Evaluation of Professional Development for Language Teachers in California

Albert S. Lozano  
Stanford University

Hyeyoung Sung  
Stanford University

Amado M. Padiña  
Stanford University

Duarte M. Silva  
Stanford University

Abstract: The California Foreign Language Project (CFLP), established in 1988 by the California legislature, is a professional development program designed to improve and expand foreign language teaching in California while promoting access and equity within educational institutions for every student. As one of the nine consent areas of the California Subject Matter Project, CFLP is a voluntary project that served 600 elementary, secondary, and postsecondary foreign language teachers from 43 counties in 1998/1999. This article describes the origin and rationale of the California Foreign Language Project and discusses the importance of professional development programs, a topic of growing interest given the nationwide focus on student performance and school reform. Finally, the components of professional development program evaluation, and specifically of CFLP Evaluation Design, will be presented, along with the findings from the 1998/1999 program year.

Introduction

The importance of increasing students' foreign language abilities has been discussed in several influential documents on the condition of American education. Reports such as The President's Commission on Foreign Language and International Studies (1979), A Nation at Risk: The imperative for Educational Reform (1983), and Standards for Foreign Language Learning in the 21st Century (1990) have argued that American public schools should expand their focus on foreign language education. Former Secretary of Education Richard Riley stated that increasing foreign language instruction would help build a better workforce, ensure national security, and improve other areas of education (Puzehol et al., 2020). The California legislature recognized this need and in 1988 passed Senate Bill 1882, establishing the California Foreign Language Project (CFLP) as one of the nine consent areas of professional development pro-

Albert S. Lozano (PhD candidate, Stanford University) is a graduate student in the School of Education, Stanford University, Stanford, California. Hyeyoung Sung (PhD, Stanford University) is a Research Associate in the School of Education, Stanford University. Amado M. Padiña (PhD, University of New Mexico) is Professor of Psychological Studies in Education, Stanford University. Duarte M. Silva (EdD, University of San Francisco) is the Executive Director of the California Foreign Language Project at Stanford University.
programs (the California Subject Matter Projects [CSMP]). The CFEP is a professional development program for language teachers, supports efforts to involve all language teachers in a supportive professional community that respects diverse identities; provides opportunities for leadership, promotes linguistic and cultural competence, and advocates the attention, expansion, and articulation of foreign and indigenous language offerings across educational levels beginning in the elementary grades. CFEP also strives to promote access and equity within educational institutions for every student.

In this report, we describe the origin and rationale of the California Foreign Language Project. We also discuss the importance of professional development programs, a topic that has received increased attention due to the widespread focus on student performance and school reform. Finally, the components of professional development, program evaluation, and specificity of CFEP's Evaluation Design, will be presented, along with the findings from the 1998/1999 program year.

California Foreign Language Project

During the 1980s, numerous reports described the poor performance of students in America's schools (e.g., A Nation at Risk, 1983). In response to these reports, the California legislature established several councils to design professional development programs. One of these programs, the Foreign Language Curriculum Implementation Centers (FLCICs), is the educational foundation for CFEP's current vision. The FLCIC program is focused on assisting language teachers to reorient their language programs and lessons towards proficiency-oriented instructional, to strengthen their own abilities to implement communication-based instruction, and to enable them to assist their colleagues in the implementation of communicative learning strategies. (Hiro, et al., 1994, p. 17)

The motivation behind the FLCIC project was to change the way that foreign languages were taught, that is, to shift the orientation from the monolingual method to communication-based instruction (CBI), which California's foreign language framework defines as teaching that helps students to develop the skills necessary to produce and receive understandable messages. "is found in programs that emphasize an authentic exchange of meanings in the foreign language and the language being learned" (California State Department of Education, 1989, p. 7). CFEP, the project created in 1986 that originated from the former FLCIC, also ensures that participants are familiar with communication-based instruction and focuses on the Standards for Foreign Language Learning, a document that guides CFEP's efforts in professional development for language teachers.

CFEP is a professional development program designed to improve and expand foreign language teaching in California while promoting access and equity within educational institutions for every student. As one of the nine content areas of the CSMP, CFEP has evolved into a voluntary project, consisting of a central office and nine regional sites that organize workshops, seminars, and other professional opportunities. Moreover, legislation passed in 1998 (Senate Bill #1734) mandates that all CSMPs, including CFEP, focus more money and efforts on schools designated as "low-performing." Thus, the focus of providing foreign language teachers with pedagogical knowledge based on partnering students' communicative skills, along with the increased focus on low-performing schools, forces CFEP's professional development goals of:

1. strengthening academic content knowledge
2. developing teacher leadership
3. serving low-performing schools and districts
4. establishing partnerships with low-performing schools
5. establishing and maintaining a professional community and teacher networks
6. evaluating student learning

The Importance of Professional Development Programs

Professional development programs are currently receiving more emphasis from both educational policy makers and researchers, and much of this emphasis arises from the importance given to comprehensive reform and school improvement. As Guskey (2000, p. 16) states, "High-quality professional development is at the center of every modern proposal to enhance education. Regardless of how schools are formed, reformed, structured, or restructuring, the renewal of staff members' professional skills is considered foundational to improvement." Programs related to the implementation of the Standards have been document in subject areas such as foreign language (Kern & Lively, 1995; Kern & Morin, 1996), mathematics (Svec, 1997; Stead, 1990), and science (National Academy of Sciences, 1996; Hawley and Valla (1999) state that professional development programs are a vital component of successful educational reform. As the American Federation of Teachers notes:

"Without professional development school reform will not happen... The nation can adapt rigorous standards, set forth a visionary scenario, compile the best research about how students learn, change the nature of textbooks and assessment of students, and change all the other elements involved in systemic reform. But... unless the classroom teachers understand and are committed to the plan and knows how to make it happen, the dream will come to naught. (cited in Howley and Valla, 1999, p. 129)"
The Evaluation of Professional Development Programs

Although the availability of literature on professional development exists, "researchers have tried unsuccessfully to determine the true impact of professional development in education" (Guskey, 2000, p. 32). Guskey argues that evaluations of professional development programs are often erroneous due to confusion about the criteria of effectiveness, a misguiding search for main effects between programs, and a neglect of issues concerning the quality of components. Nevertheless, Guskey states that sound professional development evaluations include five components: (1) participants' reactions to a workshop, program, or series of events; (2) participants' learning of new skills, knowledge, and/or attitudes; (3) information on organization and support change; (4) participants' use of new knowledge and skills; and (5) student learning outcomes. CFLPs' Evaluation Design incorporates several of these components to obtain a better picture of the program's effectiveness.

CFLP's Evaluation Design

To evaluate how well each regional site implements professional development services that fulfill CFLPs' six goals, an Evaluation Design was established for each site so that it could collect information about its professional development activities. Sites are required to construct an Evaluation Team to assist in the CFLP evaluation effort and establish a plan of events/activities for the upcoming year. Each site develops the implementation of their yearly plan by providing the CFLP central office with a site portfolio. The portfolio has three main sections: (1) site performance, (2) teacher performance, and (3) student performance.

In the first session of the portfolio, site performance, CFLP requires that sites provide program agendas, copies of handouts, materials used during program activities, and samples of written or technology-based materials created by participants during the programs offered. Sites are also instructed to submit videotaped samples that are representative of their program activities. At all CFLP professional development programs, participants are asked to fill out program evaluation surveys to assess how well each workshop program was delivered and how much each participant learned through their participation. Sites are asked to submit a summary of the participants' evaluation results from each workshop. Finally, sites are to include a description of their efforts in collaboration with low-performing schools.

In the teacher performance section of the portfolio, regional sites are asked to provide: (1) sample lesson plans created by teachers who participated in the programs, and (2) videotapes and accompanying handouts (if any) of participants teaching a lesson in their own classrooms. The CFLP Evaluation Team developed a protocol to examine how well the lesson plans and videotaped classrooms teaching samples reflect aspects of the proficiency-oriented language instruction, one of the goals of CFLP professional development programs.

The student learning section of the portfolio focuses on how the teaching practices of participating teachers influence student outcomes, and sites are requested to provide oral and written samples of student work. The sites assess students' oral proficiency using the Classroom Oral Competency Interview (COCI) (California Foreign Language Project, 1993) and the Standard Foreign Language Oral Skills Evaluation Matrix (FLOSEM) (Padilla & Sung, 1999) and student writing proficiency by means of the Classroom Writing Competency Assessment (CWCA) (California Foreign Language Project, 1996).

Sites are required to submit portfolios demonstrating evidence of meeting CFLPs' goals but are not mandated to use certain prescribed methods in their teacher development programs (e.g., a set number of workshops). That is, each site adapts to professional development series to more effectively meet the needs of its respective constituents. For example, many sites have focused attention on new foreign language teachers at the secondary level, whereas others have maintained a strong relationship with postsecondary foreign language educators. Sites also differ with regard to the total number of participants, number of languages represented, and geographic regions served.

CFLP's current Evaluation Design has been modified from previous years to encompass evaluative elements on multiple levels. In the past, CFLP relied on participants' workshop evaluations, briefings conducted between site-affiliated personnel and CFLP members, and observations made at site-sponsored workshops. However, the updated Evaluation Design calls for each site to assemble a Site Portfolio that best captures program activities and their effects on teacher and student performance throughout the program year. Regional site evaluation teams are instructed on how to organize the site portfolio to include specific evidence that demonstrates each site's effectiveness in implementing its approved professional development program. Site portfolios are collected at the end of the program year and then analyzed by the CFLP evaluation staff. Each site receives a written feedback report based on the analysis of data from the site portfolio, including recommendations for improving program activities and data collection for the following year. The report is given to the site director, evaluation liaison personnel, and other members of the site's leadership team.

Findings

Site Performance

Professional development workshop activities. During
the 1998/1999 program year, the CFIP network statewide offered over 170 days of professional development program sessions for more than 500 language educators. Most of these programs focused on: (1) strategies for implementing proficiency-oriented language instruction, (2) assessment protocols, (3) lesson planning and design, (4) standards-based language education, (5) incorporating technology in language instruction, and (6) teaching language to heritage learners. In addition, regional sites held numerous leadership teacher meets to plan and evaluate their programs.

According to information from the site portfolios, the most prevalent topic addressed by the CFIP sites was the Standards for Foreign Language Learning, which targets the areas of Communication, Culture, Connections, Comparisons, and Communities (the "five Cs"). All nine sites offered participants information on the Standards either by creating workshops that specifically addressed them and/or by using the Standards as the "organizing principle" of their workshop series. In fact, one site offered sessions pertaining to the Standards in all three of their workshops, and participants' comments indicate that information about the Standards was well presented and informative. Another prevalent topic was information and instruction on using several student evaluation instruments endorsed by CFIP. More specifically, seven of the nine sites presented participants with information and instruction on at least one of three methods of student assessment: (1) COI, (2) CWCA, and (3) FOLSEM.

Even though each site adapts programs and workshops to the needs of its participants, all nine sites reported that participants were pleased with the workshop presentations. Moreover, participants felt that the cumulative effect of the program activities enhanced their theoretical and pedagogical knowledge in foreign language education. As one participant wrote, "After two years, I fully understand and (am) able to apply the five-corp lesson plans, literacy lessons, COI, CWCA, and Standards. I will always enjoy going over my folders. I look forward to preparing more lessons."

Several sites created workshops and group activities in the standards based on the needs of their specific cohorts. Two sites offered workshops on the Standards that were designed specifically for new teachers, while other sites designed workshops/seminars that provided leadership team members with additional information and instruction, knowledge that is critical for team leaders, to encourage them to assist less-experienced foreign language teachers. As one site leader stated, "it will be exciting the teachers in San Francisco, and while I don't feel like an expert, you workshop has given me a degree of confidence I lacked before." Several sites have also incorporated workshops that assist language educators who teach Spanish for Spanish Speakers and the Less Commonly Taught Languages (e.g., Chinese, Korean).

Collaborative efforts with partnership schools. The diversity of a relationship in which two supports types of collaborative efforts that were created between low-performing schools and CFIP affiliates. What is most evident is that in partnerships that empowered teachers from low-performing schools to restructure the curriculum. For example, one site established a partnership with a middle school (third through) and its articulated high school (seven teachers) in which teacher in collaboration with CFIP staff were able to write a vision statement of their goals and objectives. CFIP also provided those teachers with workshops pertaining to the Standards and communication-based pedagogy. Another site established partnerships with teachers from one middle and two different high schools. In this middle school, teachers expressed a desire to meet with high school teachers in order to provide better student stimulation, and this site was instrumental in arranging this meeting. Moreover, several teachers from this same middle school participated in the site's CFIP-sponsored workshops and were sponsored with stipends to attend a California Language Teachers Association (CLTA) conference. This site also supported teachers from the middle school in writing a Middle School Demonstration Program grant and held a workshop on the Standards that included teachers from all three partnership schools. While several sites had an impact on low-performing schools at the staff level, another site successfully implemented a relationship in which two participating sites of a low-performing school were provided mentoring from a CFIP-affiliate staff member, that individualized endow-er that also produced positive results. As one of the partic- ipants wrote, "I really understand, I saw the peer coaching experience as a resource to ask questions, a piece to feel vulnerable as well as nurtured, and a part of the back for my hard work and dedication to the process of learning."

Even though the above examples demonstrate that the collaborative partnerships have been effective, other sites have had difficulty in relating their services to the needs of some low-performing schools. The problems experienced by some sites are due to a disconnect between the types of services offered and the types of services teachers from low-performing schools felt would improve their performance. For example, one site formed a collaborative partnership with the site of the tenth school of a secondary school district and offered services that included: (1) training in Standards-based lessons with experienced teachers, (2) opportunities for teachers in the low-performing schools to observe experienced teachers delivering Standards-based lessons, and (3) peer coaching. Unfortunately, these ser- vices were not fully utilized by teachers at the night low-performing schools because, "very few teachers were able to take advantage of these services extended to them."
Another CLEP site established a partnership with one low-performing school; however, "teachers were invited to attend the standard workshops on assessment instruments, but they failed to attend." In these instances, sites did not receive input from low-performing schools and developed services that were not matched to the needs of these low-performing school teachers.

Finally, one site was able to secure partnerships with three low-performing high schools and offered these workshops focused on themes researched by representatives of the partners' schools. Evaluation summaries from both CLEP staff and low-performing school teachers indicate that the participants may not have met the participants' needs. Since then, the site has responded by working collaboratively with partners to redesign their workshops so that they better address the needs of partnership-school teachers.

**Teacher Performance**

To determine how the program activities influence participating language teachers in their classrooms, each site was asked to provide sample lesson plans of participating teachers and videotapes of participating teachers' classroom instruction. Five hundred twenty-five (85.7%) participants were classroom teachers, and of those who indicated their teaching level and language taught, the majority taught at the high school level (77%) and taught Spanish (77%).

Lesson plan samples. In the portfolio, all nine regional sites provided lesson plans that incorporated workshop materials. Several CLEP workshops focused on providing participants with instruction on creating five-step lesson plans, with two elements (guided practice and comprehensible input) receiving the most emphasis. These two elements were generally the strongest components in participants' lesson plans, an indication to CLEP sites that participants were implementing workshop materials. In addition, the study found that the CLEP's thematic organization was effective in organizing materials for integration into existing curriculum. Participants were able to integrate the materials into their classrooms.

**Student Performance**

Oral proficiency improvement measured by the Stanford FLOSEM. Each regional site was asked to select at least two schools (one low-performing school and one other school) and to collect students' oral proficiency data by means of the Stanford FLOSEM. Proficiency ratings were collected from at least one classroom for each level of the language taught at the school twice per year (once after the first month of instruction and again at the end of the school year). Although not every site was able to provide all of the required data, many sites were successful in collecting the TLOQSM data from various language programs with different instructional levels.
The Stanford FLOSEM assigns oral proficiency scores from 5 (earliest beginning level) to 30 (native-like proficiency). Oral proficiency scores were collected from Spanish, French, German, Italian, Japanese, Arabic, and ESL (English as a Second Language) classes and from many different instructional levels. A total of 1,160 scores were collected in the fall semester and 1,128 scores in the spring, and of these, two matched rating scores were available for 983 students. A paired t-test was performed for these 983 students' FLOSEM scores to determine if students' oral proficiency increased from fall to second ratings. The mean score of the second FLOSEM ratings (M = 12.94, SD = 6.37) was significantly higher than that of the first ratings (M = 10.63, SD = 6.55). The mean difference was statistically significant; t = 6.02, p < .001. This finding indicates that students made significant progress across all language levels in their oral proficiency during the school year.

Oral and writing proficiency using the COCI/CWCA. Each regional site collected oral and writing proficiency data from students in advanced foreign language classes (instructional levels 3 and 4 or higher) by means of the COCI and the CWCA. The sites were asked to submit 20 student samples of the COCI and CWCA from teachers in every language they served. Brief descriptions of the assessment instruments are as follows.

Learning in the Formulasic Range use unstated chunks of language (words, phrases, some sentences). Learners in the Formulasic Range recall and use chunks and recombine them to create their own sentences. Finally, learners in the Formulasic Range organize created utterance into paragraphs, expressing more complex meaning. With each range, there are three levels. In the Low level, learners show minimal ability to sustain performance with the language type. The quantity and quality of the language type increase in the mid level. In the High level, learners show unsaturated ability to perform at the next range.

The majority of the sites succeeded in collecting the data to meet this requirement. A total of 357 student COCI oral interview samples and 446 CWCA writing samples were collected. The regional sites were able to collect both oral and writing proficiency samples from only 272 students. The majority of student proficiency data was collected from other Spanish or French classes. Of the 357 oral samples, 156 (43.7%) were in Spanish and 115 (32.2%) were in French. The other language data included Japanese (8.1%), ESL (9.9%), and German (3.0%) samples. Regardless of the writing proficiency data, 170 writing samples (38.1%) out of 446 were in Spanish and 132 samples (29.6%) were in French. There were also German (15.5%), Japanese (7.4%), and ESL (14.4%) samples.

Figure 1 presents the number of students whose COCI and CWCA ratings fall into each range (Formulasic, Creativis, or Manipulat and level (Low, Mid, or High). Students' oral and writing proficiency varied from Formulasic to Planned High. However, many students were in the Created range and attained an average of Created Low/Mid for their oral and writing proficiency. According to the COCI and CWCA Manuals, students in advanced language classes (level 3 or higher) are expected to have proficiency ratings in the Created Low to Mid range. Of the 357 students, 239 (67.4%) achieved or exceeded the expected oral proficiency rating of Created Low or Mid, and 293 out of 446 students (66%) achieved at the expected level or higher in their second language writing proficiency.

The analysis of the COCI and CWCA results indicates that there was a significantly high correlation between oral proficiency and writing proficiency in a second language. \( r = .579, p < .001 \). For example, among the 272 students who were assessed on both the COCI and CWCA instruments, 72 students (26.7%) received equivalent ratings from both assessments, while 162 students (59.6%) had
COCl and CWCA ratings that differed by one or two levels (e.g., Created Low on the COCl with a CWCA rating of Created Med or High). An examination of the correlation between COCl and CWCA by different languages indicated that there was a significantly high correlation in most of the language programs: $r = .946, p < .0001$ for Spanish ratings with 117 students; $r = .560, p < .0001$ for French with 115 students; $r = .752, p < .0001$ for Japanese with 28 students; and $r = .994, p < .0001$ for German ratings with 13 students.

Writing proficiency ratings ($M = 4.43, SD = 1.79$) were, in general, significantly higher than the oral proficiency ratings ($M = 4.21, SD = 1.99$, $t(271) = -2.12, p < .05$). However, when the analyses were done separately for the different language programs, this significant difference was only found in the French language program. Students' writing proficiency in French ($M = 4.63, SD = 1.66$) was significantly higher than their oral proficiency ($M = 4.01, SD = 1.70$, $t(85) = -1.63, p < .0001$). In the other language programs (except ESL), the students' writing proficiency was consistently higher than their oral proficiency but the difference was not statistically significant. On the other hand, the ESL program students' oral proficiency ($M = 4.11, SD = 1.45$) was higher than their writing proficiency ($M = 3.89$, $t(62) = 5.10, p < .0001$).
oral and writing proficiency ratings on the COC and CWCA by percent of students falling into the various proficiency categories

![Graph showing oral and writing proficiency ratings on the COC and CWCA by percent of students falling into the various proficiency categories.]

25%
20%
15%
10%
5%
0%
Low
Mid
High
Low
Mid
High
Low
Mid
High
Portraits

Writing Proficiency Ratings on the CWCA

30%
25%
20%
15%
10%
5%
Low
Mid
High
Low
Mid
High
Low
Mid
High
Portraits

The difference between the low-performing and high-performing schools in terms of the students' oral and writing proficiency was examined. The results showed that student proficiency at the high-performing schools was significantly higher than at the low-performing schools (see Table 1).

Separate analyses by different language programs and different instructional levels demonstrated that the significant differences in oral and writing proficiency between the high- and low-performing schools were confined to a few advanced levels of instruction, as noted in Table 2.

Overall, an analysis of student writing proficiency data revealed a surprising result. In writing proficiency studies, from the low-performing schools significantly more students from the high-performing schools in Spanish 3 and 4 as well as in ESL-2 (see Table 3).

Discussion

The evaluation results indicate that regional site programs met the majority of CEP goals and objectives. For example, the Standards for Foreign Language Learning was used as an organizing principle to identify every site, indicating that CAEL's efforts are helping educate participants on proficiency-oriented instruction. Moreover, sites are required to have participants evaluate each workshop (participants' feedback) included in the portfolio:

1. summaries of the survey information (both quantitative and qualitative), and
2. the site's responses to participants' needs and concerns. Verbs from the participants indicated that overall, the workshops were positive endeavors.

One participant wrote, "I will never serve my students so that they may be successful in learning language."

Sots also asked to assess participants' learning, with instruments such as a Performance-Based Instructional Techniques survey developed by the regional site. The survey, which served as an evaluative measure and learning tool, asked participants to rate their implementation of CEP pedagogical approaches in each of the five lesson plan stages.

In addition, language-specific group workshops were utilized to help enhance teachers' language proficiency. Site portfolios also revealed that not all goals were met to the satisfaction of the teachers or CEP technical staff. For instance, teacher development efforts were not provided during the 1990/1991 program year. This is due to the fact that many sites secured much more on incorporating the Standards in their program activities, assessing students' oral and writing proficiency and for the first time on establishing partnerships with low-performing schools. In effect, not all the sites addressed teacher leadership development to the same extent in their yearly activities and site portfolios. Although sites have some flexibility as
to how very approach teacher leadership development. All sites are expected to schedule a focus on high performing schools. Finally, site portfolios indicated that service to low-performing school districts required a challenge, as was including more teacher participants from these schools.

The success of CFLPs professional development programs is most pronounced when analyzing the final two goals of the evaluation design: (1) teacher performance, which addresses participants' use of new knowledge and skills in their teaching, and (2) student performance, which focuses on student learning outcomes. Lesson plans that were submitted from each site showed how well participants were incorporating what they had learned through participation in the site workshops. Despite the fact that not every site submitted the requested number of participants (2) teaching in their classrooms, the videotapes (and accompanying lesson plans) that were submitted indicated that participants were able to express new knowledge to their lesson plans and to incorporate this knowledge into classroom teaching. Results also indicated that CFLPs effects have influenced students' foreign language abilities. On all three measures, students' language proficiency increased during the school year.

Our results, along with information about teacher development, do not offer "proof" of CFLPs effectiveness, but do provide evidence that the program is having a positive impact on participating teachers and their students. Finally, while CFLP anticipates that student learning outcomes will remain fairly consistent in subsequent years, data from the 1998/1999 program year provided information on the effectiveness of professional development efforts and indicators of areas that need improvement, such as gathering more teacher performance data.

An analysis of CFLPs professional development program indicates that the organization communicates consistent goals and objectives to its sites, while allowing the sites to address these goals as they see fit. CFLPs utilization of a site portfolio provides sites with the opportunity to describe their workshops and partnerships, giving both sites and CFLP evaluators the opportunity to work constructively and cooperatively to improve foreign language education in California. CFLPs evaluation design coincides with current research on the evaluation of professional development, with argues that evaluation should occur at different levels, including student learning (Guskey, 2000). However, what is most promising about CFLPs Evaluation Design is that it allows regional sites and CFLP evaluators to both plan and refine programs and activities to best serve participating teachers, most notably to serve to low-performing schools. In the past, CFLP has had more experience in working with teachers and administrators from high-performing schools. The increased focus of service to low-performing schools, however, has meant that CFLP and site affiliates have had to learn how to reach our recruit, and offer relevant professional development programs to teachers and administrators in low-performing schools. To accomplish this, each site is permitted to coordinate partnerships with low-performing schools and teachers in whatever manner works best, depending on local circumstances.

Finally, while some details, such as the program goals and objectives, are specific to CFLP, developers of other professional development programs may use the model presented here to guide assessment of their own programs' effectiveness. CFLPs Evaluation Design can easily be generalized and/or modified to fit a specific context, especially when evaluating a program's impact on multiple levels; that is, the effect on participants in their classrooms, along with the impact on student learning. Moreover, teachers can anticipate some of the difficulties that may arise and adjust their evaluation designs accordingly.

The need to increase students' foreign language abilities is not a novel idea. More than 20 years ago, educational scholars wrote that the "study of a foreign language introduces students to non-English speaking cultures, heightens awareness and comprehension of one's native tongue, and serves the Nation's needs in commerce, diplomacy, defense, and education" (National Commission on Excellence in Education, 1983, p. 20). The current focus on providing foreign language teachers with the pedagogical tools necessary to produce students capable of enhanced "communication skills in a foreign language is especially important today, given America's presence in the global economy. "For American students, the ability to function competently in at least one language other than English will become increasingly important in the rapidly shrinking, multilingual world of the twenty-first century" (National Standards in Foreign, Language Education Project, 1999, p. 39). The success of a professional development program lies in its ability to implement new skills and knowledge and students benefiting positively from this process. By this principle, CFLP serves to create an evaluation design that incorporates reflection and modification, while best addressing educational needs.

Notes
1. The California Department of Education identifies a school as "low-performing" if scores on the Stanford Achievement Test-Version 9 (SAT-9) are in the bottom 40%. CFLP uses a school SAT-9 scores in English/Language Arts.
2. The protocol can be obtained through the California Foreign Language Project.
Table 1
SIGNIFICANT DIFFERENCES IN STUDENTS’ ORAL/WRITING PROFICIENCY BETWEEN HIGH- AND LOW-PERFORMING SCHOOLS

<table>
<thead>
<tr>
<th></th>
<th>Number</th>
<th>Mean</th>
<th>SD</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oral Proficiency</td>
<td>High-performing</td>
<td>173</td>
<td>4.82</td>
<td>1.56</td>
</tr>
<tr>
<td></td>
<td>Low-performing</td>
<td>184</td>
<td>3.72</td>
<td>2.10</td>
</tr>
<tr>
<td>Writing Proficiency</td>
<td>High-performing</td>
<td>202</td>
<td>4.66</td>
<td>1.86</td>
</tr>
<tr>
<td></td>
<td>Low-performing</td>
<td>244</td>
<td>4.08</td>
<td>1.97</td>
</tr>
</tbody>
</table>

*Significance level: \( p < .0001 \)

Table 2
SIGNIFICANT DIFFERENCES IN STUDENTS’ ORAL/WRITING PROFICIENCY BETWEEN HIGH- AND LOW-PERFORMING SCHOOLS BY PROGRAM LEVEL AND LANGUAGE

<table>
<thead>
<tr>
<th></th>
<th>Number</th>
<th>Mean</th>
<th>SD</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oral Proficiency</td>
<td>Spanish 3</td>
<td>High-performing</td>
<td>27</td>
<td>5.00</td>
</tr>
<tr>
<td></td>
<td>Low-performing</td>
<td>32</td>
<td>3.59</td>
<td>1.29</td>
</tr>
<tr>
<td></td>
<td>French 3</td>
<td>High-performing</td>
<td>25</td>
<td>4.36</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Low-performing</td>
<td>17</td>
<td>3.35</td>
</tr>
<tr>
<td></td>
<td>French 4</td>
<td>High-performing</td>
<td>26</td>
<td>4.92</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Low-performing</td>
<td>34</td>
<td>4.38</td>
</tr>
<tr>
<td></td>
<td>ESL 2</td>
<td>High-performing</td>
<td>10</td>
<td>4.50</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Low-performing</td>
<td>8</td>
<td>3.50</td>
</tr>
<tr>
<td>Writing Proficiency</td>
<td>German 3</td>
<td>High-performing</td>
<td>27</td>
<td>4.07</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Low-performing</td>
<td>32</td>
<td>3.16</td>
</tr>
<tr>
<td></td>
<td>German 4</td>
<td>High-performing</td>
<td>16</td>
<td>5.19</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Low-performing</td>
<td>36</td>
<td>4.67</td>
</tr>
</tbody>
</table>

*Significance level: \( p < .0001 \)

Table 3
INSTANCES OF LOW-PERFORMING SCHOOLS OUTSCORING HIGH-PERFORMING SCHOOLS IN WRITING PROFICIENCY

<table>
<thead>
<tr>
<th></th>
<th>Number</th>
<th>Mean</th>
<th>SD</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spanish 3</td>
<td>High-performing</td>
<td>30</td>
<td>4.00</td>
<td>1.74</td>
</tr>
<tr>
<td></td>
<td>Low-performing</td>
<td>32</td>
<td>4.77</td>
<td>1.43</td>
</tr>
<tr>
<td>Spanish 4</td>
<td>High-performing</td>
<td>56</td>
<td>4.75</td>
<td>1.74</td>
</tr>
<tr>
<td></td>
<td>Low-performing</td>
<td>23</td>
<td>5.20</td>
<td>1.04</td>
</tr>
<tr>
<td>ESL 2</td>
<td>High-performing</td>
<td>9</td>
<td>3.44</td>
<td>1.67</td>
</tr>
<tr>
<td></td>
<td>Low-performing</td>
<td>8</td>
<td>4.38</td>
<td>0.92</td>
</tr>
</tbody>
</table>

*Significance level: \( p < .0001 \)