Research and Policy Considerations for English Learner Equity

Joseph P. Robinson-Cimpian, University of Illinois at Urbana-Champaign

Karen D. Thompson, Oregon State University

Ilana M. Umansky, University of Oregon

Published in (2016). *Policy Insights from the Behavioral and Brian Sciences, 3*(1), 129-137.

Contact author information:

Joseph P. Robinson-Cimpian, Ph.D.

Department of Educational Psychology

University of Illinois at Urbana-Champaign

210F Education Bldg., 1310 S. 6th St.

Champaign, IL 61820

jpr@illinois.edu

217-333-8527 (phone)

Note: We thank Diane August for helpful comments. This research was funded by a National Academy of Education/Spencer Foundation Postdoctoral Fellowship awarded to Robinson-Cimpian, as well as the U.S. Department of Education Institute of Education Sciences grant R305H140072 for which Thompson is the PI. The research presented here does not necessarily reflect the opinions of the funding agencies.

Abstract

English learners (ELs), students from a home where a language other than English is spoken and who are in the process of developing English proficiency themselves, represent over 10% of the US student population. Oftentimes education policies and practices create barriers for ELs to achieve access and outcomes that are equitable to those of their non-EL peers. We discuss how recent education research—often using experimental and quasi-experimental designs—provides new insights on how to evaluate EL policies, as well as how best to alter current policies to yield more equitable outcomes for ELs. Topics discussed include (1) EL classification and services, (2) language of instruction, (3) access to core content, and (4) assessments.

Key Points

- Research suggests that many currently implemented education policies likely contribute to inequitable access and outcomes for English learners (ELs).
- Recent studies—often using experimental and quasi-experimental designs—suggest ways in which these policies can be altered to improve equity.
- New research suggests how policies governing EL classification, as well as the corresponding settings and services provided, can be rigorously evaluated and amended.
- Recent rigorous research on bilingual programs shows positive long-term student outcomes, particularly in dual immersion programs.
- Tracking ELs into low-level classes and supplanting academic content with English support services limit ELs' access to core curricular content.
- Because test scores can have high-stakes consequences for students, teachers, and schools, assessments for ELs must be valid and reliable.

Tweet: New paper suggests equitable education policies for English learner students

Research and Policy Considerations for English Learner Equity

Currently, one in five students in U.S. public schools speak a language other than English at home (Ryan, 2013). Roughly half of these students, over four million children, are in the process of developing proficiency in English and are classified as English learners (ELs, U.S. Department of Education, 2015). Forty years ago, in the landmark case *Lau v. Nichols* (1974), the Supreme Court ruled, "[T]here is no equality of treatment merely by providing students with the same facilities, textbooks, teachers, and curriculum; for students who do not understand English are effectively foreclosed from any meaningful education." Citing Title VI of the Civil Rights Act (1964), which prohibits discrimination on the basis of "race, color, and national origin" in any federally-funded program, the Court held that school districts were obligated to take "affirmative steps" to effectively educate students acquiring English. But which "affirmative steps" best enable school systems to meet ELs' needs has long been contested.

Here we provide an overview of empirical research on four topics crucial to ensuring an equitable education for English learners: (1) policies for classifying students as ELs and reclassifying students as English proficient; (2) the use of students' primary languages for instruction; (3) access to grade-appropriate instruction in the content areas while students are in the process of acquiring English; and (4) the design of meaningful assessment and accountability systems for ELs. While other topics, such as school funding and teacher education, are also important in providing an equitable education for English learners, we focus on the four areas above because the empirical research base in these areas is the most robust. After reviewing the research in each area, we discuss policy implications.

Classification as EL and Reclassification as English Proficient

Determining which students should be considered ELs and which services they should receive is one of the most fundamental, yet challenging, issues for policymakers. Research on the topics of initial *classification* (as EL) and *reclassification* (as English proficient) has focused on two primary policy-relevant questions: (1) Once a student is classified as EL, *how many years* does it typically take for the student to attain English proficiency and thus be reclassified? and (2) How do policymakers establish *appropriate criteria* for initial classification and subsequent reclassification to ensure that students who need services are receiving them?

Research on the question of time to reclassification suggests that the answer rests on a number of factors, including characteristics of the student and the criteria used—but in general, attaining English proficiency takes time. A frequently cited study by Hakuta, Butler, and Witt (2000) used cross-sectional data and concluded that oral English proficiency took 2-5 years for the majority of students, while proficiency in English language arts (ELA) took about 4-7 years for most students. More recently, researchers have used survival analysis to examine time to a particular milestone with longitudinal student-level data (e.g., Conger, 2009; Thompson, 2015a; Umansky & Reardon, 2014). Although these survival analysis-based studies used similar methods, their varying conclusions illustrate how the criteria established by different districts affect the expected time horizons. For example, Conger (2009) found that, although some demographic groups took more or less time on average, the median time to attain the required level of English language proficiency in New York City was about 3 years. By contrast, the two different large urban districts in California studied by Thompson (2015a) and Umansky and Reardon (2014) required several more criteria, including a core content test of ELA and teacher evaluations. These California studies found that median time to reclassification was about 6-6.5

years, and in each district, more than one-quarter of students were not reclassified after nine years. Taking these studies together, Hakuta et al.'s (2000) estimated timeframes remain consistent with the new research findings, suggesting that it take most students multiple years to be reclassified and that timing to reclassification varies considerably due to both individual and structural factors.

Turning now to the question of how to establish appropriate criteria for classification and reclassification, the above discussion of time to reclassification makes clear that criteria vary substantially across states and even across districts within states (Linquanti & Cook, 2015; National Research Council, 2011). However, the types of criteria used typically consist of (1) measures of English language proficiency (ELP), (2) measures of academic achievement, and (3) teacher input. A measure of ELP is the most common and basic requirement to attain English proficient status. Such criteria are in place to ascertain whether or not a given student requires ongoing English support (as an EL) or can be mainstreamed in school (as a former EL). Although we must remember that the test-developer intent for an ELP assessment is to measure the construct of ELP, not to attach interpretation to a particular test value or to create a binary category of EL/non-EL to receive different services (Haertel & Ho, in press), research suggests that ELP tests should be the primary factor in reclassification decisions (Linquanti & Cook, 2015; Umansky et al., 2015).

The academic content-area criteria for reclassification are more controversial. Most common is the inclusion of a standardized measure of ELA achievement. Less frequent is the inclusion of standardized measures of math achievement and/or grades. The main arguments for including achievement measures are to ensure (1) that a given student is academically prepared to succeed in a mainstream environment, and (2) that EL programming provides sufficient

academic content to EL students (Linquanti, 2001; Ragan & Lesaux, 2006). Critics counter (1) that ELs should not be required to meet academic requirements that native-English speakers need not meet in order to be in mainstream classes, (2) that EL students should not be held accountable for poor academic performance that may, in part, stem from the provision of less-than-adequate educational opportunities as ELs, and (3) that academic assessment of ELs is plagued with validity and reliability issues (e.g., Abedi, 2004; Solórzano, 2007). Several studies also found that academic achievement measures (e.g., ELA tests) tend to take on a more prominent role in restricting reclassification eligibility at higher grade levels (Robinson, 2011; Umansky & Reardon, 2014); thus, it is often not the case that high-performing long-term ELs lack English proficiency, but rather core content tests impede reclassification.

Another important dimension of whether criteria are "appropriate" extends beyond psychometric and philosophical concerns, focusing instead on *evaluating the effects* of existing classification and reclassification criteria on subsequent achievement and graduation. Robinson (2011) argued that effects of reclassification at a test-based, policy-specified threshold suggest that there is misalignment between the services/settings provided to ELs before and after they are reclassified; and thus the ideal situation would be no effects of reclassification, which would suggest a smooth transition from EL to reclassified status. More specifically, Robinson argued that policymakers should consider the student's linguistic needs and services/settings provided to different groups of students when establishing thresholds for reclassification eligibility on tests of English proficiency. For example, if the policy sets the reclassification bar low in terms of English language proficiency, when students are still benefiting from services intended for ELs, then we would expect reclassification to have negative effects on subsequent achievement and graduation. Conversely, if the policy sets the bar too high, when English supports are not needed

and time might be better spent on other learning opportunities, then students who barely failed to meet the criteria will underperform relative to their otherwise identical peers who were reclassified.[1]

Using regression discontinuity designs (RDDs) to compare the outcomes of students who just barely met or failed to meet the reclassification criteria, researchers have provided the most rigorous evidence to date on the effects of reclassification policies (e.g., Robinson, 2011). These RDD studies compare the outcomes of nearly identical sets of students who receive different treatments, thereby removing selection bias from the estimated effects of reclassification. More recent work has demonstrated that, when there is evidence of misalignment in services at the policy threshold, policymakers can *shift the threshold* in the appropriate direction to produce the more desirable outcome of smooth transitions (Robinson-Cimpian & Thompson, in press). But shifting the threshold is not the only way to alter reclassification effects—educators can work with existing thresholds, and instead realign instructional services. Ongoing work with one state education agency has used multi-site RDDs to identify effects of reclassification at both the state level and district level to assess (1) whether the state threshold is appropriate on average and (2) whether it works well in all district contexts. This research suggests that, even at the *same* state-specified threshold, the effects of reclassification can vary widely across districts (Robinson-Cimpian, Thompson, & Makowski, revision under review). The above studies have all examined

[1] This example illustrates how language proficiency can interact with the policy threshold to produce academic disruptions. However, language proficiency is not the only reason why some ELs may do poorly once reclassified. If students did not have the opportunity to build foundational content knowledge while classified as ELs, they may struggle in content-area courses after reclassification. Importantly, the RDD methods described in this section help to identify if there is evidence of misalignment between the services/settings provided to students at the policy threshold, but they do not identify the misalignment source (e.g., language demands, access to core content). Follow-up RDDs or in-depth qualitative analysis may help to identify these sources (Robinson-Cimpian et al., revision under review).

the effects of reclassification, but emerging research is examining the effects of initial classification as EL (Umansky, 2013).

The collection of RDD studies on reclassification effects suggest that (1) by setting test-based thresholds, policymakers have tremendous influence over when a student is reclassified; (2) given the services available in a district, a misplaced threshold can lead to substantial negative effects on achievement, course-taking, and graduation, for either the students who remain ELs or those who are reclassified; (3) policymakers can shift the thresholds to change the effects of reclassification; and (4) even at the same threshold, different districts can have different effects depending on their unique set of services and circumstances.

Research-based Policy Recommendations Regarding (Re)classification

- Recognize that students vary in the time required to reach English proficiency, but that most research suggests the average time to proficiency is between 4 to 7 years.
- Avoid setting a pre-determined maximum number of years for receiving EL services.
- When making classification and reclassification decisions, emphasize more construct-relevant factors (e.g., English language proficiency scores) and deemphasize less relevant ones (e.g., academic test scores).
- Understand that reclassification can have effects on subsequent student outcomes.
- Do not assume that schools or districts with higher reclassification rates are necessarily serving students better—in fact, they may be removing beneficial services too soon, and in turn, causing lower graduation rates.
- Evaluate criteria used in reclassification decisions. Use rigorous RDDs for these evaluations whenever possible, then follow-up with districts identified as reclassifying students too soon or too late given the services/settings available.

- Adjust reclassification thresholds/criteria, realign services, and provide additional supports to struggling schools and districts accordingly.

Language of Instruction

Perhaps the most heated debate in EL policy has been whether and to what extent students' primary languages should be used for instruction. Proponents of English-only models suggest that if students are exposed to more English, they will learn English more quickly (e.g., Rossell & Baker, 1996). These arguments undergird the restrictive language policies enacted by several states, which prohibit the use of languages other than English for instruction (Gándara & Hopkins, 2010). On the other hand, proponents of bilingual education assert that by learning academic content in their primary language while simultaneously developing English proficiency, students will be able to understand content-area instruction and ultimately transfer skills and knowledge from their primary language to English (e.g., Cummins, 2000).

On balance, the vast literature on this question suggests that in the medium to long-term bilingual programs have, at best, moderate positive effects, and at worst, no negative effects on students' acquisition of English and on their content-area achievement in English. Five meta-analyses conducted over the past 30 years all concluded that bilingual programs had significant small to moderate positive effects on outcomes in English, including English proficiency, English language arts, and math (August & Shanahan, 2006; Greene, 1997; Rolstad, Mahoney, & Glass, 2005; Slavin & Cheung, 2005; Willig, 1985). A separate meta-analysis conducted by Rossell & Baker (1996) came to the conclusion that bilingual programs had negative effects on student outcomes, but Greene (1997) demonstrated a variety of problems with the inclusion criteria for this meta-analysis. When considering only studies using experimental methods,

bilingual education showed positive effects of approximately 0.3 SDs on English language outcomes (Greene, 1997; Slavin et al., 2005).

Four studies that occurred too recently to be included in these meta-analyses provide important additional information about whether and to what extent ELs' primary languages should be used for instruction. Slavin and colleagues (2011) conducted a randomized controlled trial in which students were enrolled in either a transitional bilingual program or an English immersion program. While students in the English immersion setting scored higher on English reading assessments in the primary grades, by fourth grade, there were no significant differences on these assessments for the students in the two programs.

In two other studies, researchers also found initial advantages on outcome measures in English for ELs in English-only programs. However, these two studies were able to analyze student outcomes over a longer period of time, and found that at the secondary level, ELs in bilingual programs ultimately outperformed their peers who received English-only instruction (Umansky & Reardon, 2014; Valentino & Reardon, 2015). In the first of these studies, Umansky and Reardon (2014) analyzed the time necessary for Latino ELs to be reclassified as fully English proficient when enrolled in English-only instruction or three different types of bilingual programs. They found that students in English-only instruction were initially more likely to reach English proficiency, but by high school, students in bilingual programs had surpassed this group in English proficiency likelihood. In the second study, Valentino and Reardon (2015) used growth models to compare ELs' performance in ELA and math through middle school. Similarly, they found that ELs in bilingual programs had lower ELA and math scores in early elementary school than ELs receiving English-only instruction, but test score growth rates of ELs

in bilingual programs, particularly dual immersion programs[2], exceeded growth rates for ELs in English-only classrooms. This led to better long-term outcomes for ELs in bilingual programs, though results varied somewhat by subject, ethnicity, and type of bilingual program. While these two studies are not experimental, they use a rich set of controls, including demographic characteristics, students' initial English proficiency, school-level fixed effects, and parents' preferences for school and program type.

Finally, a recent large-scale quasi-experimental study analyzes the causal effect of dual immersion programs on outcomes for ELs and native English speakers using data from Portland, Oregon, which uses a lottery to assign students to immersion programs (Steele et al., 2015). Findings from this study show significant positive effects of dual immersion on English reading outcomes for both ELs and native English speakers, ranging from 0.13 standard deviations in 5th grade to 0.22 standard deviations in 8th grade. By middle school, ELs in immersion programs are significantly more likely to be reclassified as English proficient than their peers who applied for but did not win slots in immersion programs, and this effect is stronger for ELs whose native language matches the partner language used in the immersion program (i.e., for Spanish-speaking ELs enrolled in Spanish-English dual immersion programs). Students winning slots in immersion programs have scores on math and science assessments administered in English that are not statistically distinguishable from their peers in English-only classrooms even though students in immersion program receive math and science instruction at least partially in the partner (non-English) language through fifth grade. Given the promise of dual language immersion programs several states including New York and Oregon are currently funding their expansion (Manning, 2014; Harris, 2015).

[2] Dual immersion programs enroll both ELs and native English speakers, with the goal of developing bilingualism and biliteracy for all students.

While bilingual programs' effects on outcomes in English are important, it is also useful to consider other outcomes. Not surprisingly, research shows that students in bilingual programs have significantly higher outcomes on assessments given in the partner languages used in bilingual programs than students in English-only programs do (Barnett et al., 2007; Greene, 1997). These positive outcomes on assessments in other languages are important in light of recent studies demonstrating that full bilingualism is associated with a variety of positive long-term outcomes. For example, bilingualism is associated with lower dropout rates, higher earnings, and higher educational attainment (Callahan & Gandára, 2014). Additional experimental research has shown that bilingualism produces a variety of cognitive health benefits, including stronger executive function and lower incidences of Alzheimer's (Adesope, Lavin, Thompson, & Ungerleider, 2010; Craik, Bialystok, & Freedman, 2010).

Research-based Policy Recommendations Regarding Language of Instruction

- Eliminate restrictive language policies currently in place in several states, which prohibit the use of languages other than English for instruction.
- Given the particular promise of dual immersion programs, consider incentivizing the development and/or expansion of dual immersion programs.
- Ensure that evaluations of bilingual programs consider long-term student outcomes, at least past elementary school, to avoid drawing inaccurate conclusions about program effectiveness.

English Learners' Access to Core Content

Access to core content lies at the heart of federal law concerning the education of English learner students. Law and regulation regarding the education of ELs are framed around ELs'

twin rights: to support learning English, and to accessible grade-level core content. Yet ensuring students' right to equitable and full access to core content has proved elusive. Research identifies four main ways in which access to core content is frequently limited for ELs: (1) English-only instruction without appropriate accommodations, (2) weak and/or slow-paced curriculum in separated classes for ELs, (3) tracking into low-track (low-level) classes, and (4) exclusion from core subject area classes.

English-only instruction without appropriate accommodations. Despite the *Lau v. Nichols* ruling that English-only instruction without accommodations effectively bars ELs from access to content, ELs continue, at times, to be placed into just such classrooms. Teachers widely state that they feel insufficiently prepared to work with their EL students (Gándara, Maxwell-Jolly, & Driscoll, 2005). Furthermore, research suggests that the "sink or swim" placement of EL students may be more acute in some core subject areas, like math (Hopkins, Lowenhaupt, & Sweet, 2015). The frequency of this practice also varies considerably by school and district. When it occurs, however, it severely limits ELs' ability to access and learn content, particularly among ELs with low levels of English proficiency.

Weak and/or slow-paced curriculum in separated classes for ELs. English learners are often placed into classrooms (at the elementary school level) and core content area classes (at the middle and high school levels) that enroll only or primarily other EL students. The purpose of this placement is to ensure that ELs are in classes that use pedagogical practices that are accessible to ELs. Yet often these classes offer diminished, slower-paced, and/or less rigorous content (Dabach, 2014; Harklau, 1994). Teachers struggle to provide grade-level core content instruction in English to students who are not English proficient (Gándara et al., 2005). Teachers may also have lower expectations of their EL students resulting in inferior instruction and

content-coverage (Blanchard & Muller, 2015; García-Nevarez, Stafford, & Arias, 2005). In addition, EL-specific classes tend to be taught by less-qualified and less experienced teachers (Dabach, 2015; Gándara, Rumberger, Maxwell-Jolly, & Callahan, 2003). Finally, EL students in these classes have little exposure to English-speaking peers and meaningful content-based dialogue in English, key to English acquisition (Dabach, 2014). Isolation in these classes with inferior content directly influences students' opportunity to learn, and it also often generates social stigma toward English learners (Thompson, In Press).

Tracking into low-track classes. ELs tend to be over-represented in remedial and low-track classes and under-represented in advanced placement, honors, and other upper-track classes compared to their English proficient peers. In addition, ELs are more likely to be in slower versus accelerated course sequences, such as math course sequences in middle and high school (Thompson, 2015). Part of this disproportionality is due to ELs' lower average academic performance, which results in placement into lower-track classes. English learners may have lower academic outcomes for a host of reasons. One reason is limited understanding of material taught in English without sufficient modifications. Indeed, this is one of the main rationales for bilingual instruction; students are less likely to fall behind academically if they have access to content area instruction in a language they understand while they are acquiring English.

Some of the disproportionality in course placement, however, is a direct result of classification as an English learner (Umansky, 2014). For example, EL-classified students may be ineligible for advanced or grade-level classes and even reclassified students (former ELs) may be automatically routed into remedial level courses (Kanno & Kangas, 2014). This is problematic because lower-track classes have been shown to offer less exposure to content, and to employ more passive and rote pedagogical practices (Oakes, 2005). While providing rigorous

core content, such as high-track math courses, in ways that allow ELs access to the material can be a technical challenge, particularly for newcomer students, case studies of particular schools and programs offer examples of how this can be possible. For example, a group of California high schools successfully used online math and science curriculum in Spanish with newcomer students, enabling students to learn grade-level content in their primary language while they learned English during other parts of the school day (Hopkins, Martinez-Wenzl, Aldana, & Gándara, 2013).

Exclusion from core subject area classes. There is wide variation in course placement practices for ELs by school and district (Estrada, 2014). In some cases, ELs have full access to core content instruction and in others, their EL status prevents or limits enrollment in core courses (Callahan, 2005; Callahan, Wilkinson, & Muller, 2010). For example Arizona's policy to place ELs in four hours of daily English language development (ELD) instruction severely limits the amount of time students can be exposed to academic content (Lillie, Markos, Arias, & Wiley, 2012). Even in states and districts with far less extreme English language instruction policies, daily ELD classes often crowd out or replace core content area classes and instruction, especially English language arts. In middle school, evidence from one school district in California shows that on average over one third of EL students are not enrolled in a full course load in any given semester (Umansky, 2014). Exclusion from core instruction and core subject areas can severely curtail students' ability to meet graduation and post-secondary enrollment requirements, and can slow students' progression through school.

Research-based Policy Recommendations Regarding Core-Content Access

- Provide more guidance, monitoring and accountability to ensure that ELs are provided with equitable access to core content, including college-track and advanced level courses.

- Provide support and evaluate efforts to avoid crowding out of academic access by English language instruction. Two possibilities that require more evaluation are extending the school day/year for ELs and integrating language and content instruction into the same classes.
- When possible, consider making core content instruction available in students' primary languages while students are in the process of learning English, particularly for newcomer students.
- Provide targeted professional development on ways to provide grade-level content to students who are acquiring English proficiency both in separated EL classes and in mainstream classes.
- Learn from districts and schools that have implemented models enabling ELs to enroll in full course loads and college-track courses.

Assessments and Accountability for EL Students

In accordance with federal law under *No Child Left Behind* (NCLB; 2002), all students in grades 3-8 must be assessed annually in ELA and math. In addition to meeting "adequate yearly progress" for the population of students in a school, the law stipulates that annual growth targets must be met for subpopulations, such as ELs. This subpopulation requirement for school-level accountability brought revitalized attention to many longstanding concerns regarding the validity and reliability of standardized academic assessments for ELs (Abedi, 2004). Now, in addition to the student-level implications of invalid and unreliable assessments, there are school-level implications as well because schools with lower-than-expected achievement gains for ELs could be labeled "in need of improvement." In this section, we focus on three prominent assessment

issues: (1) how assessments of ELs relate to accountability, (2) consequences of invalid and unreliable assessments, and (3) how to make assessments for ELs more valid and reliable.

As noted above, the standardized assessment scores of ELs are used for school-level accountability, both through contributing to the school's overall score and to the subpopulation score. Unlike some subpopulations (e.g., racial minorities), however, the label of EL is transitory for most students, which creates a state of constant flux as students move in and out of this category. This category instability presents challenges for assessing subpopulation growth, as the highest achieving students tend to exit the EL category each year. Abedi (2004) notes that several states proposed using a "once LEP [limited English proficient], always LEP" accounting strategy, but this was not allowed. Ultimately, the federal law was amended to include in the EL accountability category students who are currently ELs or were reclassified in the previous two years. Research demonstrates that failing to include former ELs as part of the "ever EL" category leads to overestimation of achievement gaps and underestimation of progress made by students who were once ELs (Hopkins, Thompson, Linquanti, Hakuta, & August, 2013; Saunders & Marcelletti, 2013). Moreover, because ELs' content-area assessment scores are highly correlated with their English proficiency levels, ELs at the beginning stages of learning English are extremely unlikely to meet grade-level standards on English language arts and math assessments, impacting schools accountability ratings (Hopkins et al., 2013). Therefore, researchers have argued that more realistic expectations for ELs' content-area assessment scores should be established, taking into account their level of English proficiency and their time in U.S. schools (Hopkins et al., 2013).

In addition to school-level accountability concerns related to which students to include in the EL subgroup, the validity of assessments for ELs can have a substantial effect on measures of

"teacher value-added" (i.e., the average amount of year-to-year achievement gains students experience with a specific teacher). Value added modeling has its own criticisms, which we will not discuss here (but see, e.g., Reardon & Raudenbush, 2009; Rothstein, 2010). Instead, we highlight some recent criticisms directed at the inclusion of EL scores in value-added estimates: (1) most tests are less reliable at the lower and upper end of the achievement distribution, and ELs tend to be concentrated in the lower end; (2) inconsistent use of accommodations across time and location adds variation in the measures within teachers over time as well as between teachers; (3) the responsibility for educating an EL is often shared across a number of teachers (e.g., a classroom teacher and an ESL teacher), thus making it difficult to determine the precise contributions of each teacher to a student's growth; and (4) the influence ELs have on value-added estimates depends in part on the methodology used for calculating teacher value-added (Jones, Buzick, & Turkan, 2013; Lakin & Young, 2012).

Beyond accountability consequences, assessments have a direct impact on the education of individual ELs themselves. For example, assessments are typically used to determine (1) EL and English proficient status, (2) special education identification, and (3) academic track (remedial, grade level, honors, etc.) placement in core content area classes. Invalid or unreliable assessment results among ELs can jeopardize their appropriate and equitable placements in all three of these areas (Linan-Thompson, 2010).

Given the importance of assessments to schools, teachers, and individual students, we now focus on a principal method for improving assessment validity and reliability for ELs—accommodations. Testing accommodations come in a variety of forms, including extra time, bilingual dictionaries, and test translations. Different accommodations vary substantially both in terms of their effectiveness (Kieffer, Lesaux, Rivera, & Francis, 2009) and of the advantages

they might provide to ELs over other students (e.g., dictionaries; Abedi, Hofstetter, & Lord, 2004). Across the various forms of accommodations, research suggests that test translations may lead to the biggest improvements in validity on average (Kieffer et al., 2009; Robinson, 2010). However, there are notable limitations and obstacles with translations, including: (1) the language of instruction should match the language of the test, and thus translations may not be appropriate for ELs largely instructed in English (Hofstetter, 2003); (2) ensuring that a translated test assesses the same construct as the English version requires substantial time, effort, and resources (Hambleton, Merenda, & Spielberger, 2004); (3) some states do not permit the use of test translations on standardized exams (Rivera & Collum, 2014); and (4) the method of creating a translated test (e.g., back-translation, forward-translation) varies across states and tests, potentially affecting quality (Rivera & Collum, 2014).

One promising approach to accommodations incorporates computers in assessing the needs of individual ELs and in administering the appropriate accommodations. Using a computer-based algorithm to determine the most appropriate accommodation for each student, Kopriva and colleagues (2007) found that students who were provided with the algorithm-determined accommodation performed substantially better than if they were given a random accommodation. Technology may also help in other ways—for example, Abedi (2009) found that providing pop-up definitions (when hovering over selected words) to both ELs and non-ELs assessed via a computer can improve assessment validity for ELs while simultaneously not providing them with an unfair advantage.

Research-based Policy Recommendations Regarding Assessments

- For school accountability, consider using the category of "ever EL" instead of "currently (or recently) EL."

- Establish academic achievement expectations that take into account students' English proficiency and time in the school system.
- If using value-added methods for teacher or school accountability, consider the consequences of using invalid and unreliable assessments of ELs, and adjust accordingly, either by not including ELs' scores or by improving validity and reliability through accommodations.
- Enact policies that allow for a wide range of accommodations to be used because no single accommodation works for all students.
- Assess students to determine the most appropriate accommodation(s) and provide them as needed.

Conclusion

English learners' access to equitable education is affected by a wide-ranging set of education policies, including those governing the type of instruction they receive and the language of that instruction, their access to curriculum, the ways that they are assessed, and even when they are no longer considered ELs and no longer subject to these policies. These policies frequently vary across schools, districts, and states, and can change substantially over time, adding further instability to the educations of this habitually underserved group. In this article, we highlighted findings from rigorous research studies—often using experimental or quasi-experimental designs—to study the implications of these policies and practices for ensuring ELs have equitable educational opportunities and experiences. Our review suggests that many current policies and practices should be reconsidered in light of the research conclusions.

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. (2008). Measuring students' level of English proficiency: Educational significance and assessment requirements. *Educational Assessment*, *13*(2-3), 193-214.

Abedi, J. (2009). Computer testing as a form of accommodation for English language learners. *Educational Assessment*, *14*(3-4), 195-211.

Abedi, J., Hofstetter, C., & Lord, C. (2004). Assessment accommodations for English language learners: Implications for policy-based empirical research. *Review of Educational Research, 74*(1), 1–28.

Adesope, O.O., Lavin, T., Thompson, T., & Ungerleider, C. (2010). A systematic review and meta-analysis of the cognitive correlates of bilingualism. *Review of Educational Research, 80*(2), 207-245.

August, D., & Shanahan, T. (Eds.) (2006). *Developing literacy in second-language learners*. Mahwah, NJ: Erlbaum.

Barnett, W.S., Yarosz, D. J., Thomas, J., Jung, K., & Blanco, D. (2007). Two-way and monolingual English immersion in preschool education: An experimental comparison. *Early Childhood Research Quarterly*, *22*(3), 277-293.

Blanchard, S., & Muller, C. (2015). Gatekeepers of the American dream: How teachers' perceptions shape the academic outcomes of immigrant and language-minority students. *Social Science Research, 51*, 262-275.

Callahan, R. (2005). Tracking and high school English learners: Limiting opportunity to learn. *American Educational Research Journal, 42*(2).

Callahan, R. & Gándara, P. (2014). *The bilingual advantage: Language, literacy, and the labor market* (Eds.). Clevendon, UK: Multilingual Matters.

Callahan, R., Wilkinson, L., & Muller, C. (2010). Academic achievement and course taking among language minority youth in US schools: Effects of ESL placement. *Educational Evaluation and Policy Analysis, 32*(1), 84-117.

Craik, F.I.M., Bialystok, E., & Freedman, M. (2010). Delaying the onset of Alzheimer disease: Bilingualism as a form of cognitive reserve. *Neurology, 75*, 1726-1729.

Cummins, J. (2000). *Language, power, and pedagogy: Bilingual children in the crossfire.* Clevedon, UK: Multilingual Matters.

Dabach, D. (2014). I Am Not a Shelter!" Stigma and social boundaries in teachers' accounts of students' experience in separate "sheltered" English learner classrooms. *Journal of Education for Students Placed at Risk, 19*(2), 98-124.

Dabach, D. (2015). Teacher Placement Into Immigrant English Learner Classrooms: Limiting Access in Comprehensive High Schools. *American Educational Research Journal, 52*(2), 243-274.

Estrada, P. (2014). English learner curricular streams in four middle schools: Triage in the trenches. *The Urban Review, 46*(5).

Gándara, P., & Hopkins, M. (2010). *Forbidden language: English learners and restrictive language policies.* New York: Teachers College Press.

Gándara, P., Maxwell-Jolly, J., & Driscoll, A. (2005). Listening to Teachers of English language learners: A survey of California teachers. Policy Analysis for California Education.

Gándara, P., Rumberger, R., Maxwell-Jolly, J., & Callahan, R. (2003). English learners in California schools: Unequal resources, unequal outcomes. *Education Policy Analysis*

Archives, 11(36), 1-54.

García-Nevarez, A.G., Stafford, M.E., & Arias, B. (2005). Arizona elementary teachers' attitudes toward English language learners and the use of Spanish in classroom instruction. *Bilingual Research Journal, 29*(2), 295-317.

Greene, J.P. (1997). A meta-analysis of the Rossell & Baker review of bilingual education research. *Bilingual Research Journal, 21* (2/3).

Haertel, E.H., & Ho, A. (in press). Fairness using derived scores.

Hambleton, R.K., Merenda, P.F., & Spielberger, C.D. (Eds.). (2004). *Adapting educational and psychological tests for cross-cultural assessment*. Mahwah, NJ: Psychology Press.

Harklau, L. (1994). Tracking and linguistic minority students: Consequences of ability grouping for second language learners. *Linguistics and Education, 6*(3), 217-244.

Hofstetter, C.H. (2003). Contextual and mathematics accommodation test effects for English-language learners. *Applied Measurement in Education*, *16*(2), 159-188.

Hopkins, M., Lowenhaupt, R., & Sweet, T.M. (2015). Organizing English learner instruction in new immigrant destinations: District infrastructure and subject-specific school practice. *American Educational Research Journal, 52*(3), 408-439.

Hopkins, M., Thompson, K.D., Linquanti, R., August, D., & Hakuta, K. (2013). Fully accounting for English learner performance: A key issue in ESEA reauthorization. *Educational Researcher 42*(2): 101-108.

Jones, N.D., Buzick, H.M., & Turkan, S. (2013). Including students with disabilities and English learners in measures of educator effectiveness. *Educational Researcher*, *42*(4), 234-241.

Kanno, Y., & Kangas, S. (2014). "I'm not going to be, like, for the AP": English language learners' limited access to advanced college-preparatory courses in high school.

American Educational Research Journal, 51(5), 848-878.

Kieffer, M.J., Lesaux, N.K., Rivera, M., & Francis, D.J. (2009). Accommodations for English language learners taking large-scale assessments: A meta-analysis on effectiveness and validity. *Review of Educational Research*, *79*(3), 1168-1201.

Lau v. Nichols, No. 414 U.S. 563 (1974).

Lillie, K.E., Markos, A., Arias, M.B., & Wiley, T.G. (2012). Separate and not equal: The implementation of structured English immersion in Arizona's classrooms. *Teachers College Record, 114*(9), 1-33.

Linquanti, R. & Cook, H.G. (2015). Re-examining reclassification: Guidance from a national working session on policies and practices for exiting students from English learner status. Washington, DC: Council of Chief State School Officers.

Manning, R. (2014, July 11). Oregon educators prioritize bilingual education. Oregon Public Broadcasting.

National Research Council. (2011). *Allocating Federal Funds for State Programs for English Language Learners*. Washington, DC: The National Academies Press.

No Child Left Behind. (2002). *No Child Left Behind (NCLB) Act of 2001*, Pub. L. No. 107–110, § 115, Stat. 1425.

Oakes, J. (2005). *Keeping track: How schools structure inequality*. New Haven: Yale University Press.

Reardon, S.F., & Raudenbush, S.W. (2009). Assumptions of value-added models for estimating school effects. *Education Finance and Policy*, *4*(4), 492-519.

Rivera, C., & Collum, E. (2014). *State assessment policy and practice for English language learners: A national perspective*. New York: Routledge.

Rothstein, J. (2010). Teacher quality in educational production: tracking, decay, and student achievement. *Quarterly Journal of Economics*, *125*(1), 175-214.

Robinson, J.P. (2010). The effects of test translation on young English learners' mathematics performance. *Educational Researcher, 39*(8), 582-590.

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.

Robinson-Cimpian, J.P., & Thompson, K.D. (in press). The effects of changing test-based policies for reclassifying English learners. *Journal of Policy Analysis and Management.*

Robinson-Cimpian, J.P., Thompson, K.D., & Makowski, M.B. (revision under review). Evaluating English learner reclassification policy effects across districts. Centennial Issue of the *American Educational Research Journal*.

Rolstad, K., Mahoney, K., & Glass, G.V. (2005). The big picture: A meta-analysis of program effectiveness research on English language learners. *Educational Policy*, *19*(4), 572-594.

Ryan, C. (2013). *Language use in the United States: 2011.* Washington, DC: U.S. Census Bureau.

Saunders, W.M., & Marcelletti, D.J. (2013). The gap that can't go away: The catch-22 of reclassification in monitoring the progress of English learners. *Educational Evaluation and Policy Analysis*, *35*(2), 139-156.

Slavin, R.E., & Cheung, A. (2005). A synthesis of research on language of reading instruction. *Review of Educational Research, 75* (2), 247-284.

Slavin, R.E., Madden, N., Calderón, M., Chamberlain, A., & Hennessy, M. (2011). Reading and language outcomes of a multiyear randomized evaluation of transitional bilingual

education. *Educational Evaluation and Policy Analysis 33*(1), 47-58.

Solórzano, R.W. (2008). High stakes testing: Issues, implications, and remedies for English language learners. *Review of Educational Research*, *78*(2), 260-329.

Steele, J.L., Slater, R.O., Zamarro, G., Miller, T., Li, J., Burkhauser, S., & Bacon, M. (2015). The effect of dual-language immersion on student achievement: Evidence from lottery data. Working paper.

Thompson, K.D. (2015a). English learners' time to reclassification: An analysis. *Educational Policy*. Advance online publication. doi: 10.1177/0895904815598394

Thompson, K.D. (2015b). Questioning the long-term English learner label: How classification and categorization can blind us to students' abilities. *Teachers College Record*.

Thompson, K.D. (2015c). *What blocks the gate? Exploring current and former English learners' math course-taking in secondary school*. Paper presented at the American Educational Research Association, Chicago, IL.

Umansky, I.M. (2014). *Leveled and exclusionary tracking: English learners' access to core content in middle school*. Stanford University Center for Education Policy Analysis.

Umansky, I.M. (2013). *Peeling back the label: Do classifications and specialized services help or hurt language minority students?* Paper presented at the Segregation, Immigration, and Educational Inequality Conference, Ghent, Belgium.

Umansky, I.M. & Reardon, S.F. (2014). Reclassification patterns among Latino English learner students in bilingual, dual immersion, and English immersion classrooms. *American Educational Research Journal 51*(5), 879-912.

Umansky, I.M., Reardon, S., Hakuta, K., Thompson, K.D., Estrada, P., Hayes, K., … Goldenberg, C., (2015). Improving the opportunities and outcomes of California's

students learning English: Findings from school district-university collaborative partnerships. Sacramento, CA: Policy Analysis for California Education.

U.S. Department of Education, National Center for Education Statistics. (2015). *The Condition of education 2015* (NCES 2015-144), English Language Learners.

Harris, E.A. (2015, January 15). New York City education department to add or expand 40 dual-language programs. *The New York Times*. A24.

Valentino, R.A., & Reardon, S.F. (2015). Effectiveness of four instructional programs designed to serve English learners: Variation by ethnicity and initial English proficiency. *Educational Evaluation and Policy Analysis*.

Willig, A. (1985). A meta-analysis of selected studies on the effectiveness of bilingual education. *Review of Educational Research, 55*(3), 269-317.