

# Evaluation of the Quality Teaching for English Learners (QTEL) Program

Supporting Statement Part B

## **Prepared for**

US. Department of Education  
Institute of Education Sciences  
555 New Jersey Avenue, NW, Room 506 E  
Washington, DC 20208

## **Prepared by**

Neal Finkelstein, Ph.D.  
Regional Educational Laboratory (West)  
WestEd  
730 Harrison Street  
San Francisco, CA 94107

Raquel Sanchez, Ph.D.  
Berkeley Policy Associates  
440 Grand Ave, Suite 500  
Oakland, CA 94610

April 20, 2007

## Introduction

The *Quality Teaching of English Learners* (QTEL) program, is a model of professional development for teachers of secondary English Language Learners (ELL), developed by WestEd's Teacher Professional Development Program with the funding from Regional Educational Laboratory (West) (REL West). The QTEL program of teacher professional development is based on the premise that improving the education of secondary ELLs requires teachers to:

- 1) Develop deep knowledge of what it means for ELLs to participate in academic activity;
- 2) Understand, take part in, and reflect on research-based practices that support students' development of academic literacy in English as well as deep content knowledge in academic subjects; and
- 3) Receive the necessary support to change their classroom practice.

WestEd's QTEL staff work with middle school teachers in sustained collaborative relationships to introduce them into communities of educators with a shared vision of quality education and how to enact it.

The primary purpose of the study is to measure the impact of the QTEL professional development for teachers on student achievement. A secondary study goal is to examine the extent to which schools and teachers receive the QTEL training and coaching as intended and the extent to which the QTEL model is implemented as intended. This study will also serve to inform future program improvement and replication of the QTEL program. Moreover, it will provide documentation of the details of QTEL implementation for use by other institutions and entities. WestEd and its partner, Berkeley Policy Associates (BPA), are conducting this study for the Institute of Education Sciences (IES) of the US Department of Education.

At the request of educators in San Diego County, approximately 50 middle schools, 600 teachers, and about 16,000 sixth, seventh, and eighth grade students will be involved in a three-year longitudinal study of the impact of QTEL on student achievement patterns. The study will also assess program impacts on teacher and classroom outcomes. Eligible middle schools in participating districts, those schools with at least 10 percent of the classified as Limited English Proficient (LEP) or Redesignated as Fluent English Proficient (RFEP), will be grouped together by their location, and by their student, teacher, and school characteristics. The 50 middle schools selected and their participating teachers will then be randomly assigned within the resulting strata to treatment and control conditions.

## 1. Universe of Respondents/Sampling Methods

This study will use a cluster random assignment design with schools as the unit of random assignment. The potential universe of schools includes a large number of middle schools in the Western Region. Responding to the demand for these services by educators, the setting we selected for the study is San Diego County, which includes San Diego, the second-largest city in California, and a large number of other cities and districts. San Diego County has close to three million inhabitants (2005 Census estimate<sup>1</sup>) and is the sixth largest county in the U.S. One third of all county residents are Latino and close to 40 percent do not speak English at home. Therefore, the instruction of ELLs is a major concern for San Diego County schools.

The universe of teachers consists of those ESL/ELA teachers in the sixth, seventh, and eighth grades in the above schools. Similarly, the universe of students includes all students enrolled in the sixth, seventh, and eighth grades of the middle schools in the San Diego County. (Some districts only have seventh and eighth grades in their middle schools, in which case the study will include only students in those grades). Among these students, the study will focus especially on those who are classified as ELLs.

At least 50 schools in San Diego County will be involved in data collection that have high proportions of ELLs. For the study, we are defining the sampling frame as middle schools that have a minimum of 10 percent of students who are classified as LEP for purposes of No Child Left Behind. In practice, the proportion of students who are ELLs (which include those reclassified as “English proficient” in elementary schools) is significantly higher than the proportion that is officially classified as LEP.<sup>2</sup>

Five considerations will drive our data collection activities in schools. First, as part of the REL West contract, the focus of the study is the Western region, which includes the states of California, Arizona, Nevada, and Utah. The sample will be located in California. Second, to estimate the net impact of QTEL it is important that schools participating in the study are not already part of an ongoing QTEL program or similar targeted professional development for teachers of ELL students. Third, we want to select schools that are broadly representative of the region and serve students who are similarly representative of the growing ELL student population in the region. Fourth, the schools and their teachers must have the potential to benefit from QTEL. This means that the schools must be sufficiently stable to allow teachers to fully participate in QTEL, including the coaching and collaborative lesson planning activities during the school year. Teacher assignments also need to be stable so that teachers who participate in a summer institute are likely to be at the school for the entire school year and likely to remain at the school for several consecutive years. The QTEL program is intensive and requires strong buy-in from school principals and district administrators. All of this means that districts and schools that have unstable leadership, may be subject to state takeover, or are in other ways unstable will not be included in the study. Lastly, by concentrating in a large but distinct geographic area with large numbers of ELL students, the study is able to collect data in sufficient numbers of schools while keeping the cost of implementation and research activities manageable.

We expect a 100 percent response rate for schools/districts to deliver the research team student achievement data. This estimate is reasonable since it is part of requirements for districts/schools to participate in the study. We expect an 85 percent response rate for teachers. This estimate is based on several factors that will be discussed in Section 3 below.

---

<sup>1</sup> <http://www.census.gov/popest/counties/CO-EST2005-08.html>, accessed on December 9, 2006.

<sup>2</sup> The San Diego County Office of Education estimates that at least 25 percent of middle school students in the county are ELLs.

Twenty-five of these schools will be randomly assigned to a treatment group, which will be eligible to participate in QTEL and twenty-five will be assigned to a control group, which will be excluded from QTEL for three years. The schools will be randomized within their school district, so that each district has at least one treatment school and one control school.<sup>3</sup> Such stratification by district also helps to minimize random variation in background characteristics between students and teachers in the treatment and control groups. All ESL and ELA teachers in treatment schools will participate in the program and all students in grades 6, 7, and 8 will be part of the study. It is expected that the overall study sample will have approximately 600 teachers and approximately 50,000 students, of whom an estimated 12,500 will be ELLs. These teachers and students will roughly be equally distributed among grades 6, 7, and 8.

The purpose of this sample design is to produce impact estimates that have sufficient statistical precision so that impacts that are practically meaningful will also be statistically significant. Middle school educational interventions are usually found to have impacts in the order of 0.30 to 0.60 standard deviations (Bloom & Lipsey, 2005), so it is reasonable to design our study to at least be able to detect impacts of that magnitude on student outcomes in key student subgroups. Most importantly, we will want to be able to distinguish the program's impacts on ELL students, assuming that those students make up 25 percent of a school's student body, on average.

Exhibit 5 shows the expected minimum detectable effect sizes (MDES) for our study design, which assumes a sample of 50 schools. We do not think that increasing the sample beyond 50 schools would be useful because, even in San Diego County, the number of middle schools with significant ELL populations is limited. Therefore, we would either have to add an additional county, which is very expensive, or we would have to sample down to schools with lower ELL populations, for which the intervention would be less relevant. Neither one of these options is acceptable.

To calculate student-level MDES, we are using empirical data reported by Bloom et al. (2005) to determine appropriate values for the intra-class correlation (ICC) and explanatory power of our impact regressions. The ICCs and  $R^2$  statistics reported by Bloom et al. appear to be correlated. In two middle school districts they find values of 0.17 and 0.77 for ICC and  $R^2$  in one district and 0.23 and 0.91 in the other. This makes intuitive sense. If student outcomes are more strongly clustered by school one might expect the year-over-year predictive power of student outcomes to be greater as well. Given these results, we believe an ICC 0.15-0.2 and an  $R^2$  of 0.65-0.70 to be reasonable. At the student level, we limit the analyses to a single grade of 320 students overall and 80 ELLs (25 percent). This is estimated average based on the available information.

For calculating teacher-level MDES we unfortunately have no reliable data on the explanatory power of teacher-level covariates in impact regressions and no reliable data on teacher-level ICCs. Thus, we used ICCs of 0.05 and 0.20 to bracket a likely empirical value and we use an R-squared of 0, which is as conservative as it can be.

As Exhibit 5 shows, these parameters produce student-level MDES between 0.19 and 0.22 for the full sample, between 0.20 and 0.23 for ELLs, and between 0.29 and 0.33 for ELLs in half the schools. Even if the cost of the intervention is amortized over many years, we believe that these MDES are sufficiently small for an ambitious and intensive intervention like QTEL.

Teacher-level MDES are substantial, ranging from 0.29 to 0.42 for the full sample of teachers, using our assumptions. Outcomes measured via classroom observations exhibit MDES estimates ranging from 0.49-0.55. However, we do not expect a teacher-professional development intervention like QTEL to produce impacts of 0.20 on students without at least a 0.40 or higher impact on teachers. Also, we believe that the student-level MDES should be the primary driver for the statistical power calculations since student-level outcomes are ultimately the primary focus of the intervention and the study.

---

<sup>3</sup> The exception is very small districts with only one middle school, of which there will likely be a few in the sample.

**Exhibit 5: Statistical Power Calculations**

	Schools	Teachers/ school	Students/ grade	Case 1	Case 2	
<b>MDES for student outcomes</b>						
Assumptions						
				ICC	0.15	0.20
				R2	0.65	0.65
MDES						
Full Sample	50		320	0.19	0.22	
ELL students	50		80	0.20	0.23	
ELL students in half the schools	25		80	0.29	0.33	
<b>MDES for teacher outcomes</b>						
Assumptions (no baseline data)						
				ICC	0.05	0.20
				R2	0.00	0.00
MDES						
Full Sample	50	12		0.29	0.42	
Classroom Observation Data	50	3		0.49	0.55	
Teachers in half the schools	25	12		0.42	0.60	
Subgroup of 50 percent of teachers	50	6		0.37	0.47	

**2. Information Collection Procedure**

Of the schools requesting the program, about half will be randomly selected to participate in the QTEL training, and the remaining half will be assigned to a control group. The random assignment will be conducted at BPA and a research staff member will notify each school of their group assignment.

Each school that volunteers to participate in this study will be asked to distribute information/assent forms to the students in their ESL and ELA classes. The teachers who agree to participate in the study will be asked to complete an annual teacher survey about their professional preparation, teaching experience and attitudes toward working with ELLs as well as an annual test of their pedagogical knowledge. These surveys/tests will be completed in a web-based format. Log-in information will be sent to teachers via their school email account. If teachers prefer to complete a paper version of the survey/test, the research team can send them a hardcopy (of the survey/test) and an addressed, stamped envelope in which to return the survey/test to the research team. This information will be given to the teachers prior to the survey/test administration (see an example of “Dear Teacher” letter as shown in Appendix C).

Each selected teacher will participate in a classroom observation in the spring of the program year (6<sup>th</sup> grade in 2008, 7<sup>th</sup> grade in 2009, 8<sup>th</sup> grade in 2010). Some classrooms may also be selected to participate in videotaped observations (in 10 treatment classrooms and 5 control classrooms). In addition, there will be data collection through observations and focus groups (with coaches or teachers) that are related to teachers’ professional development in QTEL program.

Student archived data will be annually collected through the participating districts/schools. These data include the California Standards Test (we plan to focus on all content areas, but especially on ELA), the California English Language Development Test (CELDT), measures of attendance, student grades (GPA), and grade promotion.

We will follow up with teachers who do not complete surveys/tests or confirm observation appointments in a timely manner. These follow-ups will be conducted bi-weekly first by email, then

by telephone, and last by letter. If needed, we will solicit the help of school principals in encouraging teachers to respond. We will also follow up (weekly) with the districts/schools until they deliver the required student data to the research team.

The data collection procedure and detailed timeline is also discussed under A2, "Purposes and Uses of the Data", as well as A16, "Data Collection Schedule."

### **3. Methods for Maximizing Response Rates**

In this study, the new data collection will mainly come from teachers (in particular, teacher survey and test of pedagogical knowledge). Several methods will be used to ensure an 85 percent response rate for collecting teacher surveys/tests:

- Prior to each survey/test administration, the research team will contact each school principal that encourages teacher participation in the surveys/tests.
- Prior to each survey/test administration, a "Dear Teacher" letter will be sent (via email or regular mail) to each participating teachers to inform the upcoming data collection activity.
- The surveys/tests will be conducted using a web application, allowing for ease of completion. (As mentioned earlier, a hardcopy of the survey/test will be made available upon request.)
- The research team will compensate teachers for their time to complete these data collection activities. This compensation plan is discussed under A9, "Payments to Respondents".
- Teachers who fail to complete the survey/test in a timely manner will be followed-up with every other week first by email, then by telephone, and last by letter.

No new data collection will take place from students. Only extant state-referenced achievement data and other administrative data will be collected from the districts/schools. To ensure the delivery of student data, a district level memorandum of understanding (MOU) will be sent to each district outlining the support they will receive for participating in the study, the roles and responsibilities of both research staff and district/school site staff, and estimates of the time required to collect data. A copy of MOU is listed in Appendix D.

We will be in frequent contact with teachers by email and telephone throughout the study to assess the stability of their participation. The support provided by the program implementation team is designed to meaningfully help teachers with their instructional program; we believe this has inherent benefits that will help to retain study participants. Specifically, there will be several focus group sessions (with coaches or teachers) planned throughout the study that will enable and facilitate ongoing communication with the teachers.

Although this study includes a plan to monitor and ensure implementation fidelity, it is possible that some participants assigned to the treatment group will not participate in all intervention activities. Nonparticipation by significant numbers of those targeted to receive the intervention would likely dilute potential program impacts. Extensive efforts will be made to collect data from such non-participants, and levels of participation in the intervention will be monitored through surveys and records. So as not to bias impact estimates, all such participants will be kept in the impact analysis in their original, assigned groups to avoid sample selection bias. That is, an intention-to-treat analysis (ITT) will be performed. ITT refers to the fact that random assignment only establishes an "intention to treat," but does not actually guarantee that those assigned to the program experience it.

## 4. Expert Review and Piloting of Instruments

All data collection instruments went through a series of reviews either by the research team or the TWG members to ensure that the instruments are reliable and valid. In addition, all new data collection instruments were tested with 9 or fewer teachers participating in WestEd's implementation of the QTEL program in San Jose, California, during the 2006-2007 school year. Because of the small sample size, item level statistical analysis was not conducted. However, changes were made to improve the clarity of survey/test items and increase the relevance of observation protocol items. Testing did not lead to significant changes in the length or content of instruments. The time burden estimate associated with each instrument was obtained through this piloting.

The SIOP© observation protocol (Pearson Education, Inc., 2004) will be used for classroom observations. This instrument has been well validated by the developers in a variety of settings with ELLs (Echevarria, Vogt and Short, 2004).

## 5. Statistical Consultants

Exhibit 6 below lists the name and telephone number of individuals consulted on statistical/data analytical aspects of the study.

**Exhibit 6: Statistical Consultants**

<b>Name</b>	<b>Affiliation</b>	<b>Contact Information</b>
Hans Bos, Ph.D.	Berkeley Policy Associates	(510) 465-7884
Neal Finkelstein, Ph.D	WestEd	(877) 938-3400
Lorena Ortiz Adams	Berkeley Policy Associates	(510) 465-7884

## References

- Allwright, D. (1988). *Observation in the language classroom*. London: Longman.
- Allwright, D., & Bailey, K. (1991). *Focus on the language classroom*. Cambridge, England: Cambridge University Press.
- Angrist, J. D., Imbens, G. W., & Rubin, D. B. (1996). Identification of causal effects using instrumental variables. *Journal of the American Statistical Association*, 91, 444-472.
- August, D., & Hakuta, K. (Eds.). (1997). *Improving schooling for language-minority children*. Washington, DC: National Academy Press.
- Bloom, H. S., Bos, J. M., & Lee, S. (1999). Using cluster random assignment to measure program impacts. *Evaluation Review*, 23(4), 445-469.
- Bloom, H. S., & Lipsey, M. W. (2005). *Project on Uses and Abuses of Effect Size Measures: Introduction to the Issues*. Powerpoint presentation to U.S. Department of Education, July 28. New York: MDRC.
- Bloom, H.S., Richburg-Hayes, L. and Rebeck Black, A. (2005). Using covariates to improve precision: Empirical guidance for studies that randomize schools to measure the impacts of educational interventions. MDRC Working Papers in Research Methodology, November, 2005.
- Bruner, J. (1983). *Child's talk: Learning to use language*. Oxford, England: Oxford University Press.
- Candlin, C., & Murphy, D. (Eds.). (1987). *Language learning tasks*. Englewood Cliffs, NJ: Prentice Hall.
- Echevarria, J., Vogt, M. and Short, D. (2004). *Making Content Comprehensible for English Learners: The SIOP Model*. Boston: Pearson Education Inc.
- Edwards, A. D. and Westgag, D. (1987). *Investigating classroom talk*. Philadelphia: Falmer.
- Fix, M., & Zimmermann, W. (1993). *Educating immigrant children: Chapter 1 in the changing city*. Washington, DC: Immigrant Policy Program, The Urban Institute.
- Gibbons, P. (2002). *Scaffolding language, scaffolding learning: Teaching second language learners in the mainstream classroom*. Portsmouth, NH: Heinemann.
- Gibbons, P. (2003). *Scaffolding academic language across the curriculum*. Presentation at the annual meeting of the American Association for Applied Linguistics, Arlington, VA.
- Hall, R. (2000). Videorecording as theory. In *Handbook of research design in mathematics and science education*. A. E. Kelley and R. A. Lesh, Eds. Mahwah, NJ: Lawrence Erlbaum.
- Hammond, J. (Ed.). (2001). *Scaffolding: Teaching and Learning in language and literacy education*. Sydney: Australian Primary English Teaching Association.
- Kumpulainen, K. and Mutanen, M. (2000). Mapping the dynamics of peer group interaction: A method of analysis of socially shared learning processes. In *Social interaction in learning*



- and instruction: *The meaning of discourse for the construction of knowledge*. H. Cowie, and G. M. van der Aalsvoort, Eds. Oxford: Pergamon.
- Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge, England: Cambridge University Press.
- Lemke, J. L. (1990). *Talking science: Language, learning and values*. Norwood, NJ: Ablex Publishing.
- McGonigal, J. (1997). Using oral discourse in literacy studies. In B. Davies & D. Corson (Eds.), *Encyclopedia of language and education: Vol. 3. Oral discourse and education*. Dordrecht, the Netherlands: Kluwer Academic Publishers.
- National Center for Education Statistics (NCES). (2002). *Schools and staffing survey, 1999-2000: Overview of the data for public, private, public charter, and Bureau of Indian Affairs elementary and secondary schools*. Retrieved May 1, 2004, from <http://nces.ed.gov/pubs2002/2002313.pdf>
- 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 from [http://nces.ed.gov/programs/quarterly/Vol\\_3/3/q3-3.asp](http://nces.ed.gov/programs/quarterly/Vol_3/3/q3-3.asp)
- Resnick, L. B., & Nelson-Le Gall, S. (1997). Socializing intelligence. In L. Smith, J. Dockrell, & P. Tomlinson (Eds.), *Piaget, Vygotsky, and beyond* (pp. 145-158). London: Routledge.
- Rivera, H., Tharp, R. G., Youpa, D., Dalton, S., Guardino, G. and Laskey, S. (1999). ASOS: Activity Setting Observation System coding and rulebook. Santa Cruz: Center for Research on Education, Diversity & Excellence, University of California.
- Rivera, H. and Tharp, R.G. (2004). Sociocultural activity settings in the classroom: A study of a classroom observation system. In *Observational Research in U.S. Classrooms*, H.C. Waxman, R. G. Tharp and R.S. Hilberg (Eds.). New York: Cambridge University Press.
- Rogoff, B. (2003). *The cultural nature of human development*. New York: Oxford University Press.
- Ruiz-de-Velasco, J., & Fix, M. (2000). *Overlooked and underserved: Immigrant students in U.S. secondary schools*. Washington, DC: The Urban Institute.
- Seedhouse, P. (2004). *The interactional architecture of the language classroom: A conversation analysis perspective*. Malden, MA: Blackwell.
- Shulman, L. S. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review*, 57(1), 313-333.
- Shulman, L. S. (1995). Fostering a community of teachers and learners. Unpublished progress report to the Mellon Foundation.
- Shulman, L.S. and Shulman, J. H. (2004). How and what teachers learn: a shifting perspective. *Journal of Curriculum Studies*, 36 (2), 257 – 271.
- Van der Aalsvoort, G. M. and Harinck, F. (2000). Studying social interaction in instruction and learning: Methodological approaches and problems. In *Social interaction in learning and instruction: The meaning of discourse for the construction of knowledge*. H. Cowie, and G. M. van der Aalsvoort, (Eds.). Oxford: Pergamon.

- van Lier, L. (1988). *The classroom and the language learner*. London: Longman.
- van Lier, L. (1996). *Interaction in the language curriculum: Awareness, autonomy and authenticity*. London: Longman.
- van Lier, L. (2004). *The ecology and semiotics of language learning*. Dordrecht, the Netherlands: Kluwer Academic.
- 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: Delta Systems for the Center of Applied Linguistics.