Educational Evaluation and Policy Analysis
June 2013, Vol. 35, No. 2, pp. 139–156
DOI: 10.3102/0162373712461849

The Gap That Can't Go Away: The Catch-22 of Reclassification in Monitoring the Progress of English Learners

William M. Saunders
Talking Teaching Network and UCLA

David J. Marcelletti
Talking Teaching Network

When English Learners (ELs) demonstrate English language proficiency, they are reclassified as Fluent English Proficient (RFEP). Subsequently they are often left out of the analysis of EL progress because they are, technically, no longer ELs. This article examines the effects of including and excluding RFEPs from the analysis of EL progress. Based on statewide achievement data from California including ELs, RFEPs, IELs (all initially identified English Learners: ELs + RFEPs), and English-only students (EOs), the analysis demonstrates that focusing on current ELs and excluding RFEPs (a) underestimates the population of IELs, (b) overestimates the achievement gap between IELs and EOs, and (c) decreases the likelihood of detecting progress when positive changes in achievement have taken place over time. Implications are discussed.

Keywords: *English language learners, reclassification, accountability, academic achievement*

English Learners (ELs) are students whose families speak a language other than English and, based on language proficiency assessments, are limited English proficient. Over the course of their schooling, when ELs demonstrate English language proficiency in accordance with criteria established by their state and district, they are reclassified as fluent English proficient (RFEP). At any particular grade level and across grade levels, evaluating the progress of "English Learners" might include those that remain ELs (current ELs), reclassified ELs (reclassified Fluent English Proficient, RFEPs), and the combination of the two, which includes all students initially classified as ELs (IELs = ELs + RFEPs). This article analyzes and discusses the importance of evaluating the progress of all three groups and illustrates a simple but often unrecognized Catch-22.[1]

Among *all* the students who are *initially* classified as ELs (IELs), those who are most successful—those who develop and demonstrate proficiency in English and are reclassified (RFEPs)—typically do not factor into evaluations of English Learner progress. RFEPs typically do not factor into evaluations of English Learner progress because their reclassification—their success—makes them no longer an EL. That is the Catch-22: Those that succeed—RFEPs—are typically excluded from the analysis

We wish to thank Tom Johnstone and Chris Jones for their early help in framing the problem; Karen Givvin, Brad Ermeling, Claude Goldenberg, Shelly Spiegel-Coleman, Kenji Hakuta, Robert Linquanti, and Leslie Laine for their feedback on initial drafts; and the journal reviewers who provided extensive, productive comments and suggestions.

of progress (Grissom, 2004; Parrish et al., 2006).

For example, in most states, ELs who reclassify are included in the No Child Left Behind (NCLB) EL subgroup for at most 2 years, after which their success has no bearing whatsoever on the EL accountability criteria their school, district, and state must meet. Unlike other NCLB subgroups formed on the basis of ethnicity or race, the EL subgroup does not remain stable over time: Successful ELs reclassify and move out of the EL subgroup (Abedi, 2004; Ramsey & O'Day, 2010). Schools and districts are under significant pressure to demonstrate progress with ELs. If they actually produce that progress and effectively reclassify their ELs, sometime thereafter (2–3 years) the progress of those former ELs moves off the radar.

As another example, the National Assessment of Educational Progress (NAEP) includes two categories of language proficiency: EL and Non-EL. Because states vary in terms of when and how they reclassify students and whether they assign and preserve in their database systems an RFEP classification, NAEP does not have the ability—across all states—to maintain an RFEP category. As a consequence, the achievement levels of ELs that have reclassified and moved out of the EL category are rarely reported. This significantly limits the utility of NAEP data because NAEP data cannot be used to analyze the progress of RFEPs and/or all initially classified ELs (EL + RFEP = IELs), only those that remain ELs.

Indeed, policy discussions on improving accountability for ELs under the reauthorization of the Elementary and Secondary Education Act (ESEA) have called attention to this "revolving door" EL cohort phenomenon and called for more precise accountability measures that consider the total cohort of students entering school needing to learn English (Working Group on ELL Policy, 2011). As Ramsey and O'Day (2010) describe,

> Membership in the EL accountability subgroup, however, changes in systematic and predictable ways over time. Higher-performing students are systematically removed from the subgroup when their English language skills reach a certain level, while less proficient students are constantly moving into the group, as newly-arrived ELs enter the system. . . . Indeed, subgroup progress is systematically underestimated because the more advanced students are no longer included in the determinations. (pp. 6-7)

This article seeks to inform these important policy discussions and also to inform the larger educational community, including school and district officials evaluating the progress of ELs at the local level and also federal and state department officials and researchers evaluating EL outcomes at the state and national level. As elaborated in the next section, others have called attention to the Catch-22 (Abedi, 2004; Grissom, 2004; Linquanti, 2001; Parrish et al., 2006); however, we know of no study that has examined empirically the effects of excluding and including RFEPs in the analysis of EL achievement. We know of no study that has examined achievement levels for current ELs, RFEPs, and IELs (ELs + RFEPs) to explicitly compare the extent to which achievement levels (a) differ among the groupings and (b) compare to non-EL populations like students from English-only backgrounds (EOs).

This article analyzes statewide achievement results from California. California data are particularly suited to this analysis because (a) the state disaggregates achievement results for current ELs and RFEPs across all tested grades (2nd through 11th), providing the opportunity to examine variation grade-by-grade across almost the entire K-12 span; (b) the state's recommended procedures for identifying and reclassifying ELs are among the most comprehensive in the country (see Background section); and (c) California provides a very large and relatively diverse sample of ELs, representing approximately 36% of all ELs educated in the United States and more than 60 different home languages (Wolf, Kao, Griffin, et al., 2008).

The Background section reviews current research on the complexities involved in identifying and reclassifying ELs and four recent studies that have examined EL achievement and the gaps that exist between ELs and non-ELs. Two studies do not include RFEPs, and two studies do include RFEPs. The contrasting results from these four studies contextualize the current study's research questions, which are provided at the end of the Background section.

The Methods section describes the datasets and statistics used in the analysis, as well as demographic data on California's K-12 enrollment, in general, and its EL population, specifically. To examine the effects of including RFEPs in the analysis of EL progress, the Results section analyzes (a) the proportions of ELs and RFEPs among all IELs across Grades 2 through 11; (b) achievement gaps between ELs and EOs, RFEPs and EOs, and IELs and EOs; and (c) changes in those proportions and gaps across a 5-year period. Based on the results of the analysis, the article closes with a discussion of implications for policy, assessment, research and evaluation, and other issues related to educating ELs.

Background

Identification and Reclassification of English Learners

The identification and reclassification of ELs is complex. Ongoing challenges include establishing and maintaining a clear and comprehensive definition of an EL (Abedi & Gándara, 2006), developing valid and reliable measures of English language proficiency (Abedi, 2008b; Abedi & Gándara, 2006; Genesee, Lindholm-Leary, Saunders, & Christian, 2006), and implementing and monitoring effective and consistent identification and reclassification criteria and procedures (Abedi, 2008a; Linquanti, 2001; Mahoney & MacSwan, 2005; Parrish et al., 2006). To elaborate on these challenges, we first explain current identification and reclassification procedures in California, and then we summarize research that delineates the complexities of specific aspects of those methods. We concentrate on California procedures as an instructive example and to provide additional background on the California data to be analyzed subsequently. Additionally, we provide national data that help contextualize California's practices.

Currently in California, ELs are identified and reclassified based on the following instruments and procedures (California Department of Education, 2011). Upon enrollment, parents complete a Home Language Survey. If parents indicate that a language other than English is spoken in the home, students are administered the California English Language Development Test (CELDT). The CELDT produces an overall proficiency score and level and separate scores and levels in each of four domains: Speaking, Listening, Reading, and Writing. The five levels are beginning, early intermediate, intermediate, early advanced, and advanced. Based on the initial CELDT administration, students are identified as ELs if their overall proficiency level falls below early advanced, and/or any one of their domain scores falls below intermediate.

Near the beginning of each year, the CELDT is administered to all identified ELs. When an EL produces an overall proficiency level of Early Advanced and all domain levels are at Intermediate or better, she or he is considered for reclassification. In addition to CELDT results, schools must also factor into the reclassification decision teacher evaluation of curriculum mastery, parental opinion and consultation, and a measure of basic skills in English, for which the California Department of Education recommends districts use the California Standards Test in English Language Arts (CST ELA). The recommended criterion for the CST ELA is performance in the basic range, where basic is the third of five levels of achievement: far below, below, basic, proficient, and advanced. Districts retain final authority over all reclassification decisions and can set their own CST cut-point within or above the basic range for CST ELA.

The instruments and procedures described above represent progress in the identification and reclassification of ELs, progress that is due primarily to the requirements of NCLB. Abedi (2008b) provides a thorough description of the NCLB requirements and processes through which states joined together in consortiums to meet those requirements. In sum, NCLB required states to develop standards, assessments, and accountabilities in accord with both Title I and Title III—that is, for all students and ELs. States had to develop English Language Development (ELD) standards and English language proficiency assessments (ELP) aligned with their content standards and assessments. California developed the CELDT and the CST. This provided all districts in California with a common ELP assessment and a common assessment of academic achievement. Prior to NCLB, states and districts chose from a wide range of commercially produced ELP assessments and

academic achievement assessments, and in many cases used only academic achievement measures to identify and reclassify ELs (Mahoney & MacSwan, 2005).

Despite the increased national accountability achieved over the past decade, states vary in terms of the design and rigor of their ELD standards; the weighting applied to the speaking, listening, reading, and writing portions of their ELP assessments; and the cut-points (standards setting) used to reclassify ELs (Abedi, 2008b, Ramsey & O'Day, 2010). They also vary in terms of the sources they use as part of their identification and reclassification procedures (Wolf, Kao, Herman, et al., 2008). Regarding the identification of ELs, like California, 46 of the 49 states and the District of Columbia report using a Home Language Survey, and 34 of the 50 entities use a sinle common ELP assessment. However, 16 states allow districts to choose their ELP assessment from an approved list. The number of levels of proficiency used by each state range (minimum of three and a maximum of six levels) and the criterion established to identify ELs varies, as well (Wolf, Kao, Griffin, et al., 2008). Regarding reclassification, only 9 other states use all four of the sources used by California: 46 states include *ELP test scores*, 28 include *academic achievement test scores*, 21 include *teacher evaluation*, and 15 incorporate *parent/guardian input* (Wolf, Kao, Griffin, et al., 2008).

Even within California, Parrish et al. (2006) found considerable variation within the state as to where, for the purposes of reclassification, districts set their cut-points on the CST. Some districts are more conservative and set their cut-point higher in an effort to ensure that students are successful upon exit from EL services and entry into all mainstream programming. Others are less conservative and set their cut-point lower in an effort to ensure that students have ready access to mainstream programming and college-preparatory coursework.

At the heart of the reclassification challenge is the matter of language proficiency and academic achievement and the critical, complex, and overlapping relationship between the two. Measures of language proficiency must address and measure *academic* language proficiency because such proficiency is critical for academic success in the mainstream classroom (Genesee et al., 2006; Goldenberg, 2008). As such, valid and reliable measures of academic language proficiency are pivotal to reclassification decisions. However, even the post-NCLB generation of ELP assessments are not yet providing such measures, and there remains a good deal of research and development needed to evaluate and systematically improve the actual validity and reliability of these assessments (Abedi, 2008b, Wolf, Kao, Herman, et al. 2008). In response to the NCLB mandates, states and consortia of states committed to the development and administration of ELP assessments that attempt to incorporate the construct of academic language, and by 2008 all states had implemented an ELP assessment (Ramsey & O'Day, 2010). With implementation achieved, states also need to commit to the long-term evaluation and refinement of the assessments (Abedi, 2008b; Wolf, Kao, Griffin, et al., 2008). The driving questions are as follows: To what extent are ELP assessments validly and reliably measuring not just language proficiency, in general, but *academic* language proficiency, specifically? How can the validity and reliability of these *academic* language proficiency assessments be improved? How can higher levels of consistency in criteria and measures be achieved across states (Wolf, Kao, Herman, et al., 2008)?

Beyond the assessments themselves, however, identification and reclassification are ultimately based on the professional judgments of teachers, coordinators, and administrators at the school sites. Abedi (2008a) tests empirically the benefits of an "augmented" system for identifying and reclassifying ELs. Based on a statistical analysis that generated composites across multiple measures of language proficiency and academic achievement and a step-wise selection process, the analysis demonstrates that the validity of EL classification can be improved given the deliberate and systematic use of multiple measures. Robinson (2011) puts forth and analyzes statistically a unique and provocative approach to reclassification: Reclassification decisions should be informed by local resources and weighted towards the best possible transition from services designed for EL to those services and resources available to support RFEPs. Outcomes should be carefully tracked and monitored to

systematically calibrate the threshold for retaining students in EL programming or reclassifying them. Through their statistical modeling, both Abedi (2008a) and Robinson (2011) demonstrate the complex decision making that identification and especially reclassification requires and the understanding and benefits that might be achieved by making that process the object of extensive, systematic implementation studies.

Reported Gaps Between Current ELs and Non-ELs or EOs

To contextualize the results reported in this article, we draw upon two examples from the literature that provide an estimate of the achievement gaps between EL and EOs or Non-ELs. Both studies reveal large gaps between current ELs and their counterparts; however, neither study addresses the RFEP population, and each therefore leaves open to question the achievement levels of all initially classified ELs (IELs = ELs + RFEPs).

Hemphill and Vanneman (2011) report gaps between Whites, Hispanic non-ELs, and Hispanic ELs in reading and math for Grades 4 and 8 across several NAEP administrations. Among other things, the analysis found *narrowing* gaps between Hispanic non-ELs and Whites and *widening* gaps between Hispanic-ELs and both Whites and Hispanic non-ELs. However, the report makes no mention of RFEPs and provides no explanation as to whether RFEPs were included in the EL or non-EL subgroup. That confounds the interpretation of the widening and narrowing gaps. Assuming Hispanic RFEPs are part of the Hispanic non-EL group (states identify ELs for NAEP, but not necessarily RFEPs), it is possible that Hispanic RFEPs (former ELs) are contributing to the narrowing gaps between Hispanic non-ELs and Whites. Moreover, if the proportion of Hispanic RFEPs relative to Hispanic ELs is increasing, that might explain the widening gap between Hispanic-ELs and both Whites and Hispanic non-ELs: More successful ELs (RFEPs) are moving into the non-EL group and reducing achievement levels in the remaining EL group. This is the "skimming" effect noted by Grissom (2004), Parrish et al. (2006), and Ramsey and O'Day (2010).

Aguilar (2010) reports achievement results for English language arts based on the 2009 and 2003 administration of the CST. The analysis focuses on the percentage of students scoring proficient or advanced (meeting NCLB requirements) and includes gaps between current ELs and EOs across Grades 3, 5, 8, and 10. In 2009, results reveal a gap of 33% between ELs and EOs at Grade 3 (20% of current ELs and 53% of EOs scored proficient or advanced), a 43% gap at Grade 5 (19% ELs and 62% of EOs); a 49% gap at Grade 8 (8% EL vs. 57% EO), and a 44% gap at Grade 10 (6% ELs vs. 50% EOs). Moreover, the gaps in 2009 tended to be 4 to 8 points larger than they were in 2003: an increase from 29% to 33% at Grade 3; 35% to 43% at Grade 5, 34% to 49% at Grade 8, and 37% to 44% at Grade 10. These California results are to be taken seriously as they portray very low levels of achievement among those that remain ELs (see Olsen, 2010: Long-Term English Learners). However, without the inclusion of RFEPs, these data might not provide a complete picture of achievement levels among all students initially classified as ELs (IELs = ELs + RFEPs). They only portray the relative standing of those that remain ELs. If from 2003 to 2009, larger proportions of ELs were reclassified as RFEPs and larger proportions of RFEPs scored proficient or advanced, these data would not reveal that positive change because they do not include RFEPs.

Studies That Have Included RFEPs

Two additional California studies contribute to this background: Grissom (2004) and Parrish et al. (2006). Whereas neither study focused on gaps between ELs and EOs or Non-ELs per se, both studies illustrate the importance of including RFEPs in the analysis. The Grissom study rebutted claims of proponents of California's Proposition 227, a state referendum passed in 1998 that virtually eliminated bilingual programming for ELs in favor of what is termed Structured English Immersion (SEI). After the first year of SEI implementation, the leading proponent of Proposition 227, Ron Unz, promoted an analysis that, among other things, purported to show significant gains in achievement among ELs. As reported in Grissom (2004), the purported gains were an artifact of comparing post-227 results

that combined ELs and RFEPs to pre-227 results that included only ELs. Grissom analyzed achievement results for three successive cohorts of initially classified ELs (ELs + RFEPs) over a 3-year period spanning Grades 2 through 5 and found no significant increases in EL, RFEP, or EO achievement levels over the 3-year period. However, more germane to this analysis, Grissom's results show RFEPs and EOs scoring slightly above national norms and ELs scoring well below national norms on standardized measures of English reading achievement. Given that RFEPs have to meet specific language proficiency and academic achievement criteria to reclassify, these results are not surprising. Assuming students are reclassified based on such criteria, RFEPs should theoretically always score substantially higher on measures of achievement than those ELs who are not meeting the reclassification criteria. Although differences between RFEP and EL levels of achievement may seem obvious, or even a tautology (of course, RFEPs will score substantially higher than ELs), Grissom's dataset makes that difference apparent empirically. In addition, Grissom reports the proportion of RFEPs and ELs grade-by-grade for each of his three cohorts. By 5th grade, almost one third of IELs had been reclassified (RFEPs). Had Grissom excluded RFEPs, he would have ignored almost a third of the IEL sample and distorted the portrait of their achievement levels.

Commissioned by the California Department of Education, the Parrish et al. (2006) study evaluated the implementation and impact of Proposition 227 on the education of ELs. Parrish et al. is the only study we have located that reports achievement data for EOs, ELs, RFEPs, and also IELs (ELs + RFEPs).

> Combining ELs and RFEP students into one group avoids the bias and distortion caused by "skimming" the best-performing ELs out of the EL category when they are redesignated as RFEPs. In addition to the EL/RFEP combined subgroup, we are reporting ELs and RFEPs separately. Our goal in presenting these data in two ways is to convey progress made by all students "ever EL" (including those former ELs who have been reclassified) and to highlight the performance of the RFEPs as a subgroup. (Parrish et al., 2006, p. III-15)

Whereas Parrish et al. (2006) made a methodological decision to include RFEPs to avoid "the bias and distortion" caused by excluding RFEPs, they do not make that bias and distortion the object of their analysis. Their report focuses on the impact of Proposition 227. However, data included in their report provide an estimate of the distortion they describe—the same distortion to be investigated in this article.

Parrish et al. (2006) calculated an annual standardized (reported in standard deviation units) gap size for RFEPs and EOs, ELs and EOs, and IELs (EL + RFEP) and EOs averaged across Grades 2 through 11 for each of several years of data based on both norm-referenced tests (SAT-9 and CAT6: language arts, reading, and math, 1998–2004) and criterion-reference tests (CST: reading/English language arts and math, 2002–2004). For all years and all measures, IEL-EO gaps are smaller than EL-EO gaps. In 2004, IEL-EO gaps were approximately 0.20 standard deviations smaller than EL-EO gaps. For example, on the CST of reading/English language arts, the gap averaged across Grades 2 through 11 between IELs and EOs was –0.61 standard deviations, whereas the gap between ELs and EOs was –0.84 standard deviations (0.23 difference). On CAT6 measures of English language arts, the average gap between IELs and EOs was –0.50, and the gap between ELs and EOs was –0.70 standard deviations (0.20 difference). On the criterion-referenced CST of mathematics, the average gap between IELs and EOs was –0.38 and the gap between ELs and EOs was –0.57 standard deviations (0.19 difference). On the norm-referenced CAT6, the average gap between IELs and EOs was –0.40 and the gap between ELs and EOs was –0.59 standard deviations (0.19 difference). If one is intending to evaluate the status of EL achievement in comparison to EOs and seeks to generalize to all initially classified ELs, the Parrish et al. dataset indicates that EL-EO gaps overestimates the IEL-EO gap by about 0.20 standard deviations.

In sum, the identification and reclassification of ELs is complex. Despite noteworthy progress since the introduction of NCLB requirements, significant challenges still remain in developing and implementing valid and reliable instruments and procedures for identifying and reclassifying ELs (Abedi, 2008b). These complexities and challenges notwithstanding, this article seeks to

examine the impact of including RFEPs in the analysis of EL progress. Conceptually, not including RFEPs underestimates the population of all students initially classified as ELs (Ramsey & O'Day, 2010). Empirically, existing studies suggest that not including RFEPs in the analysis of achievement outcomes overestimates achievement gaps and makes the evaluation of progress over time highly problematic.

We analyze statewide achievement data from California, and we report results for EOs, ELs, RFEPs, and IELs. First, we analyze data from the 2010 administration of the CST ELA, and then we compare 2010 and 2005 results. The analysis is framed around two questions:

1. What is the size of the gap between EOs and (a) current ELs, (b) RFEPs, and (c) IELs?
2. What, if any, changes are evident over time in these gap estimates between EOs and (a) current ELs, (b) RFEPs, and (c) IELs?

Our questions focus on gaps because the intent to close achievement gaps is a defining characteristic of the current reform era (NCLB, 2002) and because the analysis and reporting of EL-Non-EL or EL-EO gaps is widespread (Aguilar, 2010; August & Shanahan, 2006; Genesee, Lindholm-Leary, Saunders, & Christian, 2006; Goldenberg, 2008; Hemphill & Vanneman, 2011). Both questions focus on comparisons between EOs and ELs, RFEPs, and IELs to demonstrate the importance of the three comparisons, individually and in combination. Question 1 focuses on the gaps themselves. Question 2 focuses on identifying changes in those gaps over time, an undertaking that is influenced by the inclusion or exclusion of RFEPs.

Methods

All data were collected from the California Department of Education Standardized Testing and Reporting website (http://www.cde.ca.gov/ta/tg/sr/). Data include all students tested in the state of California at Grades 2 through 11 from the following three mutually exclusive groups: EOs, ELs, and RFEPs. In California, students who reclassify retain the RFEP designation in all subsequent years.[2]

The analysis includes the same academic achievement statistics that are currently reported by California in accord with No Child Left Behind accountabilities: the percentage of students at each of the five achievement levels: far below basic, below basic, basic, proficient, and advanced. In particular, we examine the cumulative percentage of students scoring proficient or advanced, the same criterion used for NCLB annual objectives. In some analyses, we also report the cumulative percentage of students scoring basic, or proficient or advanced (basic or better). All results are based on the CST ELA.

The analytic method is descriptive and designed to examine academic achievement gaps between ELs and EOs, RFEPs and EOs, and IELs (EL + RFEP) and EOs. The combined achievement results for IELs (EL + RFEP) are not reported at the CDE STAR website. To calculate IEL results we summed the raw numbers of ELs and RFEPs in each achievement band and then calculating the corresponding percentage for each band and combination of bands (proficient and advanced). Results from 2010 are reported first in order to investigate the magnitude of current gaps between the groups mentioned above (Research Question 1). Results from 2005 are then compared to those from 2010 to identify changes in gaps over time and the extent to which such changes are evident when comparisons do or do not include results for RFEPs (Research Question 2).

Results from 2010 are based on 4,197,225 second through 11th-grade students who took the ELA portion of the CST, of whom 1,675,446 were classified as ELs or RFEPs (39.9%) and 2,521,779 were classified as EOs (60.1%). Results from 2005 are based on 4,388,245 second through 11th-grade students who took the ELA portion of the CST, of whom 1,660,620 were classified as ELs or RFEPs (37.8%) and 2,727,625 were classified as EOs (62.2%). The California DOE reported testing 99% of English Learners in Grades 2 through 11 in both 2005 and 2010 (http://ayp.cde.ca.gov/reports/AcntRpt).

The California state reporting site disaggregates current EL results by those who have been enrolled in U.S. schools for 12 months or more

TABLE 1
Percentage of EL and RFEP Among All IEL by Grade Levels in the 2010 Administration of the CST ELA

Group	2nd	3rd	4th	5th	6th	7th	8th	9th	10th	11th
EL	95	84	74	62	52	47	43	45	43	40
RFEP	5	16	26	38	48	53	57	55	57	60

Note. EL = English Learners; RFEP = Reclassified as Fluent English Proficient; IEL = Initially Identified English Learners (ELs + RFEPs); CST ELA = California Standards Test in English Language Arts.

and those who have been enrolled less than 12 months. We did not utilize this disaggregation in our analysis. Rather, we analyzed the results for all current ELs to represent the entire current EL sample, which is—to the best of our knowledge—consistent with other analyses that have examined EL-EO gaps (e.g., Aguilar, 2010; Parrish et al., 2006). However, for context, in the 2010 sample, ELs enrolled in U.S. schools less than 12 months represent, on average, 2.0% of the IEL (EL + RFEP) sample at each grade level (range across grade levels is 1.6% to 2.1% but 3.1% at Grade 9) and, on average, 3.8% of the current EL sample at each grade level (range is 2.2% to 4.8% but 7.0% at Grade 9). In the 2005 sample, ELs enrolled in U.S. schools less than 12 months represent, on average, 4.3% of the IEL (EL + RFEP) sample at each grade (range is 2.8% to 5.6% but 6.1% at Grade 9) and, on average, 6.4% of the current ELs sample at each grade level (range is 5.0% to 7.4% but 10.3% at Grade 9).

The actual percentages of RFEPs and ELs per grade level for 2010 and 2005 are reported in the Results section. To calculate the percentage of RFEPs or ELs among all IELs at each grade level, we divided the number of each by the number of both (e.g., percent RFEP = [RFEP / EL + RFEP] × 100).

The data used for this analysis (aggregate results by language proficiency and grade level) do not allow us to disaggregate the EO, EL, and/or RFEP samples by other important demographic variables, including socioeconomic status, ethnicity, home language, and years in the U.S. As such, our analysis addresses the research questions at the most global level, a limitation we elaborate upon in the Discussion section. For the purpose of contextualizing these samples, however, the general characteristics of California's K-12 enrollment for 2010 are as follows (2005 data are in parentheses). Fifty-six percent of California's K-12 enrollment participates in the Free or Reduced Meals program (50%). The largest ethnicity groups are: 50% Hispanic/Latino (47%), 27% White (31%), 9% Asian (8%), 7% African American (8%), and 3% Filipino (3%). Students classified as English Learners constitute 24% of K-12 enrollment (25%); K-12 enrollment data for RFEPs are not available. Among those classified as ELs, the largest language groups include: 85% Spanish (85%), 3% Vietnamese (2%), 1% Filipino (1%), 1% Cantonese (1%), 1% Hmong (1%), 1% Korean (1%), 1% Mandarin (1%), and 1% Arabic (0.5%). The remaining ELs represent more than 50 other languages (http://www.cde.ca.gov/ds/sd/cb/dataquest.asp).

Results

Percentages of ELs and RFEPs Among All IELs and Achievement Levels of EOs, ELs, and RFEPs

Table 1 reports the percentage of ELs and RFEPs as a function of all IELs (EL + RFEP) that participated in the 2010 administration of the CST ELA. From Grades 2 through 11, the percentage of RFEPs increases while the percentage of ELs decreases. At 2nd grade in 2010, of all the students classified as EL or RFEP, 95% are EL and 5% are RFEP; at 5th grade, 62% are EL and 38% are RFEP; at 7th grade, 47% are EL and 53% are RFEP; and in 11th grade, 40% are EL and 60% are RFEP. The data in Table 1 provide an indication of the proportion of initially classified ELs that are left out of the analysis when the sample includes only "current" ELs and excludes RFEPs. By 5th grade, more than a one third (38% RFEP) are excluded. From Grades 7 through 11, more than

TABLE 2
Percentage of Grade 10 EO, EL, and RFEP by Achievement Bands, CST ELA, 2010

	Combined achievement bands			
Group	Far below & below	Basic	Proficient & advanced	Basic or better
EO	21	27	52	79
EL	65	29	6	35
RFEP	13	37	50	87

Note. In the 10th-grade cohort, RFEP represents 57% of all IEL (EL + RFEP). EL = English Learners; RFEP = Reclassified as Fluent English Proficient; EO = English-only; CST ELA = California Standards Test in English Language Arts; IEL = Initially Identified English Learners (ELs + RFEPs).

TABLE 3
Percentage of Grade 8 EO, EL, and RFEP by Achievement Bands, CST ELA, 2010

	Combined achievement bands			
Group	Far below & below	Basic	Proficient & advanced	Basic or better
EO	15	24	61	85
EL	53	36	11	47
RFEP	9	31	60	91

Note. In the 8th-grade cohort, RFEP represents 57% of all IEL (EL + RFEP). EL = English Learners; RFEP = Reclassified as Fluent English Proficient; EO = English-only; CST ELA = California Standards Test in English Language Arts; IEL = Initially Identified English Learners (ELs + RFEPs).

TABLE 4
Percentage of Grade 5 EO, EL, and RFEP by Achievement Bands, CST ELA, 2010

	Combined achievement bands			
Group	Far below & below	Basic	Proficient & advanced	Basic or better
EO	11	23	66	89
EL	36	42	22	64
RFEP	2	22	76	98

Note. In the 5th-grade cohort, RFEP represents 38% of all IEL (EL + RFEP). EL = English Learners; RFEP = Reclassified as Fluent English Proficient; IEL = Initially Identified English Learners (ELs + RFEPs); EO = English-only; CST ELA = California Standards Test in English Language Arts; IEL = Initially Identified English Learners (ELs + RFEPs).

half are excluded (53% RFEP in 7th and 60% RFEP in 11th).

Tables 2, 3, and 4 report achievement results for EOs, ELs, and RFEPs in Grades 10, 8, and 5. The purpose here is to examine differences in the distributions among the three groups across achievement levels and at grade levels with differing proportions of ELs and RFEPs. Section 2, then, examines achievement gaps across all grades, 2 through 11. To best illustrate the differences among the three groups, we have collapsed five achievement bands into three: (a) far below basic and also below basic, (b) basic, and (c) proficient and also advanced. Table 2 reports the percentage of 10th-grade EOs, ELs, and RFEPs scoring within each of these three bands.

We would like to make three points based on these 10th-grade comparisons. First, by 10th grade, current ELs and RFEPs have dramatically different levels of achievement. Fifty percent of RFEPs but only 6% of ELs are performing at proficient or advanced levels in English Language Arts. Moreover, 87% of RFEPs but only 35% of ELs are performing at basic, proficient, or advanced levels (basic or better). The low levels of achievement evident among current

ELs at the secondary school level have been documented elsewhere. Olsen (2010) reports on the characteristics and low levels of achievement among "Long-Term English Learners," that is, ELs that never reclassify. That this problem has been documented, named, and is being studied is clearly important, as the low levels of achievement are extreme among current ELs at the secondary level who represent approximately 40% of all IELs.

Second, as illustrated in Table 2, there is no achievement gap between EOs and RFEPs: 52% of EOs and 50% of RFEPs are performing at proficient or advanced levels in English Language Arts, and 79% of EOs and 87% of RFEPs are performing at basic or better. In a subsequent section, we will discuss some of the issues to keep in mind when comparing RFEP results to those of EOs (RFEPs have to score at the basic level of the CST and also demonstrate English language proficiency on the state's English Language Development Test to attain their RFEP classification; EOs do not), but the point here is that there is no perceptible gap.

Third, these 10th-grade data illustrate how reporting achievement levels—especially at the higher grades—of currently classified ELs without attention to RFEP results misrepresents—by omission—the performance of many—in fact a majority (57%)—that were initially classified as ELs.

The 10th-grade results from 2010 are not anomalous. The same pattern emerges throughout the grades. We review results from Grades 8 and 5 to illustrate the common and distinct patterns across middle and elementary school grades: RFEPs are much more likely than currently classified ELs and just as likely as EOs to score proficient or advanced and also basic or better on the ELA portion of the CST. Table 3 reports results for Grade 8, which reveal a pattern that is almost identical to the one observed in the 10th-grade data.

Table 4 reports results for Grade 5. At the elementary grades the pattern changes just a bit. RFEPs outperform EOs, and, in comparison to Grades 6 through 11, a larger percentage of ELs score proficient or advanced and basic or better. This is an artifact of the changing nature of the currently classified EL group and the RFEP group over Grades 2 through 11. As a true cohort of initially classified ELs changes from 100% EL (they are all ELs initially) to, say, 60% RFEPs and 40% current ELs in Grade 11, each successive grade skims off the most successful ELs for reclassification. As such, achievement levels tend to be very high for RFEPs at the elementary grades because the sample of RFEPs represents a small proportion of the most successful ELs. In turn, achievement levels among ELs tend to be higher at the elementary grades than at subsequent grades, at least in part, because many successful ELs who will reclassify within a year or two are still in the ELs sample. Recall from Table 1 that the proportion of RFEPs doubled across Grades 4 to 7, from 26% to 53%, respectively.

Even at the elementary grades the point remains clear: Analyzing results for only currently classified ELs and excluding RFEPs underestimates the population of all initially classified ELs and therefore does not provide an accurate picture of the achievement levels of IELs. Moreover, these 5th-grade results help accentuate the Catch-22 of current NCLB accountability practices. Seventy-six percent of these 2010 Grade 5 RFEPs scored proficient or advanced. In California, if these RFEPs continue to score proficient or advanced in ELA in Grades 6 and 7 (and/or have already done so as RFEPs), they will subsequently be removed from the EL subgroup and play no role in their middle or high school's EL accountabilities.

Question 1: What is the size of the gap between EOs and ELs, RFEPs, and IELs (ELs + RFEPs)?

Table 5 provides the percentage of EOs, RFEPs, ELs, and IELs (EL + RFEP) scoring proficient or advanced in 2010 for each of Grades 2 through 11. These grade-by-grade results provide an overall sense of the pattern across grades, including the overall achievement trends for the ELA portion of the California Standards Test as represented by the EO samples (i.e., how likely are native English speakers to score proficient or advanced at each grade level?). Grade-by-grade results also reveal the changing pattern of RFEP and EL results as increasing numbers of students in each successive grade level are RFEP. The data also provide

TABLE 5

Percentage Proficient or Advanced, EO, EL, and IEL (EL + RFEP), by Grade Level, CST ELA, 2010

Group	2nd	3rd	4th	5th	6th	7th	8th	9th	10th	11th	Row mean
EO	60	52	70	66	64	63	61	62	52	49	59.9
RFEP	80	70	88	76	64	63	60	61	50	44	65.6
EL	39	21	34	22	16	12	11	10	6	5	17.6
IEL (EL+RFEP)	41	29	48	42	39	39	39	38	31	29	37.5
Gap: RFEP vs. EO	+20	+18	+18	+10	0	0	–1	–1	–2	–5	+5.7
Gap: EL vs. EO	– 21	– 31	– 36	– 44	– 48	– 51	– 50	– 52	– 46	– 44	– 42.3
Gap: IEL vs. EO	– 19	– 23	– 22	– 24	– 25	– 24	– 22	– 24	– 21	– 20	– 22.4

Note. EL = English Learners; RFEP = Reclassified as Fluent English Proficient; IEL = Initially Identified English Learners (ELs + RFEPs); EO = English-only; CST ELA = California Standards Test in English Language Arts.

TABLE 6

Percentage Basic or Better, EO, EL, and IEL (EL + RFEP), by Grade Level, CST ELA, 2010

Group	2nd	3rd	4th	5th	6th	7th	8th	9th	10th	11th	Row mean
EO	84	82	89	89	89	86	85	84	79	74	84.1
RFEP	96	97	99	98	94	94	91	90	87	81	92.7
EL	71	58	71	64	57	49	47	44	35	26	52.2
IEL (EL+RFEP)	72	64	78	76	75	73	72	69	65	60	70.4
Gap: RFEP vs. EO	+12	+15	+10	+9	+5	+8	+6	+6	+8	+7	+8.6
Gap: EL vs. EO	– 13	– 24	– 18	– 25	– 32	– 37	– 38	– 40	– 44	– 48	– 31.9
Gap: IEL vs. EO	– 12	– 18	– 11	– 13	– 14	– 13	– 13	– 15	– 14	– 14	– 13.7

Note. EL = English Learners; RFEP = Reclassified as Fluent English Proficient; IEL = Initially Identified English Learners (ELs + RFEPs); EO = English-only; CST ELA = California Standards Test in English Language Arts.

answers to Question 1: What is the size of the gap between EOs and (a) current ELs, (b) RFEPs, and (c) IELs (EL + RFEP)? As shown in Table 5, the gaps between RFEPs and EOs are either in favor of RFEPs (2nd–5th) are nonexistent (6th–9th) or are small (10th–11th). The pattern is exactly the opposite for ELs and EOs between whom gaps increase from –21 in Grade 2 to between –44 and –52 across Grades 5 through 11. Averaging across the grade levels, the gap between currently classified ELs and EOs is –42.3. Both RFEPs and ELs show a pattern of decreasing percentages scoring proficient or advanced across Grades 2 through 11. As discussed earlier, this is greatly influenced by the changing composition of each group from grade to grade, with increasing numbers of ELs entering the RFEP group, making the RFEP group less selective across the grades, and increasing numbers of successful ELs leaving the EL group across the grades, leaving the least successful ELs to populate that group. One should also note, however, that from Grades 9 to 10 and 11, results also drop considerably for EOs, suggesting changes in the difficulty level of the 10th- and 11th-grade assessments and/or other variables at play (e.g., student motivation).

The only way to estimate achievement levels and gaps for all initially classified ELs (IELs) is to include all current ELs and RFEPs (IELs). As shown in Table 5, combining ELs and RFEPs (IELs) yields very different estimates of achievement gaps than the ones calculated based on results only for those that remain ELs. Whereas the EL-EO gaps range from –21 to –52 and average –42.3 across the grade levels, the gaps between IELs (EL + RFEP) and EOs range from –19 to –25 and average –22.4. The average gap between EOs and current ELs is 89% larger than the gap between EOs and IELs (42.3 – 22.4 / 22.4). To put it another way, the average gap between EOs and IELs is 47% smaller than the gap between EOs and current ELs (42.3 – 22.4 / 42.3).

Before moving on to Question 2, we include one additional comparison. Table 6 includes comparisons based on the cumulative percentage of students scoring basic, proficient, or advanced (i.e., basic or better). Data are reported for EOs, RFEPs, ELs, and IELs (EL + RFEP). It is informative to extend the analysis to include

all students scoring at the basic level or better because that cumulative percentage can be interpreted to mean all students who are meeting or at least approaching mastery of the ELA standards. On average, approximately 84% of EOs, 93% of RFEPs, 52% of ELs, and 70% of IELs scored basic or better. The gaps between IELs and EOs range from –12 to –18 and average –13.7, whereas the gaps between current ELs and EOs range from –13 to –48 and average –31.9. The average gap between current ELs and EOs is 133% larger than the gap between IELs and EOs (31.9 – 13.7 / 13.7); conversely, the average gap between IELs and EOs is 57% smaller than the gap between current ELs and EOs (31.9 – 13.7 / 31.9).

At least in this dataset, gaps between current ELs and EOs substantially overestimate the gap between IELs and EOs. Whether we look at any particular grade, or at the patterns across grades, or at averages taken across grades, analyzing the achievement of ELs based on data from only currently classified ELs and without RFEPs misrepresents the actual achievement status of all initially classified ELs. Indeed there are gaps between the achievement levels of such students and those students who come from English-only backgrounds. Currently, on average, 59.9% of EOs score proficient or advanced in ELA and 84.1% score basic or better. In contrast, only 37.5% of IELs (EL + RFEP) score proficient or advanced in ELA and 70.4% score basic or better. No doubt there is much work to be done to reduce the –22.4% gap at proficient or advanced and the –13.7% gap at basic or better. However, we would argue that these estimates, especially at the secondary school level and especially with all data disaggregated (EL, RFEP, and EL + RFEP), better inform that work as they provide a more complete picture of how all initially classified ELs are performing—a picture that is more complete than the one drawn by focusing solely on those students that remain English learners, which yields gaps that are approximately twice as large: on average –42.3% for proficient or advanced and –31.9% for basic or better.

Question 2: What, if any, changes are evident over time in these gap estimates between EOs and ELs, RFEPs, and IELs (ELs + RFEPs)?

In this section, we focus on Question 2: What, if any, changes are evident over time in these gap estimates between EOs and (a) current ELs, (b) RFEPs, and (c) IELs (ELs + RFEPs)? We hypothesized that positive changes over time would only be detectable in analyses involving RFEPs: Positive changes would be associated with larger proportions of RFEPs, and analyses based solely on current ELs would neglect that. In fact, a comparison of 2005 and 2010 results supports the hypothesis. The proportion of RFEPs is larger in 2010 than in 2005. The percentage of RFEPs scoring proficient or advanced on CST ELA is larger in 2010 than in 2005. The gap between EOs and IELs (EL + RFEPs) is smaller in 2010 than in 2005. However, none of these changes are evident in the gap analyses involving EOs and current ELs. In fact, the gap between EOs and ELs is larger in 2010 than in 2005. Without examining data for RFEPs and focusing only on data for current ELs, the 2005 and 2010 comparisons might lead to the faulty conclusion that all initially classified ELs are simply falling further and further behind EOs.

Table 7 includes the precentage of RFEPs at each grade level for both the 2005 and 2010 administrations of the ELA portion of CST. On average, the percentage of RFEPs among all IELs (EL + RFEP) is higher by about 12% in 2010 than in 2005. For example, in 2005 at Grade 5, RFEPs represented 25% and in 2010 they represented 38% of all IELs— a change of +13%. At Grade 8, the percentages of RFEPs in 2005 and 2010 changed by 16 points, from 41% to 57%; and at Grade 10, the percentage of RFEPs changed by 13 points from 44% to 57%. These differences between 2005 and 2010 might be attributable to any number of factors, for example, improved language learning among ELs, changes in reclassification procedures, changes in the EL population. Our intent here, however, is not to make claims about what might have caused these changes, but rather to bring to light the changes themselves. In this case, comparing 2005 to 2010 results is informed by the fact that grade level cohorts in 2010 had larger proportions of RFEPs, a fact that would go unrecognized if the analysis included only current ELs.

Table 8 reports the percentages of EOs and RFEPs that scored proficient or advanced on

TABLE 7

Percentage of RFEP Among All IEL by Grade Level, 2005 and 2010 Administrations of the CST ELA

Group	2nd	3rd	4th	5th	6th	7th	8th	9th	10th	11th	Row mean
2005 RFEP	3	9	15	25	34	39	41	41	44	48	29.9
2010 RFEP	5	16	26	38	48	53	57	55	57	60	41.5
Difference	+2	+7	+11	+13	+14	+14	+16	+14	+13	+12	+11.6

Note. EL = English Learners; RFEP = Reclassified as Fluent English Proficient; IEL = Initially Identified English Learners (ELs + RFEPs); CST ELA = California Standards Test in English Language Arts.

TABLE 8

Percentage Proficient or Advanced by Grade Level, CST, 2005 and 2010, EO and RFEP

Group/year	2nd	3rd	4th	5th	6th	7th	8th	9th	10th	11th	Row mean
EO 2005	51	41	57	52	47	51	47	51	44	42	48.3
EO 2010	60	52	70	66	64	63	61	62	52	49	59.9
Change	+9	+11	+13	+14	+17	+12	+14	+11	+8	+7	+11.6
RFEP 2005	62	54	76	63	46	54	42	48	35	35	51.5
RFEP 2010	80	70	88	76	64	63	60	61	50	44	65.6
Change	+18	+16	+12	+13	+18	+9	+18	+13	+15	+9	+14.1
Gap 2005	+11	+13	+19	+11	–1	+3	–5	–3	–9	–7	+3.2
Gap 2010	+20	+18	+18	+10	0	0	–1	–1	–2	–5	+5.7

Note. Gaps compare the RFEP percentage to the EO percentage within each year. Negative gaps indicate the RFEP percent is smaller than the EO percent; positive gaps indicate the RFEP percent is larger than the EO percent. EL = English Learners; RFEP = Reclassified as Fluent English Proficient; IEL = Initially Identified English Learners (ELs + RFEPs); EO = English-only; CST ELA = California Standards Test in English Language Arts.

CST ELA in 2005 and 2010. We already know that the grade-level cohorts in 2010 in comparison to 2005 had larger numbers of RFEPs relative to all initially classified ELs. Results in Table 8 also indicate that those larger numbers of RFEPs in 2010 were consistently more likely than the RFEPs in 2005 to score proficient or advanced on CST ELA. In 2005, on average, 51.5% of RFEPs scored proficient or advanced; in 2010, the average was 65.6%—an average increase of about 14.1 points. Positive changes are evident at every single grade level (range = +9 to +18). For example, the changes in the percentage scoring proficient or advanced among RFEPs for Grades 5, 8, and 10 are, respectively, +13 (63% to 76%), +18 (42% to 60%), and +15 (35% to 50%). Results in Table 8 also show positive changes across 2005 and 2010 for EO cohorts, among whom the average percent scoring proficient or advanced increased from 48.3% to 59.9%—an average increase of 11.6 points (range across grade levels = +7 to +17). With respect to the gap, RFEPs increased their advantage over EOs by 2.5 percentage points (3.2 to 5.7), showing larger gains than EOs at 7 out of 10 grade levels.

In contrast, comparisons in Table 9 between current ELs and EOs paint a different picture. Whereas more ELs tended to score proficient or advanced in 2010 than in 2005, those positive changes do not match the changes evident in the EO cohorts from 2005 to 2010. On average, the percentage of ELs scoring proficient or advanced increased from 10.2% in 2005 to 17.6% in 2010—an average increase of 7.4%. However, this 7-point increase among ELs is only about 2/3s the size of the increase evident among EOs: 11.6%. At all but two grade levels (2 and 4), increases among EOs from 2005 to 2010 exceed those of ELs, and the gaps between current ELs and EOs increased at each of those grade levels. Averaged across Grades 2 through 11, the gap between ELs and EOs increased from –38.1% in 2005 to –42.1% in 2010.

TABLE 9

Percentage Proficient or Advanced by Grade Level, CST, 2005 and 2010, EO and EL

Group	2nd	3rd	4th	5th	6th	7th	8th	9th	10th	11th	Row mean
EO 2005	51	41	57	52	47	51	47	51	44	42	48.3
EO 2010	60	52	70	66	64	63	61	62	52	49	59.9
Change	+9	+11	+13	+14	+17	+12	+14	+11	+8	+7	+11.6
EL 2005	22	12	19	13	7	9	6	7	3	4	10.2
EL 2010	39	21	34	22	16	12	11	10	6	5	17.6
Change	+17	+9	+15	+9	+9	+3	+5	+3	+3	+1	+7.4
Gap 2005	–29	–29	–38	–39	–40	–42	–41	–44	–41	–38	–38.1
Gap 2010	–21	–31	–36	–44	–48	–51	–50	–52	–46	–44	–42.1

Note. Gaps compare the EL percentage to the EO percentage within each year. Negative gaps indicate the EL percent is smaller than the EO percent. EL = English Learners; RFEP = Reclassified as Fluent English Proficient; IEL = Initially Identified English Learners (ELs + RFEPs); EO = English-only; CST ELA = California Standards Test in English Language Arts.

TABLE 10

Percentage Proficient or Advanced by Grade Level, CST, 2005 and 2010, EO and IEL (EL + RFEP)

Group	2nd	3rd	4th	5th	6th	7th	8th	9th	10th	11th	Row mean
EO 2005	51	41	57	52	47	51	47	51	44	42	48.3
EO 2010	60	52	70	66	64	63	61	62	52	49	59.9
Change	+9	+11	+13	+14	+17	+12	+14	+11	+8	+7	+11.6
IEL 2005	23	16	28	25	20	27	21	24	17	19	22.0
IEL 2010	41	29	48	42	39	39	39	38	31	29	37.5
Change	+18	+13	+20	+17	+19	+12	+18	+14	+14	+10	+15.5
Gap 2005	–28	–25	–29	–27	–27	–24	–26	–27	–27	–23	–26.2
Gap 2010	–19	–23	–22	–24	–25	–24	–22	–24	–21	–20	–22.4

Note. Gaps compare the IEL (EL + RFEP) percentage to the EO percentage within each year. Negative gaps indicate the IEL percent is smaller than the EO percent. EL = English Learners; RFEP = Reclassified as Fluent English Proficient; IEL = Initially Identified English Learners (ELs + RFEPs); EO = English-only; CST ELA = California Standards Test in English Language Arts.

We know from these analyses that the proportion of RFEPs was larger and the proportion of ELs was smaller in 2010 than in 2005. We know from these analyses that there were positive achievement changes from 2005 to 2010 among RFEPs, and those changes exceeded the gains of EOs, and therefore served to increase the advantage between RFEPs and EOs in favor of RFEPs. We also know that there were positive changes from 2005 to 2010 among current ELs, but those changes did not match the gains of EOs, and therefore served to increase the gap between ELs and EOs. The question now is, To what extent did these results affect the gap between EOs and IELs (EL + RFEPs)? Table 10 provides the 2005 and 2010 comparisons between EOs and IELs. At every grade level, increases from 2005 to 2010 among IELs either match or exceed those of EOs. On average, the percentage of IELs scoring proficient or advanced across Grades 2 through 11 increased from 22.0% to 37.5%, yielding an average change of +15.5 points—3.9 points greater than the average increase among EOs: +11.6 points. In 2005, the percentage of IELs scoring proficient or advanced was 26.3 percentage points less than EOs (48.3% vs. 22.0%). In 2010, the percentage of IELs scoring proficient or advanced was 22.4 percentage points less than EOs (59.9% vs. 37.5%). The gains among IELs (EL+RFEP) are a bit larger than those of EOs, yielding a 3.8 point reduction in the average achievement gap, from –26.2 to –22.4.

In sum, data from 2005 and 2010 indicate an 11.6% average increase across the grade levels in the percentage of RFEPs (among all IELs) and a 14.1% average increase in the percentage of RFEPs scoring proficient or advanced on the ELA portion of CST. That is, in 2010 there were more RFEPs, relative to ELs, and more RFEPs

performing well on the state achievement test. Independent of any causal attributions for these results, we are confident that most would view these as positive changes. However, none of these changes would have been evident if the analysis had focused solely on current ELs and neglected RFEPs. In terms of gaps, if the analysis had concentrated solely on current ELs, one would have found a 10.4% *increase* in the gap between ELs and EOs (42.1 – 38.1 / 38.1); one would not have found the 14.5% *reduction* in the gap between IELs and EOs (22.4 – 26.2 / 26.2). In short, excluding RFEPs and focus only on current ELs decreases the likelihood of detecting progress when positive changes in achievement have taken place.

Discussion and Implications

Focusing only on current ELs and excluding RFEPs in the analysis of EL progress runs the risk of (a) underestimating the population of all students initially classified as ELs (IELs), (b) overestimating the achievement gap between IELs and EOs, and (c) decreasing the likelihood of detecting progress when positive changes in achievement have taken place. The results of this analysis should inform discussions related to policy, assessment, research and evaluation, and broader discussions related to educating ELs.

First, federal and state policies should reflect greater attention to RFEPs—ELs that reclassify. Federal statute mandates specialized services for ELs, not RFEPs. However, that stipulation need not dictate who we monitor and who we do not. The Catch-22 can be eliminated. We need only officially add another category to complement the category of EL. Perhaps the most simple one to use is the one most commonly used, RFEP. A category for RFEPs needs to be included in any new or refined federal accountability system. One might consider two broad categories with two subgroups in each:

- Non-ELs (NELs): composed of students from English-only homes (EOs) and students who come to school fully proficient in English but also speak another language at home (Initially Fluent English Proficient, IFEPs), and
- Initial ELs (IELs): composed of current English Learners (ELs) and reclassified English Learners (RFEPs).

Moreover, all federal and state level data collection and assessment systems, including the National Assessment of Educational Progress (NAEP) and forthcoming assessments associated with the Common Core State Standards Initiative should track and report data for both currently classified ELs and RFEPs. One simply cannot track the progress of ELs without accounting for those ELs who schools, districts and states reclassify (RFEPs). Some might think this premature insofar as definitions and criteria for identifying and reclassifying ELs remain problematic and inconsistent from state to state and district to district. We argue that failing to track data on RFEPs only serves to perpetuate the ambiguity and inconsistency. Failing to track the academic achievement of reclassified ELs limits the empirical analysis of the ambiguity and inconsistency. Committing to tracking the progress of all initially classified ELs (current ELs and RFEPs) might provide a pathway towards reducing this ambiguity and inconsistency.

In the meantime, states and districts that have not done so should take on a thorough analysis of their EL and RFEP data, in particular the number, proportion, and achievement levels of RFEPs on a grade by grade basis. The number and proportion of RFEPs is important. Consider three hypothetical districts, each of which has only 35% of their current ELs meeting NCLB reading/language arts criteria. One of these districts has large numbers of RFEPs, most of whom are meeting criteria. Another district has large numbers of RFEPs, very few of whom are meeting criteria. The third district has almost no RFEPs at all. In terms of the success of their current ELs, they all look the same (35% meeting criteria); in terms of the success of all their initially classified ELs—including their RFEPs—they look quite different.

No doubt, the matter of reclassification, itself, is problematic insofar as criteria and assessments for reclassification vary from state to state. Moreover, districts are often given license to reclassify students based on combinations of measures and criteria. However, to be clear, we

do not advocate evaluating the progress of ELs simply by tallying and analyzing the number and percentage of RFEPs. Rather, we recommend using the RFEP category to evaluate *achievement* levels. In fact, examining the achievement levels of RFEPs is a reasonable way to evaluate the accuracy of one's reclassification process. Achievement levels of RFEPs should at least be on par with those of EOs.

Second, the research and evaluation community needs to take the illustrations provided here seriously, as well. RFEPs should be factored into study designs, data collections and explanations of results. Reporting achievement levels based on current ELs without any mention of RFEPs, the proportion they represent, and the levels of achievement RFEPs are attaining runs the risk of inadvertently misrepresenting the potential of ELs. Space limitations do not allow for a lengthy review of the numerous articles and chapters that begin with the same framing that ELs are a fast-growing and historically low-achieving population, performing well below their native English speaking peers. Without mentioning that substantial numbers of ELs do learn English and do perform on par with native English speakers, such statements and their accumulation run the risk of portraying the population of all initially classified ELs as chronic underachievers. Per statute, ELs have not yet demonstrated English language proficiency. When they do demonstrate proficiency, they are reclassified. When they reclassify, they typically no longer contribute to datasets used to estimate EL achievement. As such, most estimates of EL achievement are capped—limited by the absence of RFEPs and limited by the presence of those who do not reclassify and remain ELs. A good illustration of this comes from the NAEP results, which typically show lower reading achievement levels among ELs at Grade 8 than at Grade 4 (Goldenberg, 2008) and wider gaps between ELs and non-ELs at Grade 8 than at Grade 4 (Hemphill & Vanneman, 2011). Both the 4th and 8th grade samples are in fact English Learners, but they are comprised differently. The 8th grade sample will not include the kinds of English Learners that reclassify sometime after 4th grade and before 8th grade. Without explicating that, it simply appears that the reading achievement of ELs declines from grades 4 to 8.

Third, the results reported in this article should provoke some debate about reclassification rates, achievement levels of RFEPs, achievement levels of ELs, and comparisons between RFEPs and ELs and EOs. We reported the percentages of RFEPs at each grade level as a function of all IELs (ELs + RFEPs) in each grade level cohort throughout the state in 2010. We reported the percentages of RFEPs scoring proficient or advanced on CST ELA as a function of all RFEPs at the grade level cohort. We also reported the percentage of IELs scoring proficient or advanced on CST ELA. How will these results be viewed? Sixty percent of all IELs (ELs + RFEPs) at Grade 11 had been reclassified and 40% had not? Is that dramatically low and appalling? Or might it actually be higher than that which is going on in other states? How much do these proportions vary across the state from district to district? We actually have no way of knowing because such data are not regularly reported or discussed. Approximately 60% of students from English-only backgrounds are currently scoring proficient or advanced and 84% are scoring basic or better on the CST ELA. In contrast, among all IELs (ELs + RFEPs), approximately 38% are scoring proficient or advanced and 70% are scoring basic or better. How does that 22% gap in the percent scoring proficient or advanced or that 14% gap in the percent scoring basic or better compare to results in other states? To what extent do these gaps very across the state from district to district? Again, we have no way of knowing because such data are rarely reported. We must begin reporting and analyzing such data. Schools and districts throughout California and the nation will benefit from such analysis and the questions that drive them. What percentage of initially classified ELs should be reclassified by 5th, 8th, and 10th grade? What levels of achievement should reclassified students be expected to achieve? What is the magnitude and nature of the gaps to be closed between native English speakers and those that remain ELs?

Fourth, this analysis has brought into sharp relief the differences in achievement levels between ELs who reclassified and those who do not. Olsen's (2010) analysis has drawn important attention to the ELs at the middle and high

school levels who despite years of schooling in the U.S. do not reclassify—Long Term English Learners. While the purpose of this article is to shed light on the importance of including RFEPs in the analysis of EL progress, the data introduced add urgency to the matter of Long Term English Learners. The discrepancies in achievement between RFEPs and current ELs, especially by late middle school and into high school, are staggering and deserve to be emphasized. Recall the 2010 tenth-grade results reported in Table 2. Among current ELs in 10th grade, 65%—almost two thirds—scored below basic or far below basic on the CST ELA (46,971 of 72,263 tenth-grade ELs).

Fifth, the effort to better understand how best to educate ELs would benefit from greater attention to the progress of RFEPs. Our analysis only scratches the surface in demonstrating the importance of including RFEPs in the analysis of EL progress. There is much to be studied and learned about ELs that successfully reclassify and those that do not. By design, our study did not address the influence of socioeconomic status, ethnicity, home language, and/or years in the U.S. However, subsquent studies focused on both RFEPs and currrent ELs must begin to examine the relationships among of these demographic variables as well as program and opportunity to learn variables (e.g., teacher preparedness) as they are know to influence EL achievement (Abedi & Gándara, 2006). The RFEP results from California suggest that the population of all IELs bifurcates into Long Term English Learners and RFEPs, two subgroups characterized by dramatically different achievement levels. To what might that be attributed? What other variables best characterize each of these two subgroups? Are there student variables that distinguish RFEPs from Long Term English Learners? Are there program and opportunity to learn variables that significantly distinguish these two groups from one another? Are there early predictors that might identify ELs on course to successful reclassification and those not on course? Understanding better the student and program variables associated with successful ELs might shed important light on how best to support those that are less successful.

Notes

1. The American Heritage Dictionary defines Catch-22 as a situation in which a desired outcome or solution is impossible to attain because of a set of inherently illogical rules or conditions (http://ahdictionary.com).

2. The California Department of Education assessment site also reports results for students who come to school fully proficient in English but also speak another language at home. These students are classified as Initially Fluent English Proficient (IFEPs) and represent an important and interesting population. However, whereas they are not part of the initially classified EL population, their results are not included in this analysis.

References

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

Abedi, J. (2008a, Fall). Classification system for English language learners: Issues and recommendations. *Educational Measurement: Issues and Practices*, pp. 17–31.

Abedi, J. (2008b). Measuring students' level of English proficiency: Educational significance and assessment requirements. *Educational Assessment, 13*, 193–214.

Abedi, J., & Gándara, P. (2006, Winter). Performance of English language learners as a subgroup in a large-scale assessment: Interaction of research and policy. *Educational Measurement: Issues and Practices*, pp. 36–46.

Aguilar, V. (2010). Schooling English learners: Contexts and challenges. In California Department of Education, *Improving education for English learners: Research-based approaches*. (pp. 1–20). Sacramento, CA: California Department of Education.

August, D., & Shanahan, T. (2006). *Developing literacy in second-language learners: Report of the National Literacy Panel on Language Minority Children and Youth*. Mahwah, NJ: Lawrence Erlbaum.

California Department of Education. (2011). *California English Language Development Test: 2011-12 information guide*. Retrieved from www.cde.ca.gov/ta/tg/el/documents/celdtinfoguide1112.pdf

Genesee, F., Lindholm-Leary, K., Saunders, W. M., & Christian, D. (2006). *Educating English language learners: A synthesis of research evidence*. Cambridge, UK: Cambridge University Press.

Goldenberg, C. (2008). Teaching English language learners: What research does—and does not—say. *American Educator*, *32*(2), 8–23, 42–44.

Grissom, J. B. (2004). Reclassification of English learners. *Education Policy Analysis Archives*, *12*(36). Retrieved from http://epaa.asu.edu/epaa/v12n36/

Hemphill, F. C., & Vanneman, A. (2011). *Achievement gaps: How Hispanic and White students in public school perform in mathematics and reading on the National Assessment of Education Progress*. Washington, DC: National Center for Education Statistics, Institute of Education Sciences.

Linquanti, R. (2001). *The reclassification dilemma: Challenges and choices in fostering meaningful accountability for English learners* (UC Language Minority Research Institute Policy Report 2001-01). Davis: University of California, Davis.

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

Olsen. L. (2010). *Irreparable harm: Fulfilling the unkept promise of educational opportunity for California's long term English learners* (Research and Policy Publication). Long Beach, CA: Californian's Together.

Parrish, T., Merickel, A., Pérez, M., Linquanti, R., Socias, M., Spain, . . . Delancey, D. (2006). *Effects of the implementation of Proposition 227 on the education of English learners, K-12: Findings from a five-year evaluation* (Final Report). Palo Alto, CA: American Institutes for Research and WestEd. Available from http://www.wested.org/cs/we/view/rs/804

Ramsey, A., & O'Day, J. (2010). *Title III policy: State of the states* (Policy Brief). Washington, DC: American Institutes for Research and U.S. Department of Education.

Robinson, J. P. (2011). Evaluating criteria for English learner reclassification: A causal-effects approach using a binding-score regression discontinuity design with instrumental variables. *Educational Evaluation and Policy Analysis*, *33*(3), 267–292.

Wolf, M. K., Kao, J. C., Griffin, N., Herman, J. L., Bachman, P. L., Chang, S. M., & Farnsworth, T. (2008). *Issues in assessing English language learners: English language proficiency measures and accommodation uses—Practice review* (CRESST Report 732). Retrieved from University of California, Los Angeles, Center for Research on Evaluation, Standards, and Student Testing website: http://www.cse.ucla.edu/products/reports/R732.pdf.

Wolf, M. K., Kao, J. C., Herman, J. L., Bachman, L. F., Bailey, A. L., Bachman, P. L., . . . Chang, S. M. (2008). *Issues in assessing English language learners: English language proficiency measures and accommodation uses—Literature review* (CRESST Report 731). Retrieved from University of California, Los Angeles, Center for Research on Evaluation, Standards, and Student Testing website: http://www.cse.ucla.edu/products/reports/R731.pdf.

Working Group on ELL Policy. (2011, March 25). *Improving educational outcomes for English language learners: Recommendations for ESEA reauthorization* (Policy Brief). Palo Alto, CA: Author. Available from http://ellpolicy.org/

Authors

WILLIAM M. SAUNDERS is an Associate Research Psychologist at UCLA and Co-Chair of and a Research Director at the Talking Teaching Network, a non-profit educational research and development organization (bsaunder@ucla.edu).

DAVID J. MARCELLETTI is Co-Chair of and a Research Director at the Talking Teaching Network (dave.marcelletti@talkingteaching.org).

Manuscript received July 28, 2011
Revision received March 14, 2012
Accepted July 30, 2012