brief exchanges.	**1**		**0** Student-student exchanges are **brief, rare or entirely off task.**
	Note to the observer:						
	Comments:						

	Highly Evident		Somewhat Evident		Not Evident
11	**4** Teacher to student interaction pattern reflects **inclusion and engagement**.	**3**	**2** Teacher may engage **some students** in dialogue, but may miss opportunities to engage with others.	**1**	**0** Students may feel **excluded and/or intimidated**.
	Note to the observer: For this item, focus on the quantity of student participation in (on topic) classroom interactions. Inclusion requires a democratic level of participation. In an inclusive classroom, the teacher seems genuinely interested in soliciting a variety of student points of view. These may be whole class, small group or individual interactions with the teacher.				
	Comments:				
12	**4** **Teacher engages in intellectual dialogue with most students** at different points in the lesson.	**3**	**2** **Teacher engages in intellectual dialogue with some individual students** at different points in the lesson.	**1**	**0** Teacher's interactions with individual students lack depth/ interactions are limited to procedural/behavioral topics.
	Note to the observer: For this item, focus on the quality of the interactions between teacher and students. Intellectual dialogue is relevant and has depth of meaning. These dialogues between the teacher and individual students may take place as part of a whole class discussion, small group discussions or as individual check-ins.				
	Comments:				

	Highly Evident			Somewhat Evident			Not Evident
13	**4** In **small groups (or pairs), most students engage in dialogue with one another about academic or intellectual concepts**. Others may be passive participants, but listen attentively. Student to student interaction reflects inclusion and engagement.	**3**		**2** In **small groups (or pairs) some students actively engage in dialogue with one another about academic or intellectual concepts**. Others may be disengaged and do not participate in dialogue.	**1**		**0** **No small group (or pair) activity provided**. Students have little or no opportunity to engage in with one another about academic or intellectual concepts.
	Note to the observer: When scoring this item, focus on student participation and the content of student-student discussion. Active listening counts just as much as vocal participation.						
	Comments:						
14	**4** The lesson content focuses on **grade level content standards**	**3**		**2** The lesson content addresses **grade level content standards, but** the focus is unclear and/or the ideas are not central to the discipline.	**1**		**0** The lesson content is **not focused on grade level content standards.**
	Note to the observer: Refer to the ELA standards to verify that the lesson content is grade and subject area appropriate.						
	Comments:						

	Highly Evident		Somewhat Evident		Not Evident
	4	**3**	**2**	**1**	**0**
15	Lesson is made up of **multiple activities** that are well integrated into a coherent whole**.**		Lesson contains **some activities** that are connected to/build on each other. Others activities may be isolated exercises that do not relate to the whole. OR all activities build on each other but are not integrated into a coherent whole.		Lesson consists of only **one activity**, or **activities are a series of disconnected exercises**/procedural tasks.
	Note to the observer: When scoring this item, focus on the quantity of activities first, then the degree to which they are integrated with one another.				
	Comments:				
16	**4** The lesson promotes higher order thinking skills by **supporting all students to connect ideas and synthesize information in original or complex ways**.	**3**	**2** The lesson provides **some opportunity for students to connect ideas and synthesize information**. Students are required to connect ideas and synthesize information in ways that are somewhat formulaic.	**1**	**0** The lesson provides students with **no opportunity to connect ideas and synthesize information**.
	Note to the observer: Refer to Bloom's taxonomy in the observer training binder.				
	Comments:				

	Highly Evident		Somewhat Evident		Not Evident
17	**4** The lesson **guides students to evaluate and apply ideas** to form valid judgments and/or sound arguments. Students are also required to support their judgments/ arguments with evidence from the text.	**3**	**2** The lesson provides **some opportunity for students to evaluate and apply ideas** to form judgments and/or arguments. However, the lesson does not necessarily require students to support their judgments/ arguments with evidence from the text.	**1**	**0** The lesson provides students with **no opportunity to evaluate and apply ideas** to form judgments and/or arguments.
	Note to the observer: Focus on the support students must provide for their ideas. Are they asked to justify or explain their choices or conclusions? Are they required to explain their responses? If so, does the evidence come from a text?				
	Comments:				
18	**4** The lesson **guides all students to solve problems creatively and/or construct new meanings/understandings**.	**3**	**2** The lesson **provides some opportunity for students to solve problems and/or construct new meanings/ understandings**. Solutions may be somewhat formulaic and new understandings may lack depth.	**1**	**0** The lesson provides students with **no opportunity to solve problems or construct new meanings/understandings**.
	Note to the observer:				
	Comments:				

	Highly Evident		Somewhat Evident		Not Evident
19	**4** Teacher provides effective scaffolding to support **student use of academic language** through both oral and written activities.	**3**	**2** Teacher provides **some opportunity for students to use academic language** through either oral or written activities.	**1**	**0** Teacher provides students **little or no opportunity to use academic language** in oral or written activities.
	Note to the observer: Refer to the Observer Training Binder for examples of procedural, verbal and social-affective scaffolding.				
	Comments:				
20	**4** Students are **exposed to a variety of texts**. Some of these texts might be challenging ones, above the grade level of students and outside the standard curricular recommendations.	**3**	**2** Students are **exposed exclusively to grade level, curricular texts**.	**1**	**0** Students are **exposed to simplified texts or texts that lack richness of language**.
	Note to the observer: Texts include any media or material that contains relevant content to be interpreted (e.g., visuals, video, websites, etc.) If there is no text (such as when students are writing an essay without referring to any source material) this item should be scored as 0. Please explain your score in the comments section.				
	Comments:				

	Highly Evident		Somewhat Evident		Not Evident
21	**4** **Teacher explicitly discusses the characteristics of language and how language functions in the subject matter/ discipline**, making reference to usage, structure and conventions of the discipline. Subject matter texts are explicitly referenced and used to illustrate examples.	**3**	**2** **Teacher makes some reference to the characteristics of the language and how language functions in the subject matter/ discipline**. Subject matter texts of the discipline may or may not be explicitly referenced.	**1**	**0** **Teacher does not explicitly discuss how language is used in the content area or disciplinary discourse at all**.
	Note to the observer:				
	Comments:				
22	**4** **Teacher frequently expands communications with students** through the use of elaboration, rich examples, alternate phrasing, etc	**3**	**2** **Teacher sometimes expands communications**, or expands communications with some students and not others.	**1**	**0** **Teacher rarely expands communications** with students and tends to use simplified language structures and vocabulary/ communications are limited to simple evaluative feedback (e.g. " good job").
	Note to the observer: Teacher should paraphrase students' utterances using academic vocabulary and disciplinary language conventions. Teacher should provide examples rich vocabulary and sophisticated structures through her own language use, while taking care to scaffold student understanding.				
	Comments:				

Appendix D. Unadjusted means for primary student-level outcomes and secondary teacher-level outcomes

This appendix contains the unadjusted means for the primary student-level impacts and secondary teacher-level impacts, as shown in chapter 4. It also includes the unadjusted means for the exploratory analyses for the secondary teacher outcomes.

Table D1. Impact analysis of student achievement on the California Standards Test of English Language Arts (CST-ELA), grades 7 and 8, spring 2010

	Unadjusted means					
Sample	Intervention (standard deviation)	Control[a] (standard deviation)	Difference (standard error)	*p*-value	Effect size	Unweighted student sample size
Grade 7 students	353.09 (57.01)	353.71 (57.53)	–0.62 (5.33)	.882	–0.01	17,837
Grade 8 students	354.91 (59.10)	355.13 (59.45)	-0.22 (5.53)	.901	–0.004	18,180

Note: The unadjusted means were obtained using multilevel regression models where only the treatment variable and study design characteristics were included (without any covariates in the model). Effect sizes were calculated by dividing impact estimates by the control-group standard deviation of the outcome variable.
a. Includes a third of the students of one school that was treated as an intervention school by QTEL. These sample members were assigned to the control group for analytical purposes because their school was consolidated from three smaller schools, one of which was originally assigned to the control group.
Source: Authors' analysis of student-level data for participating districts.

Table D2. Impact analysis of student achievement on the California Standards Test of English Language Arts, grade 7 and grade 8 limited English proficient and redesignated fluent English proficient students, spring 2010

	Unadjusted means					
Sample	Intervention (standard deviation)	Control[a] (standard deviation)	Difference (standard error)	*p*-value	Effect size	Unweighted student sample size
Grade 7 limited English proficient and redesignated fluent English proficient	333.62 (52.86)	333.51 (52.24)	0.11 (3.84)	.896	0.002	7,699
Grade 8 limited English proficient and redesignated fluent English proficient	336.88 (53.92)	336.47 (54.48)	0.41 (4.01)	.678	0.007	8,098

Note: The unadjusted means were obtained using multilevel regression models where only the treatment variable and study design characteristics were included (without any covariates in the model). Effect sizes were calculated by dividing impact estimates by the control-group standard deviation of the outcome variable.
a. Includes a third of the students of one school that was treated as an intervention school by QTEL. These sample members were assigned to the control group for analytical purposes because their school was consolidated from three smaller schools, one of which was originally assigned to the control group.
Source: Authors' analysis of student-level data for participating districts.

Table D3. Impact analysis of student achievement on the California English Language Development Test (CELDT), grade 6 and grade 7 limited English proficient students, fall 2009

	Unadjusted means					
Sample	Intervention (standard deviation)	Control[a] (standard deviation)	Difference (standard error)	*p*-value	Effect size	Unweighted student sample size
Grade 6 limited English proficient students (tested fall grade 7)	547.27 (59.66)	551.46 (66.03)	–4.19 (6.24)	.51	–0.06	2,373
Grade 7 limited English proficient students (tested fall grade 8)	554.81 (70.15)	560.53 (69.81)	–5.72 (6.50)	.63	–0.08	3,456

Note: The unadjusted means were obtained using multilevel regression models where only the treatment variable and study design characteristics were included (without any covariates in the model). Effect sizes were calculated by dividing impact estimates by the control-group standard deviation of the outcome variable.
a. Includes a third of the teachers of one school that was treated as an intervention school by QTEL. These sample members were assigned to the control group for analytical purposes because their school was consolidated from three smaller schools, one of which was originally assigned to the control group.
Source: Authors' analysis of student-level data for participating districts.

Table D4. Impact analysis of teacher outcome measures, spring 2008, 2009, and 2010

	Unadjusted means					
Impact measure and sample	Intervention (standard deviation)	Control[a] (standard deviation)	Difference (standard error)	*p*-value	Effect size	Unweighted teacher sample size
Total teacher knowledge score (all surveyed teachers, Years 2–3)	22.53 (5.20)	21.28 (4.27)	1.25* (0.52)	.016	0.29	404
Average teacher attitude score (all surveyed teachers, Years 1–3)	3.12 (0.37)	3.09 (0.38)	0.03 (0.04)	.420	0.08	623
Sheltered Instruction Observation Protocol average score (all observed teachers, Years 1–3)	2.45 (0.68)	2.48 (0.65)	–0.03 (0.08)	.710	–0.05	527

Note: The unadjusted means were obtained using multilevel regression models where only the treatment variable and study design characteristics were included (without any covariates in the model). Effect sizes were calculated by dividing impact estimates by the control-group standard deviation of the outcome variable.
a. Includes a third of the teachers of one school that was treated as an intervention school by QTEL. These sample members were assigned to the control group for analytical purposes because their school was consolidated from three smaller schools, one of which was originally assigned to the control group.
Source: Authors' analysis of teacher-level data collected for study.

Table D5. Impact analysis of teacher practice: Sheltered Instruction Observation Protocol average subscale scores, spring 2008, 2009, and 2010

| Impact measure | Unadjusted means | | Difference (standard error) | *p*-value | Effect size |
	Intervention (standard deviation)	Control[a] (standard deviation)			
Average score	2.45	2.48	–0.03	.710	–0.05
	(0.68)	(0.65)			
Average preparation and background score	2.38	2.46	–0.08	.290	–0.12
	(0.73)	(0.67)	(0.08)		
Average input and interaction score	2.72	2.66	0.07	.400	0.11
	(0.65)	(0.63)	(0.08)		
Average activity score	2.43	2.41	0.02	.807	0.03
	(0.81)	(0.78)	(0.10)		
Average delivery and evaluation score	2.28	2.40	–0.12	.169	–0.16
	(0.78)	(0.75)	(0.09)		
Unweighted teacher sample size (527)					

Note: The unadjusted means were obtained using multilevel regression models where only the treatment variable and study design characteristics were included (without any covariates in the model). Effect sizes were calculated by dividing impact estimates by the control-group standard deviation of the outcome variable.

a. Includes a third of the teachers of one school that was treated as an intervention school by QTEL. These sample members were assigned to the control group for analytical purposes because their school was consolidated from three smaller schools, one of which was originally assigned to the control group.

Source: Authors' analysis of teacher-level data collected for study.

Table D6. Impact analysis of teacher practice: Program Aligned Classroom Observation average score and average subscale scores, spring 2009 and 2010

| Impact measure | Unadjusted means | | Difference (standard error) | *p*-value | Effect size |
	Intervention (standard deviation)	Control[a] (standard deviation)			
Average overall score	2.17	2.15	0.02	.875	0.03
	(0.70)	(0.73)	(0.12)		
Average activity score	2.18	2.25	–0.08	.532	–0.10
	(0.72)	(0.77)	(0.12)		
Average discussion score	2.32	2.30	0.02	.874	0.02
	(0.83)	(0.87)	(0.14)		
Average student interaction score	2.05	1.52	0.53*	.003	0.42
	(1.08)	(1.25)	(0.17)		
Unweighted teacher sample size (206)					

Note: The unadjusted means were obtained using multilevel regression models where only the treatment variable and study design characteristics were included (without any covariates in the model). Effect sizes were calculated by dividing impact estimates by the control-group standard deviation of the outcome variable.

a. Includes a third of the teachers of one school that was treated as an intervention school by QTEL. These sample members were assigned to the control group for analytical purposes because their school was consolidated from three smaller schools, one of which was originally assigned to the control group.

Source: Authors' analysis of teacher-level data collected for study.

Appendix E. Sensitivity analyses

This appendix includes the results of sensitivity analyses that examine whether impact estimates presented in the body of this report are sensitive to model specification, procedures to deal with missing data, and sample selection procedures. The findings presented in this report were not sensitive to any of the analytical decisions. The only conclusions that differed in the sensitivity analyses were for the impact of QTEL on teacher knowledge when using listwise deletion and treating the three schools that were consolidated as an intervention school. The different conclusion for teacher knowledge using listwise deletion was a result of not controlling for multiple comparisons. There was, however, a different conclusion for teacher knowledge when treating the three consolidated schools as an intervention school, even when controlling for multiple comparisons.

Listwise deletion

Tables E1 through E4 reproduce the primary confirmatory impact estimates in chapter 4 using a different method of dealing with missing data. Instead of using zeroes for missing values and including 0/1 dummy variables in the analyses to control for the absence of data, these tables present impact estimates using a model in which observations with any missing data were excluded from the analysis. (listwise deletion). For teacher knowledge, listwise deletion changed the conclusion about the impact of the intervention without adjusting for multiple comparisons (see table E4), but the conclusion did not change when adjusting for multiple comparisons.

Table E1. Impact analysis of student achievement on the California Standards English Language Arts Test, grade 7 and 8, tested in spring 2010 (listwise deletion of missing covariates)

| | Adjusted means | | | | | |
Sample	Intervention (standard deviation)	Control[a] (standard deviation)	Difference (standard error)	*p*-value	Effect size	Unweighted student sample size
Grade 7 students	355.48 (57.70)	357.57 (58.55)	–1.16 (1.82)	.526	–0.02	14,438
Grade 8 students	357.07 (59.54)	359.38 (60.11)	0.68 (2.87)	.813	0.01	14,806

Note: Data were regression-adjusted using multilevel regression models to account for differences in baseline characteristics and study design characteristics. Effect sizes were calculated by dividing impact estimates by the control-group standard deviation of the outcome variable.
a. Includes a third of the students of one school that was treated as an intervention school by QTEL. These sample members were assigned to the control group for analytical purposes because their school was consolidated from three smaller schools, one of which was originally assigned to the control group.
Source: Authors' analysis of student-level data for participating districts.

Table E2. Impact analysis of student achievement on the California Standards Test in English Language Arts, grade 7 and grade 8 limited English proficient and redesignated fluent English proficient students, tested in spring 2010 (listwise deletion method for handling missing covariates)

| | Adjusted means | | | | | |
| | Intervention (standard deviation) | Control[a] (standard deviation) | Difference (standard error) | *p*-value | Effect size | Unweighted student sample size |
Sample						
Grade 7 students classified English language learner and redesignated fluent English proficient in 2008/09	334.28 (53.57)	334.17 (52.96)	–0.68 (2.93)	.817	–0.01	6,235
Grade 8 students English language learner and redesignated fluent English proficient in 2008/09	337.48 (54.38)	338.42 (55.05)	1.11 (3.05)	.716	0.02	6,590

Note: Data were regression-adjusted using multilevel regression models to account for differences in baseline characteristics and study design characteristics. Effect sizes were calculated by dividing impact estimates by the control-group standard deviation of the outcome variable.

a. Includes a third of the students of one school that was treated as an intervention school by QTEL. These sample members were assigned to the control group for analytical purposes because their school was consolidated from three smaller schools, one of which was originally assigned to the control group.

Source: Authors' analysis of student-level data for participating districts.

Table E3. Impact analysis of student achievement on the California English Language Development Test (CELDT) for grade 7 and grade 8 limited English proficient students identified in the 2008/09 academic year, tested in fall 2009 (listwise deletion method for handling missing covariates)

| | Adjusted means | | | | | |
| | Intervention (standard deviation) | Control[a] (standard deviation) | Difference (standard error) | *p*-value | Effect size | Unweighted student sample size |
Sample						
Grade 6 limited English proficient students (tested fall Grade7)	547.11 (59.18)	542.36 (67.86)	4.75 (16.42)	.721	0.07	1,581
Grade 7 limited English proficient students (tested fall Grade 8)	553.31 (70.42)	550.19 (73.05)	3.128 (15.66)	.678	0.04	2,653

Note: Data were regression-adjusted using multilevel regression models to account for differences in baseline characteristics and study design characteristics. Effect sizes were calculated by dividing impact estimates by the control-group standard deviation of the outcome variable.

a. Includes a third of the students of one school that was treated as an intervention school by QTEL. These sample members were assigned to the control group for analytical purposes because their school was consolidated from three smaller schools, one of which was originally assigned to the control group.

Source: Authors' analysis of student-level data for participating districts.

Table E4. Impact analysis of teacher outcome measures, for grade 6 teachers in spring 2008, grade 7 teachers in 2009, and grade 8 teachers in 2010 (listwise deletion for handling missing covariates)

| Impact measure and sample | Adjusted means | | Difference (standard error) | *p*-value | Effect size | Unweighted teacher sample size |
	Intervention (standard deviation)	Control[a] (standard deviation)				
Total teacher knowledge score (all surveyed teachers, Years 2–3)	22.96 (5.02)	21.67 (4.12)	1.29* (0.64)	.043	0.31	323
Average teacher attitude score (all surveyed teachers, Years 1–3)	3.12 (0.37)	3.07 (0.37)	0.04 (0.05)	.424	0.11	501
Sheltered Instruction Observation Protocol average score (all observed teachers, Years 1–3)	2.46 (0.68)	2.47 (0.67)	0.01 (0.10)	.937	0.01	421

Note: Data were regression-adjusted using multilevel regression models to account for differences in baseline characteristics and study design characteristics. Effect sizes were calculated by dividing impact estimates by the control-group standard deviation of the outcome variable. Standard errors for the multiyear sample were adjusted for multiple teacher responses across survey years using a robust cluster variance estimator. Statistical significance levels are indicated as ** = 1 percent and * = 5 percent.
a. Includes a third of the students of one school that was treated as an intervention school by QTEL. These sample members were assigned to the control group for analytical purposes because their school was consolidated from three smaller schools, one of which was originally assigned to the control group.
Source: Authors' analysis of teacher-level data for participating districts.

Excluding schools with no grade 6

Tables E5 through E7 exclude from the student-level impact estimates 11 schools that did not offer all grades 6–8. As chapter 2 discussed, QTEL could not have impacted students in these schools for three subsequent years as it may have in schools with all grades 6–8. It does not appear that QTEL had substantially different estimated impacts once the sample was limited to schools with grades 6–8.

Table E5. Sensitivity analysis of student achievement on the California Standards Test of English Language Arts (CST-ELA), grade 7 and grade 8 students enrolled in middle schools with grades 6–8, tested in spring 2010

| Sample | Adjusted means | | Difference (standard error) | *p*-value | Effect size | Unweighted student sample size |
	Intervention (standard deviation)	Control[a] (standard deviation)				
Grade 7 students	353.72 (58.16)	350.53 (58.11)	3.19 (3.49)	.544	0.05	12,164
Grade 8 students	353.95 (59.77)	353.25 (59.52)	0.70 (2.65)	.840	0.01	12,589

Note: Data were regression-adjusted using multilevel regression models to account for differences in baseline characteristics and study design characteristics. Effect sizes were calculated by dividing impact estimates by the control-group standard deviation of the outcome variable.
a. Includes a third of the students of one school that was treated as an intervention school by QTEL. These sample members were assigned to the control group for analytical purposes because their school was consolidated from three smaller schools, one of which was originally assigned to the control group.
Source: Authors' analysis of student-level data for participating districts.

Table E6. Sensitivity analysis of student achievement on the California Standards Test of English Language Arts (CST-ELA), grade 7 and grade 8 limited English proficient and redesignated fluent English proficient students enrolled in middle schools with grades 6–8, tested in spring 2010

| | Adjusted means | | | | | |
	Intervention (standard deviation)	Control[a] (standard deviation)	Difference (standard error)	*p*-value	Effect size	Unweighted student sample size
Sample						
Grade 7 students classified as limited English proficient and redesignated fluent English proficient in 2008/09	333.50 (53.73)	329.45 (52.15)	4.05 (3.98)	.432	0.08	5,060
Grade 8 students classified as limited English proficient and redesignated fluent English proficient in 2008/09	334.44 (54.09)	335.73 (53.88)	–1.30 (3.82)	.810	–0.02	5,448

Note: Data were regression-adjusted using multilevel regression models to account for differences in baseline characteristics and study design characteristics. Effect sizes were calculated by dividing impact estimates by the control-group standard deviation of the outcome variable. Twelve schools that only had grades 7–8 were eliminated from this analysis.
a. Includes a third of the students of one school that was treated as an intervention school by QTEL. These sample members were assigned to the control group for analytical purposes because their school was consolidated from three smaller schools, one of which was originally assigned to the control group.
Source: Authors' analysis of student-level data for participating districts.

Table E7. Sensitivity analysis of student achievement on the California English Language Development Test (CELDT), grade 7 limited English proficient students enrolled in middle schools with grades 6–8, tested in fall 2009

| | Adjusted means | | | | | |
	Intervention (standard deviation)	Control[a] (standard deviation)	Difference (standard error)	*p*-value	Effect size	Unweighted student sample size
Sample						
Grade 7 limited English proficient students (tested fall Grade 8)	557.51 (65.98)	562.00 (67.19)	–4.49 (7.87)	0.880	–0.07	2,247

Notes: Data were regression-adjusted using multilevel regression models to account for differences in baseline characteristics and study design characteristics. Effect sizes were calculated by dividing impact estimates by the control-group standard deviation of the outcome variable. Twelve schools that only had grades 7–8 were eliminated from this analysis.
a. Includes a third of the students of one school that was treated as an intervention school by QTEL. These sample members were assigned to the control group for analytical purposes because their school was consolidated from three smaller schools, one of which was originally assigned to the control group.
Source: Authors' analysis of student-level data for participating districts.

Students at their middle school for three years

Tables E8 and E9 describe impacts for grade 8 students who were in the same school for three consecutive years. This is a subset of the students in tables E5 and E6. These students' potential exposure to QTEL would have been the strongest. It does not appear that QTEL had substantially different estimated impacts once the sample was limited to these students.

Table E8. Sensitivity analysis of student achievement on the California Standards Test of English Language Arts (CST-ELA), grade 8 students enrolled for three consecutive years in their middle schools, tested in spring 2010

	Adjusted means					
Sample	**Intervention (standard deviation)**	**Control[a] (standard deviation)**	**Difference (standard error)**	***p*-value**	**Effect size**	**Unweighted student sample size**
Grade 8 students	362.01 (58.09)	360.37 (57.82)	1.64 (2.67)	.461	0.03	8,720

Note: Data were regression-adjusted using multilevel regression models to account for differences in baseline characteristics and study design characteristics. Effect sizes were calculated by dividing impact estimates by the control-group standard deviation of the outcome variable.
a. Includes a third of the students of one school that was treated as an intervention school by QTEL. These sample members were assigned to the control group for analytical purposes because their school was consolidated from three smaller schools, one of which was originally assigned to the control group.
Source: Authors' analysis of student-level data for participating districts.

Table E9. Sensitivity analysis of student achievement on the California Standards Test of English Language Arts(CST-ELA), grade 8 limited English proficient and redesignated fluent English proficient students enrolled for three consecutive years in their middle schools, tested in spring 2010

	Adjusted means					
Sample	**Intervention (standard deviation)**	**Control[a] (standard deviation)**	**Difference (standard error)**	***p*-value**	**Effect size**	**Unweighted student sample size**
Grade 8 students classified as limited English proficient and redesignated fluent English proficient in 2008/09	336.21 (50.87)	340.53 (50.83)	–4.32 (6.73)	.556	–0.08	2,449

Note: Data were regression-adjusted using multilevel regression models to account for differences in baseline characteristics and study design characteristics. Effect sizes were calculated by dividing impact estimates by the control-group standard deviation of the outcome variable.
a. Includes a third of the students of one school that was treated as an intervention school by QTEL. These sample members were assigned to the control group for analytical purposes because their school was consolidated from three smaller schools, one of which was originally assigned to the control group.
Source: Authors' analysis of student-level data for participating districts.

Three schools combined into one

Tables E10 through E13 examine the sensitivity of the impact estimates to the procedure used to address the consolidation of three schools into one. Two intervention schools and one control school merged during the study and all analyses in the body of the report randomly sampled one

third of students and teachers from this merged school to serve as members of the control group, even though they were treated as in the intervention group after the schools were merged. (This makes these students and teachers control to intervention crossovers for the analyses.) This approach best maintains the integrity of the original randomization, but the impact estimates could have been sensitive to the crossover issue, which is why tables E10 through E13 treat the entire consolidated school as an intervention school. The student impact estimates were not sensitive to this decision. The teacher knowledge impact estimate was sensitive to the consolidation, changing from an impact with an adjusted *p*-value (.051) just outside the 5 percent benchmark to an impact statistically significant at the 5 percent level.

Table E10. Impact analysis of student achievement on the California Standards English Language Arts Test, grade 7 and 8: consolidation of three schools into one intervention school, tested in spring 2010

| | Adjusted means | | | | | |
Sample	Intervention (standard deviation)	Control (standard deviation)	Difference (standard error)	*p*-value	Effect size	Unweighted student sample size
Grade 7 students	353.23 (57.01)	353.56 (57.53)	–0.34 (2.68)	.748	–0.01	17,837
Grade 8 students	355.22 (59.10)	354.80 (59.45)	0.43 (1.97)	.551	0.01	18,180

Note: Data were regression-adjusted using multilevel regression models to account for differences in baseline characteristics and study design characteristics. Effect sizes were calculated by dividing impact estimates by the control-group standard deviation of the outcome variable.
Source: Authors' analysis of student-level data for participating districts.

Table E11. Impact analysis of student achievement on the California Standards Test in English Language Arts, grade 7 and grade 8 limited English proficient and redesignated fluent English proficient students: consolidation of three schools into one intervention school, tested in spring 2010

| | Adjusted means | | | | | |
Sample	Intervention (standard deviation)	Control (standard deviation)	Difference (standard error)	*p*-value	Effect size	Unweighted student sample size
Grade 7 students classified as English language learner and redesignated fluent English proficient in 2008/09	334.91 (52.86)	332.11 (52.24)	2.80 (2.86)	.257	0.05	7,699
Grade 8 students classified as English language learner and redesignated fluent English proficient in 2008/09	337.07 (53.92)	336.25 (54.48)	0.82 (2.23)	.556	0.02	8,098

Note: Data were regression-adjusted using multilevel regression models to account for differences in baseline characteristics and study design characteristics. Effect sizes were calculated by dividing impact estimates by the control-group standard deviation of the outcome variable.
Source: Authors' analysis of student-level data for participating districts.

Table E12. Impact analysis of student achievement on the California English Language Development Test (CELDT), grade 7 and grade 8 limited English proficient students: consolidation of three schools into one intervention school, tested in spring 2010

| Sample | Adjusted means | | Difference (standard error) | *p*-value | Effect size | Unweighted student sample size |
	Intervention (standard deviation)	Control (standard deviation)				
Grade 6 students classified as limited English proficient students in 2008/09	551.09 (59.74)	547.33 (66.07)	3.76 (4.15)	.425	0.06	2,381
Grade 7 students classified as limited English proficient students in 2008/09	558.41 (70.13)	556.56 (69.69)	1.85 (5.62)	.627	0.03	3,466

Note: Data were regression-adjusted using multilevel regression models to account for differences in baseline characteristics and study design characteristics. Effect sizes were calculated by dividing impact estimates by the control-group standard deviation of the outcome variable.
Source: Authors' analysis of student-level data for participating districts.

Table E13. Impact analysis of teacher outcome measures: consolidation of three schools into one intervention school, for grade 6 teachers in spring 2008, grade 7 teachers in 2009, and grade 8 teachers in 2010

| Impact measure and sample | Adjusted means | | Difference (standard error) | *p*-value | Effect size | Unweighted teacher sample size |
	Intervention (standard deviation)	Control (standard deviation)				
Total teacher knowledge score (all surveyed teachers, Years 2–3)	22.61 (5.20)	21.18 (4.27)	1.42* (0.59)	.015	0.33	404
Average teacher attitude score (all surveyed teachers, Years 1–3)	3.13 (0.37)	3.08 (0.38)	0.05 (0.05)	.319	0.13	623
Sheltered Instruction Observation Protocol average score (all observed teachers, Years 1–3)	2.47 (0.68)	2.46 (0.65)	0.00 (0.09)	.987	0.00	527

Note: Data were regression-adjusted using multilevel regression models to account for differences in baseline characteristics and study design characteristics. Effect sizes were calculated by dividing impact estimates by the control-group standard deviation of the outcome variable. Standard errors for the multiyear sample were adjusted for multiple teacher responses across survey years using a robust cluster variance estimator. Statistical significance levels are indicated as ** = 1 percent and * = 5 percent.
Source: Authors' analysis of teacher-level data for participating districts.

Additional sensitivity analyses

The results of a test of the sensitivity of the teacher attitudes impacts to the scale model used to construct the teacher attitudes outcomes are in table E14. A Rasch model was used to construct the teacher attitudes outcome, instead of the Likert scale used for the secondary impact analyses. Similar to the secondary teacher attitudes impacts result, the teacher attitudes impact using the Rasch scale was not statistically significant, indicating that teacher attitudes impacts were not sensitive to the outcome ratings scale.

Table E14. Impact analysis of teacher attitudes outcome measure using an alternate teacher attitude ability measure constructed with a Rasch rating scale model, for grade 6 teachers in spring 2008, grade 7 teachers in 2009, and grade 8 teachers in 2010

| Impact measure and sample | Adjusted means | | Difference (standard error) | *p*-value | Effect size | Unweighted student sample size |
	Intervention (standard deviation)	Control (standard deviation)				
Average teacher attitude score (all surveyed teachers, Years 1–3)	1.27 (1.08)	1.14 (1.18)	0.13 (0.14)	.380	0.11	623

Note: Data were regression-adjusted using multilevel regression models to account for differences in baseline characteristics and study design characteristics. Effect sizes were calculated by dividing impact estimates by the control-group standard deviation of the outcome variable.
Source: Authors' analysis of teacher-level data for participating districts.

The results of a sensitivity analysis in which teachers who received more than one year of the intervention were dropped from the teacher knowledge, attitudes, and practice impact analyses are in table E15. The teacher attitudes and practice impacts do not appear to be sensitive to dropping teachers who received more than one year of the intervention from the sample; those teacher impacts were similar to the impacts found for the secondary impacts and are not statistically significant. The teacher knowledge impact using the smaller sample of teachers with only one year of the intervention is statistically significant, though this result is driven by the fact that the impact was not adjusted for multiple comparisons, while the secondary teacher knowledge impact was adjusted for multiple comparisons.

Table E15. Impact analysis of teacher outcome measures excluding teachers who received more than one year of QTEL, for grade 6 teachers in spring 2008, grade 7 teachers in 2009, and grade 8 teachers in 2010

| Impact measure and sample | Adjusted means | | Difference (standard error) | *p*-value | Effect size | Unweighted teacher sample size |
	Intervention (standard deviation)	Control (standard deviation)				
Total teacher knowledge score (all surveyed teachers, Years 2–3)	22.84 (5.07)	21.48 (4.05)	1.36* (0.56)	.016	0.34	342
Average teacher attitude score (all surveyed teachers, Years 1–3)	3.13 (0.37)	3.08 (0.38)	0.05 (0.05)	.267	0.13	563
Sheltered Instruction Observation Protocol average score (all observed teachers, Years 1–3)	2.47 (0.68)	2.46 (0.64)	–0.03 (0.09)	.744	–0.05	460

Note: Data were regression-adjusted using multilevel regression models to account for differences in baseline characteristics and study design characteristics. Effect sizes were calculated by dividing impact estimates by the control-group standard deviation of the outcome variable. Standard errors for the multiyear sample were adjusted for multiple teacher responses across survey years using a robust cluster variance estimator. Statistical significance levels are indicated as ** = 1 percent and * = 5 percent.
Source: Authors' analysis of teacher-level data for participating districts.

Appendix F. Detailed tables and discussion of student sample

A sample flow diagram for English language arts achievement for the full sample of students who were in grade 8 in school year 2009/10 and who participated in the California Standards Test of English Language Arts (CST-ELA) in that year is shown in figure F1. The study began with 54 randomly assigned schools. Of these, as discussed in "Random assignment of schools," both the schools in a single district with two participating schools (one intervention school and one control school) dropped out shortly after random assignment, eliminating the entire district from the study.

Figure F1. Sample flow diagram for grade 8 students

Random assignment of schools	
54 schools	

QTEL professional development	Instruction as usual
Intervention: 27 schools	Control: 27 schools

Year 1 2007/08
Grade 6 students: n = 6,382 — 26 schools (Intervention)
Grade 6 students: n = 6,000 — 26 schools (Control)
Left study n = 1,162 / Left study n = 1,233
Joined study n = 4,010 / Joined study n = 3,505

Year 2 2008/09
Grade 7 students: n = 9,230 — 25 schools (Grade 6 to 7 stayers: n = 5,220)
Grade 7 students: n = 8,272 — 25 schools (Grade 6 to 7 stayers: n = 4,767)
Left study n = 1,310 / Left study n = 1,168
Joined study n = 1,635 / Joined study n = 1,521

Year 3 2009/10
Grade 8 students: n = 9,555 — 25 schools (Grade 6 to 8 stayers: n = 4,554; Grade 7 to 8 stayers: n = 3,366)
Grade 8 students: n = 8,625 — 25 schools (Grade 6 to 8 stayers: n = 4,166; Grade 7 to 8 stayers: n = 2,938)

Grade 8 impact analysis sample, Year 3
18,180 students, 50 schools

Note: Final student counts reflect grade 8 students with CST-ELA scores. During the baseline year, one school district containing two schools dropped out of the study. During Year 2, three schools from the same district were consolidated. At baseline and Year 1, two of these schools were in the intervention group and one was in the control group. The consolidation during Year 2 resulted in one intervention school. In the chapter 4 impact analyses, one-third of the students in this school were moved to the control group for model estimation. Grades 6–8 stayers are students who potentially had three years of exposure to QTEL teachers. Since 12 schools did not have a grade 6, the grades 6–8 stayers include only students who attended study schools with a grade 6.
Source: Student-level data for participating districts, analyzed by authors.

Next, the figure shows counts of individual students who participated in the CST-ELA.[50] In Year 1 (2007/08), the study included 6,382 intervention students and 6,000 control students in grade 6. Of these students, 81.8 percent in the intervention group (5,220) and 79.5 percent in the control group (4,767) were in grade 7 in the same schools in Year 2. The difference in one-year student retention between the intervention and control schools was not statistically significant ($p=.973$). The remaining grade 7 students in Year 2 (4,010 in the intervention group and 3,505 in the control group) were new students who had not been in grade 6 in the same school the previous year. Of the new students, 71.0 percent were enrolled in schools that did not offer a grade 6 (2,829 in the intervention group and 2,508 in the control group). The difference in new grade 7 enrollment between the intervention and control schools was not statistically significant ($p=.935$). In Year 3, the study included 71.4 percent of the original grade 6 students from Year 1 in intervention schools (4,554) and 69.4 percent of those in control schools (4,166). The difference in two-year student retention between the intervention and control schools was not statistically significant ($p=.872$). Thus, overall three-year sample attrition among Year 1 grade 6 students was 29.5 percent (28.6 percent in the intervention group and 30.5 percent in the control group).[51]

The remaining grade 8 students in Year 3 (5,001 in the intervention group and 4,459 in the control group) were students who had not been in grade 6 in the same school in Year 1. Of these grade 8 students, 34.7 percent joined the school in grade 7 (35.2 percent for intervention schools and 34.1 percent for control schools; the difference was not statistically significant; $p=.679$) and 17.4 percent joined the school in grade 8 (17.1 percent for intervention schools and 17.6 percent for control schools; the difference was not statistically significant; $p=.639$).

As discussed in "Study design," the study focused on both grade 7 and grade 8 students in Year 3. A sample flow diagram for students assessed with the CST-ELA in grade 7 in the 2009/10 school year is shown in figure F2. As shown in the figure, retention rates for grade 7 students in the 2009/10 school year were as follows: 75.9 percent of students in the intervention group (4,686) were in grade 6 in the same school in Year 2 and 80.3 percent of students in the control group (4,546) were in grade 6 in the same school in Year 2. The difference in one-year student retention was not statistically significant ($p=.789$).

---

[50] The samples for student-level English language arts outcome measures include only students who participated in the California Standards Test of English Language Arts. The study team has no data for students who did not participate in these tests. Using published data from the California Department of Education, the study team determined that aggregate rates of assessment completion ranged from 83.4 percent to 98.5 percent in the study schools, with an average rate of 92.2 percent. There were no statistically significant differences in assessment rates between the intervention and control groups ($p = .543$).

[51] The study team did not have indicators of students repeating a grade. Because of this, a student who did not show up in the subsequent grade was counted as having left the sample and a student who showed up in the same grade twice was counted as a new student.

Figure F2. Sample flow diagram for grade 7 students

```
                           ┌─────────────────────────────┐
                           │ Random assignment of schools │
                           │          54 schools          │
                           └─────────────────────────────┘
                    ┌──────────────────┴──────────────────┐
        ┌───────────────────────────┐      ┌───────────────────────────┐
        │ QTEL professional development │   │   Instruction as usual     │
        │  Intervention: 27 schools  │      │   Control: 27 schools      │
        └───────────────────────────┘      └───────────────────────────┘
```

| Year 1 2007/08 | Students had not yet entered study 26 schools | Students had not yet entered study 26 schools |

| Year 2 2008/09 | Grade 6 students: n = 6,173 / 25 schools → Left study n = 1,487 | Grade 6 students: n = 5,663 / 25 schools → Left study n = 1,117 |

| Year 3 2009/10 | Joined study n = 4,475 → Grade 7 students: n = 9,161 / 25 schools / (Grade 6 to 7 stayers: n = 4,686) | Joined study n = 4,130 → Grade 7 students: n = 8,676 / 25 schools / (Grade 6 to 7 stayers : n = 4,546) |

```
                    ┌────────────────────────────────────────┐
                    │ Grade 7 impact analysis sample, Year 3  │
                    │       17,837 students, 50 schools       │
                    └────────────────────────────────────────┘
```

Note: Final student counts reflect grade 7 students with California Standards Test of English Language Arts scores. During the baseline year, one school district containing two schools dropped out of the study. During Year 2, three schools from the same district were consolidated. At baseline and Year 1, two of these schools were in the intervention group and one was in the control group. The consolidation during Year 2 resulted in one intervention school. In the chapter 4 impact analyses, one-third of the students in this school were moved to the control group for model estimation. Grades 6–7 stayers are students who potentially had two years of exposure to QTEL teachers. Since 12 schools did not have a grade 6, the grades 6–7 stayers include only students who attended study schools with a grade 6.
Source: Student-level data for participating districts, analyzed by authors.

## Student sample for English language arts achievement outcomes

A sample flow diagram for English language learner students who were in grade 8 in the 2009/10 school year and who participated in the CST-ELA in that year is shown in figure F3.[52] For this study, English language learner students were defined as students classified as either limited English proficient or redesignated fluent English proficient at the beginning of the school year. This classification was based partly on students' test scores on the California English Language Development Test (CEDLT), which is administered at the beginning of each school year (in September or October).[53] Students who entered a grade limited English proficient were required to take the test, as were students who were new to the school system and who reported that the primary language at home was not English. These students were then classified as initially fluent English proficient, redesignated fluent English proficient, or limited English proficient, based partly on the results of the test. Because this study's English language learner student subsamples included both limited English proficient and redesignated fluent English proficient students, reclassification of students over time did not affect the overall composition of these samples. (No students are ever redesignated into initially fluent English proficient status.)

---

[52] As described previously, the samples for student-level English language arts achievement measures include only students who participated in the California Standards Test of English Language Arts. The study team has no data for students who did not participate in these tests. The California Department of Education provides enrollment information for different classifications of English proficiency at the middle school level but not at the grade level. Therefore, it was not possible to estimate the assessment completion rate for different English language subgroups by grade.

[53] More details about the process through which students are classified as limited English proficient or redesignated fluent English proficient are in "Key outcomes and measurement."

Figure F3. Sample flow diagram for grade 8 limited English proficient and redesignated fluent English proficient students
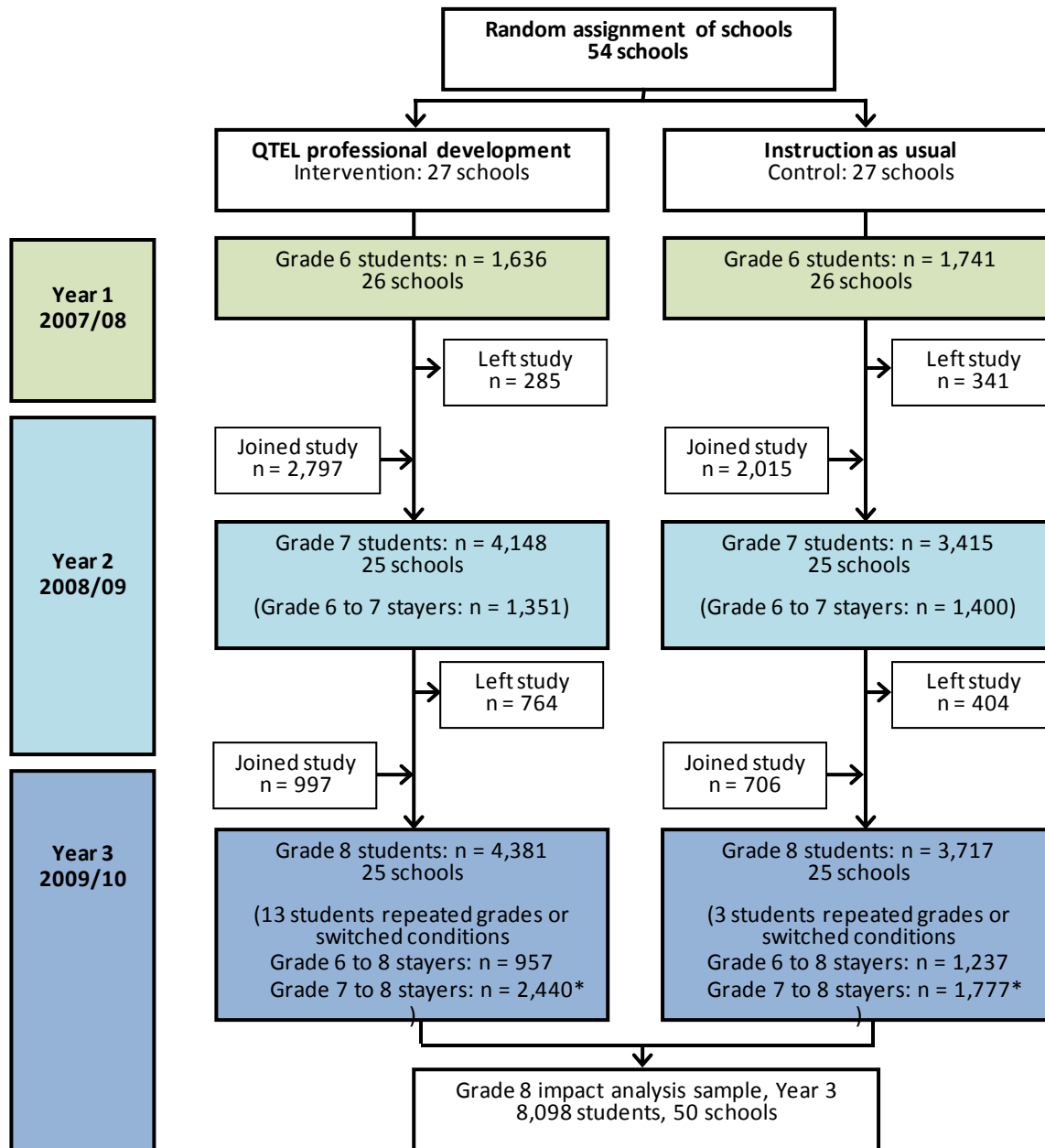
*This number does not include students who repeated grades or switched treatment conditions.

Note: Final student counts reflect grade 8 limited English proficient students and redesignated fluent English proficient students with CST-ELA scores. During the baseline year, one school district containing two schools dropped out of the study. During Year 2, three schools from the same district were consolidated. At baseline and Year 1, two of these schools were in the intervention group and one was in the control group. The consolidation during Year 2 resulted in one intervention school. In the chapter 4 impact analyses, one-third of the students in this school were moved to the control group for model estimation. Grades 6–8 stayers are students who potentially had three years of exposure to QTEL teachers. Since 12 schools did not have a grade 6, the grades 6–8 stayers include only students who attended study schools with a grade 6.
*Source:* Student-level data for participating districts, analyzed by authors.

Combined counts of individual limited English proficient and redesignated fluent English proficient students (together labeled as "English language learner students") who participated in the CST-ELA in the 2009/10 school year are shown in figure F3. In Year 1 (2007/08), the study included 1,636 intervention English language learner students and 1,741 control English language learner students in grade 6. Of these students, 82.6 percent in the intervention group (1,351) and 80.4 percent in the control group (1,400) were in grade 7 in the same schools in Year 2. The difference in one-year student retention between the intervention and control schools was not statistically significant ($p=.836$). The remaining grade 7 English language learner students in Year 2 (2,797 in the intervention group and 2,015 in the control group) were new students who had not been in grade 6 in the same school in the previous school year. Of these new students, 51.5 percent were enrolled in schools that did not offer a grade 6 (1,398 in the intervention group and 1,078 in the control group). The difference in new grade 7 enrollment of English language learner students between the intervention and control schools was not statistically significant ($p=.764$). In Year 3, the study included 58.5 percent of the original grade 6 English language learner students from Year 1 in the intervention schools (957) and 71.1 percent of the original grade 6 English language learner students from Year 1 in the control schools (1,237). The difference in two-year English language learner student retention between the intervention and control schools was not statistically significant ($p=.735$). Thus, overall three-year sample attrition among Year 1 grade 6 English language learner students was 35.0 percent (41.5 percent in the intervention group and 28.9 percent in the control group).

The remaining grade 8 English language learner students in Year 3 (3,424 in the intervention group and 2,480 in the control group) had not been in grade 6 in the same school in Year 1. Among these students, 52.1 percent joined the school in grade 7 (55.7 percent for the intervention schools and 47.8 percent for the control schools; the difference was not statistically significant; $p=.344$) and 21.0 percent joined the school in grade 8 (22.8 percent for the intervention schools and 19.0 percent for the control schools; the difference was not statistically significant; $p=.605$).

As discussed in "Study design," the study focused on both grade 7 and grade 8 English language learner students in Year 3 (2009/10). A sample flow diagram for English language learner students assessed using the CST-ELA in grade 7 in the 2009/10 school year is shown in figure F4. Retention rates for grade 7 English language learner students in the 2009/10 school year were as follows: 73.0 percent of students in the intervention group (1,858) were in grade 6 in the same school in Year 2 and 83.4 percent of students in the control group (1,900) were in grade 6 in the same school in Year 2. The difference in one-year retention between the intervention and control schools was not statistically significant ($p=.942$).

Figure F4. Sample flow diagram for grade 7 limited English proficient and redesignated fluent English proficient students in Year 2 and Year 3
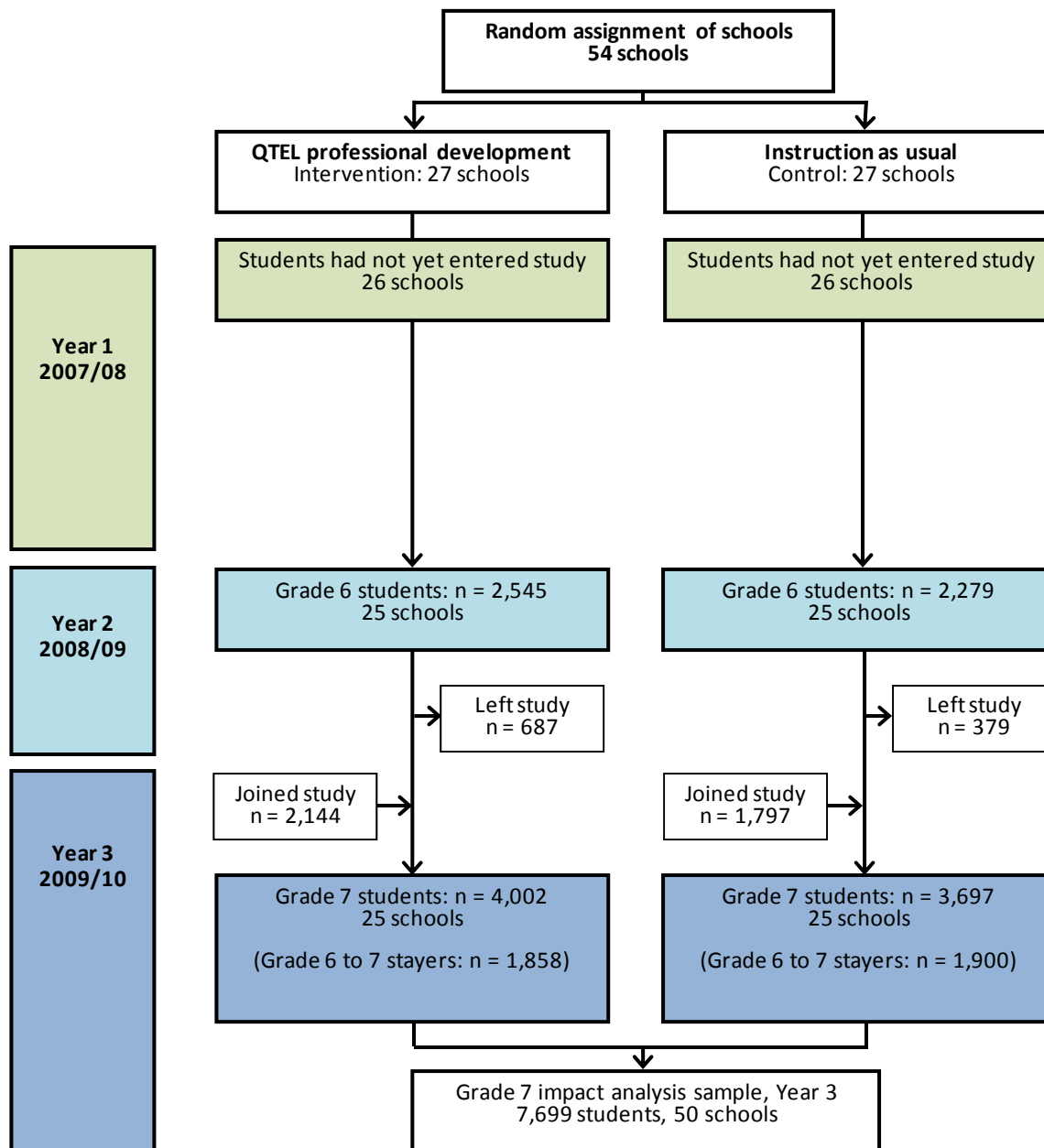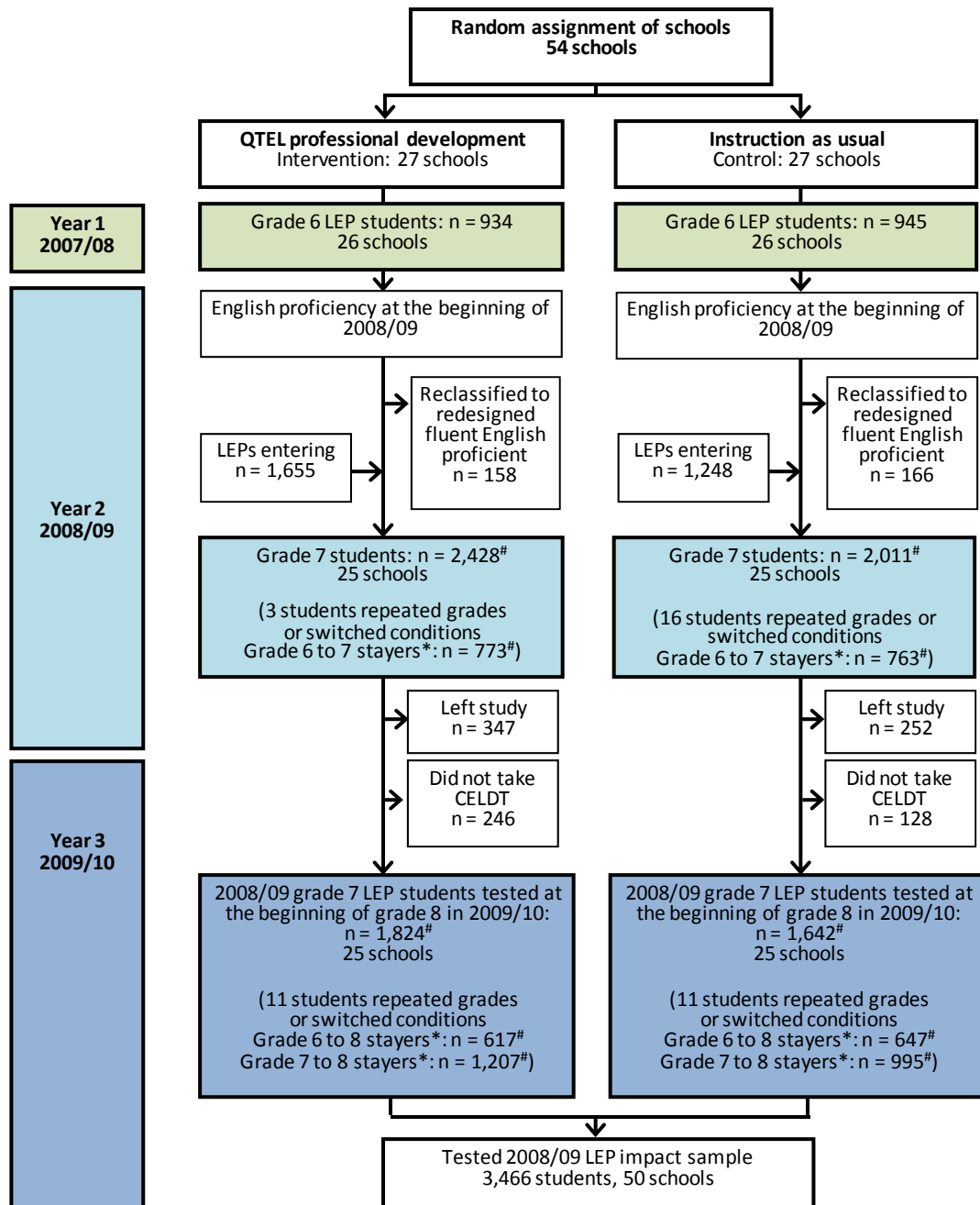

Note: Final student counts reflect grade 7 limited English proficient students and redesignated fluent English proficient students with CST-ELA scores. During the baseline year, one school district containing two schools dropped out of the study. During Year 2, three schools from the same district were consolidated. At baseline and Year 1, two of these schools were in the intervention group and one was in the control group. The consolidation during Year 2 resulted in one intervention school. In the chapter 4 impact analyses, one-third of the students in this school were moved to the control group for model estimation. Grades 6–7 stayers are students who potentially had two years of exposure to QTEL teachers. Since 12 schools did not have a grade 6, the grades 6–7 stayers include only students who attended study schools with a grade 6.
Source: Student-level data for participating districts, analyzed by authors.

Student sample for English language development outcomes

A sample flow diagram for grade 7 students who had been classified as limited English proficient in the2008/09 school year, who were therefore required to attend English language development classes during that school year, and who were assessed using the CELDT at the beginning of the 2009/10 school year is shown in figure F5. These 2009/10 CELDT scores constituted an outcome for the English language development instruction students received in the 2008/09 school year.

Figure F5. Sample flow diagram for grade 7 limited English proficient students

*Stayers is defined as students who continued to be classified as limited English proficient and stayed in the same school.
#This number does not include students who repeated grades or switched treatment conditions.
LEP is limited English proficient; CELDT is California English Language Development Test.
Note: Final student counts reflect limited English proficient students in grade 7 in 2008/09 with CELDT scores in 2009/10. During the baseline year, one school district containing two schools dropped out of the study. During Year 2, three schools from the same district were consolidated. At baseline and Year 1, two of these schools were in the intervention group and one was in the control group. The consolidation during Year 2 resulted in one intervention school. In the chapter 4 impact analyses, one-third of the students in this school were moved to the control group for model estimation purposes. Grades 6–7 stayers are students who potentially had two years of exposure to QTEL teachers. Since 12 schools did not have a grade 6, the grades 6–7 stayers include only students who attended study schools with a grade 6.
Source: Student-level data for participating districts, analyzed by authors.

Counts of grade 7 students who were classified as limited English proficient in the 2008/09 school year and who participated in the 2009/10 CELDT shortly after they entered grade8 are also shown in figure F5. In Year 1(2007/08), there were 934 limited English proficient students in grade 6 in intervention schools and 945 in control schools. Of these students, 82.8 percent in the intervention group (773) and 80.7 percent in the control group (763) were in grade 7 in the same schools in Year 2 and were still classified as limited English proficient. The difference in one-year student retention was not statistically significant ($p=.924$). Neither was the difference in one-year student reclassification from limited English proficient to redesignated fluent English proficient ($p=.840$). Moreover, at 17.2 percent, the rate of reclassification was not so large that it could have had a large effect on any subsequent outcomes measured for students who were not reclassified. That is, the inherent English language proficiency of those reclassified in the intervention group and those reclassified in the control group would have had to be very substantial for such a bias to materially affect the impact estimates based on the remaining 82.8 percent of the limited English proficient sample.

The remaining grade 7 students classified as limited English proficient in Year 2 (1,655 in the intervention group and 1,248 in the control group) were new students who had not been in grade 6 in the same school the previous year. Of these students, 46.0 percent were enrolled in schools that did not offer a grade 6 (731 in the intervention group and 604 in the control group). The difference in new grade 7 enrollment of limited English proficient students between the intervention and control schools was not statistically significant ($p=.862$). At the beginning of Year 3, the CELDT was administered to 66.1 percent of the original grade 6 limited English proficient students from Year 1in intervention schools (617) and 68.5 percent of those in control schools (647). The difference between the intervention and control schools in these students' rate of 2009/10 testing using the CELDT was not statistically significant ($p=.612$).

The remaining grade 7 students classified as limited English proficient in the 2008/09 school year were tested using the CELDT in the 2009/10 school year (1,207 in the intervention group and 995 in the control group) were not in grade 6 in the same school in Year 1. The intervention–control difference between the rates at which these students joined the sample of limited English proficient students in 2008/09 and were tested in 2009/10 was not statistically significant ($p=.637$).

As with the analysis of English language arts achievement samples, the study also focused on both grade 6 and grade 7 students classified as limited English proficient in 2008/09 and tested using the CELDT in 2009/10. A sample flow diagram for grade 6 students classified as limited English proficient in 2008/09 and assessed with the California CELDT in 2009/10 (when they had just entered grade 7) is shown in figure F6.The percentages of students classified as limited English proficient in 2008/09 who were tested in 2009/10 were 70.3 percent in intervention schools and 76.6 percent in control schools. The difference was not statistically significant ($p = .721$).

Figure F6. Sample flow diagram for grade 6 limited English proficient students
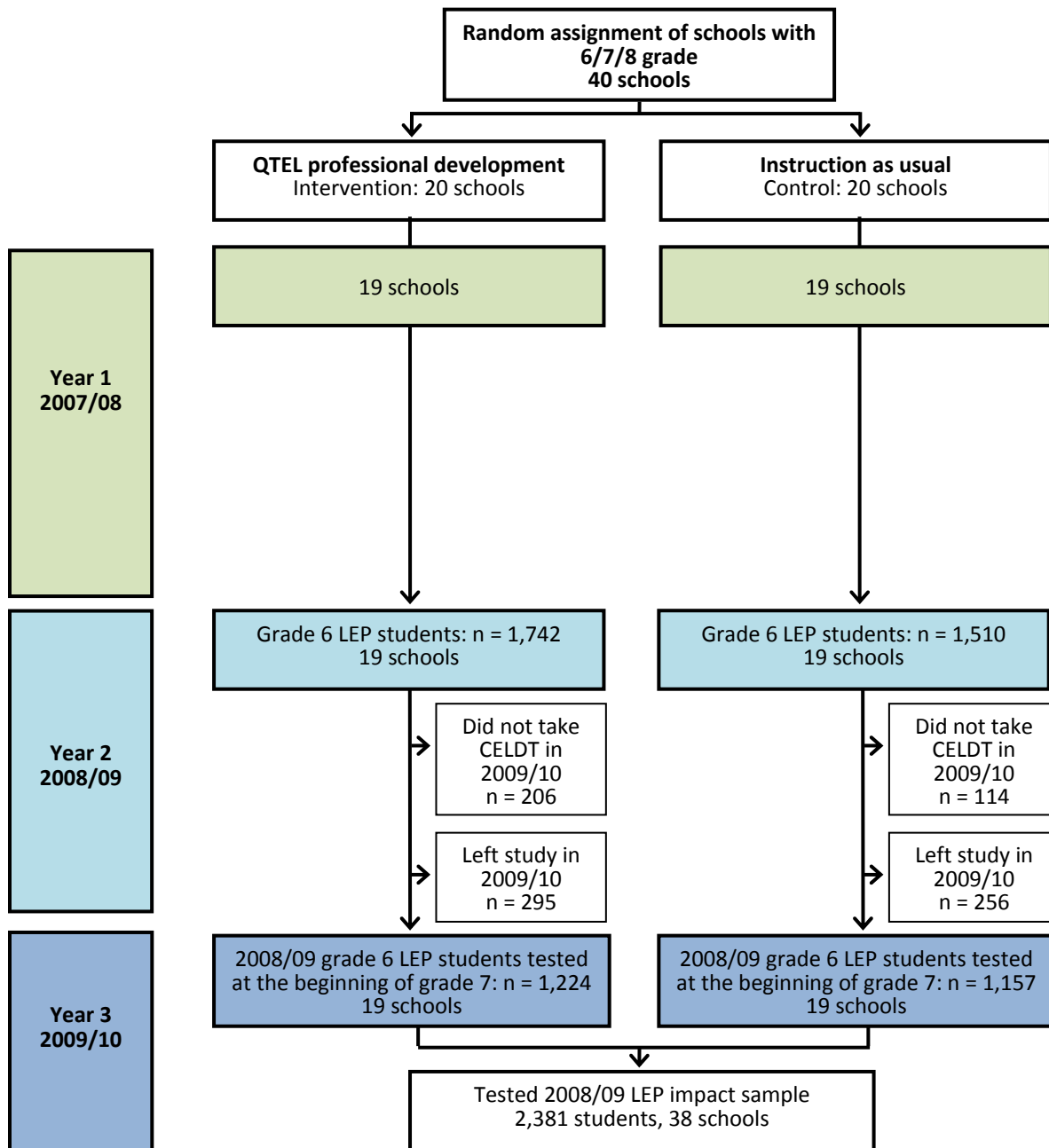
LEP is limited English proficient; CELDT is California English Language Development Test.

Note: Final student counts reflect limited English proficient students in grade 6 in 2008/09 with CELDT scores in 2009/10. During the baseline year, one school district containing two schools dropped out of the study. During Year 2, three schools from the same district were consolidated. At baseline and Year 1, two of these schools were in the intervention group and one was in the control group. The consolidation during Year 2 resulted in one intervention school. In the chapter 4 impact analyses, one-third of the students in this school were moved to the control group for model estimation. Samples include only schools that teach grade 6 as well as grades 7 and 8.

Source: Student-level data for participating districts, analyzed by authors.

References

Abedi, J., & Dietel, R. (2004, Winter). *Challenges in the No Child Left Behind Act for English language learners* (CRESST Policy Brief). Los Angeles: UCLA CSE/CRESST.

Anderson, J. A. (1977). Neural models with cognitive implications. In D. LaBerge & S. J. Samuels (Eds.), *Basic processes in reading: Perception and comprehension* (27–90). Hillsdale, NJ: Erlbaum Associates.

Angrist, J. D., Imbens, G. W., & Rubin, D. B. (1996). Identification of causal effects using instrumental variables. *Journal of the American Statistical Association, 91,* 444–455.

Arizona English Language Learners Task Force. (2007). *Structured English immersion models of the Arizona English Language Learners Task Force.* Retrieved February 2, 2010, from http://www.azed.gov/elltaskforce/2007/SEIModels9-15-07.pdf

August, D., & Hakuta, K. (Eds.). (1997). *Improving schooling for language-minority children.* Washington, DC: National Academy Press.

Batalova, J., Fix, M., & Murray, J. (2007). *Measures of change: The demography and literacy of adolescent English learners: A report to Carnegie Corporation of New York.* Washington, DC: Migration Policy Institute.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A new and powerful approach to multiple testing. *Journal of the Royal Statistical Society*, Series B, 57(1), 1289–1300.

Bloom, H. S., Bos, J., & Lee, S. (1999). Using cluster random assignment to measure program impacts: Statistical implications for evaluation of education programs. *Evaluation Review, 23*(4), 445–469.

Bolton, G (2010) *Reflective Practice, Writing and Professional Development* (3rd edition), SAGE publications, California.

Bransford, J., Brown, A., & Cocking, R. (Eds.). (2000). *How people learn: brain, mind, experience and school.* Washington, DC: National Academy Press.

California Department of Education. (2002). *Three-year plan for the development of California's assessment system.* Retrieved November 28, 2009, from http://www.cde.ca.gov/ta/tg/sa/

California Department of Education. (2007). *Standardized Testing and Reporting (STAR) program explaining 2007 STAR test results to parents and guardians.* Retrieved January 27, 2010, from http://star.cde.ca.gov/star2007

California Department of Education. (2008a). *California English Language Development blueprint preface.* Retrieved November 28, 2009, from http://www.cde.ca.gov/ta/tg/el/resources.asp.

California Department of Education. (2008b). *Technical report for the California English Language Development Test (2007–2008).* Retrieved January 27, 2010, from http://www.cde.ca.gov/ta/tg/el/techreport.asp

California Department of Education. (2009a). *California English Language Development Test (CELDT) explaining 2008/09 summary results to the public.* Retrieved January 27, 2010, from http/://www.cde.ca.gov/ta/tg/el/

California Department of Education. (2009b). *California Star Test technical report, spring 2008 administration.* Retrieved January 27, 2010, from http://www.cde.ca.gov/ta/tg/el/techreport.asp

California Education Code, Title 1. General Education Code Provisions, Division 1, Part 1, Chapter 3. English Language Education for Immigrant Children, Article 3.5.English Language Proficiency, Section 313(d). (2010). Retrieved January 27, 2010, from http://www.leginfo.ca.gov/cgi-bin/calawquery?codesection=edc&codebody=&hits=20

Capps, R., Fix, M., Murray, J., Ost, J., Passel, J. S., & Herwantoroa, S. (2005). *The new demography of America's schools: Immigration and the No Child Left Behind Act.* Washington, DC: Urban Institute.

Clegg, J. (1996). *Case studies in integrating ESL students into the mainstream curriculum.* Clevedon, UK: Multilingual Matters.

Echevarria, J., Vogt, M., & Short, D. J. (2004). *Making content comprehensible for English learners: the SIOP model* (2nd ed.). Upper Saddle River, NJ: Pearson Education, Inc.

Farr, B. (2006). *Report of a field study of quality teaching for English language learners (QTEL).* San Francisco: WestEd, Regional Educational Laboratory West.

Gándara, P., & Orfield, G. (2010). *A return to the "Mexican room": The segregation of Arizona's English learners.* Los Angeles: The Civil Rights Project.

Goldenberg, C. (2008). Teaching English language learners: What the research does—and does not—say. *American Educator, 32*(2), 8–44.

Hill, C. J., Bloom, H. S. Black, A. R., & Lipsey, M. W. (2008). Empirical Benchmarks for Interpreting Effect Sizes in Research. *Child Development Perspectives, 2*(3), 172‑177,

Hopstock, P., & Stephenson, T. (2003). *Descriptive study of services to LEP students and LEP students with disabilities. Special topic report #2: Analysis of Office of Civil Rights (OCR) data related to LEP students.* Washington, DC: U.S. Department of Education, Office of English Language Acquisition, Language Enhancement, and Academic Achievement for Limited English Proficient Students.

Jepsen, C. (2009, January). *Bilingual education and English proficiency* (University of Kentucky Center for Poverty Research Discussion Paper Series No. DP2009-01). Retrieved April 4, 2010, from http://www.ukcpr.org/Publications/DP2009-01.pdf

Masters, G.N. (1982). A Rasch model for partial credit scoring. *Psychometrika 47*, 149174.

Mora, J. K. (2010, May). *Overstated optimism: Arizona's Structured English Immersion program under Horne v. Flores.* Paper presented at the 2010 annual meeting of the American Educational Research Association, Denver, CO.

National Center for Education Statistics. (2002). *Dropout rates in the United States: 2000.* Washington DC: U.S. Department of Education. Retrieved April 17, 2010, from http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2002114

Orr, L. L. (1999). *Social experiments: Evaluating public programs with experimental methods.* Thousand Oaks, CA: Sage.

Parsad, B., Lewis, L., & Farris, E. (2001). Teacher preparation and professional development: 2000. *Education Statistics Quarterly, 3*(3), 33–36. Washington, DC: U.S. Department of Education, National Center for Education Statistics. Retrieved April 27, 2010, from http://nces.ed.gov/programs/quarterly/Vol_3/3_3/q3-3.asp

Ruiz-de-Velasco, J., & Fix, M. (2000). *Overlooked and underserved: immigrant students in U.S. secondary schools.* Washington, DC: The Urban Institute.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental design for generalized causal inference.* Boston: Houghton-Mifflin.

Schon, D. (1983) *The Reflective Practitioner, How Professionals Think In Action*, Basic Books.

Slavin, R., Madden, N., Calderon, M., Chamberlain, A., & Hennessy, M. (2010). *Reading and language outcomes of a five-year randomized evaluation of transitional bilingual education.* Retrieved April 4, 2010, from http://www.edweek.org/media/bilingual_pdf.pdf

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

U.S. Department of Education (2008). WWC Procedures and standards handbook: Version 2.0. Retrieved September 19, 2011 from http://ies.ed.gov/ncee/wwc/references/idocviewer/Doc.aspx?docId=19&tocId=4#attrition

U.S. Department of Education. (2010). *A blueprint for reform: The reauthorization of the Elementary and Secondary education Act.* Retrieved January 27, 2010, from http://www2.ed.gov/policy/elsec/leg/blueprint/blueprint.pdf

Valdes, G. (2001). *Learning and not learning English: Latino students in American schools.* New York: Teachers College Press.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes.* Cambridge, MA: Harvard University Press.

Walqui, A. (2000). *Access and engagement: Program design and instructional approaches for immigrant students in secondary school.* McHenry, IL: Center for Applied Linguistics and Delta Systems.

Walqui, A., and van Lier, L. (2010). *Scaffolding the academic success of adolescent English language learners: A pedagogy of promise.* San Francisco: WestEd.

Working Group on ELL Policy. (2009). *The American Recovery and Reinvestment Act: recommendations for addressing the needs of English language learners.* Retrieved January 8, 2010, from http://ellpolicy.org/wp-content/uploads/2009/10/ELL-Stimulus-Recommendations.pdf

Working Group on ELL Policy. (2010). *Recommendations for the reauthorization of the Elementary and Secondary Education Act.* Retrieved April 4, 2010, from http://ellpolicy.org/wp-content/uploads/2010/04/ESEA.pdf