

Latino Education
An Agenda for Community
Action Research

*A Volume of the National Latino/a Education
Research and Policy Project*

Edited by

Pedro Pedraza

Melissa Rivera

*Center for Puerto Rican Studies, Hunter College,
City University of New York*



LAWRENCE ERLBAUM ASSOCIATES, PUBLISHERS
Mahwah, New Jersey London

Senior Acquisitions Editor: Naomi Silverman
Assistant Editor: Erica Kica
Cover Design: Kathryn Houghtaling Lacey
Textbook Production Manager: Paul Smolenski
Full-Service Composer: TechBooks
Text and Cover Printer: Hamilton Printing Company

This book was typeset in 10.5/13 pt. Palatino Roman, Bold, and Italic.
The heads were typeset in Palatino, Palatino Bold, and Palatino Bold Italic.

Copyright © 2005 by Lawrence Erlbaum Associates, Inc.
All rights reserved. No part of this book may be reproduced in
any form, by photostat, microform, retrieval system, or any
other means, without prior written permission of the publisher.

Lawrence Erlbaum Associates, Inc., Publishers
10 Industrial Avenue
Mahwah, New Jersey 07430
www.erlbaum.com

Library of Congress Cataloging-in-Publication Data

Latino education : an agenda for community action research / edited by Pedro
Pedraza, Melissa Rivera.

p. cm.

"A Volume of the National Latino/a Education Research and Policy Project."
Includes bibliographical references and index.

ISBN 0-8058-4986-6 (casebound)—ISBN 0-8058-4987-4 (pbk.)

1. Hispanic Americans—Education. 2. Action research in
education—United States. 3. Educational equalization—United States.
4. Educational change—United States. I. Pedraza, Pedro, 1946–
II. Rivera, Melissa, 1971– III. National Latino/a Education Research and Policy
Project. IV. Title.

LC2669.L36 2005
371.829'68073—dc22

2005012014

Books published by Lawrence Erlbaum Associates are printed on
acid-free paper, and their bindings are chosen for strength and
durability.

Printed in the United States of America
10 9 8 7 6 5 4 3 2 1

This volume is dedicated to Dr. Enrique (Henry) Torres Trueba
for the countless hours you gave to those who needed your
help, for the compassionate manner and humane mentorship
you offered, for the humor and joy you brought to gatherings,
and for the *cariño* and *respeto* you shared with all who crossed
your path.

Student Learning and Assessment: Setting an Agenda

Robert Rueda

University of Southern California

INTRODUCTION

Hardly a day goes by when a public figure does not call out for greater accountability for public schools. Office-holders and office-seekers, from both inside and outside of education, are likely to mention the improvement of public education as part of their agenda. Most recently, this concern with education and improving public schools, the education reform agenda, has almost become synonymous with large-scale and high-stakes assessment especially as embodied in the No Child Left Behind Act of 2001 (H.R.1). Significantly, one of the four "pillars" of this national education reform blueprint is accountability and testing (the others include flexibility and local control, funding for what works, and expanded parental options). This legislation, in which accountability plays a major role, represents a comprehensive overhaul of existing federal law (Elementary and Secondary Education Act enacted in 1965 and the Improving America's Schools Act of 1995) and is the principal federal law affecting K-12 education today. Among the accountability

features of the new legislation are stringent requirements for states including:

1. States must develop "adequate yearly progress" (AYP) statewide measurable objectives for improved achievement for all students and for specific groups: economically disadvantaged students, students from major racial and ethnic groups, students with disabilities, and students with limited English proficiency
2. The objectives must be set with the goal of having all students at the "proficient" level or above within 12 years (i.e., by the end of the 2013–2014 school year)
3. AYP must be based primarily on state assessments but also must include an additional academic indicator
4. The AYP objectives must be assessed at the school level. At the end of 2 years, schools that have failed to meet their AY objective for 2 consecutive years will be identified for improvement
5. School AYP results must be reported separately for each group of students identified above so that it can be determined whether each group of students met the AYP objective
6. At least 95% of each group must participate in state assessments

A related stipulation of this legislation is that states establish performance standards for their tests. As Linn, Baker, and Betebenner (2002) note, some states set high standards not knowing that these would be used to determine AYP objectives. Moreover, virtually no state is close to meeting the goals being set, and in many cases it is not realistic to expect that states will be able to meet the standards. Although this is problematic for all states, it is especially troublesome for states with large numbers of Latino/a students where finding appropriate assessments is challenging and where the mechanisms for impacting student progress are not well conceived or implemented.

Arguments against accountability are tantamount to speaking out against motherhood or the flag. Yet, although this focus on improving schools and learning is commendable, it does have critical ramifications for Latino/a students. One, for example, is the use of English-only standardized assessments where significant consequences are attached either to students themselves, their teachers, their schools, or their districts. Although recent syntheses on teaching and learning have

suggested the importance of culture and language in the learning process, for the most part this has not translated into the current debates or implementation of assessment (American Psychological Association Board of Educational Affairs, 1995; Bransford, Brown, & Cocking, 1999; Donovan, Bransford, & Pellegrino, 1999; Lambert & McCombs, 1998). Another problematic issue in the current move toward accountability is holding students accountable to standards without also taking into account opportunity to learn (Clare Matsumura, Garnier, Pascal, & Valdes, 2002) or providing mechanisms to address low performance other than economic or other sanctions.

Although the use of high-stakes large-scale assessment for school accountability is not new, the current emphasis on their widespread use is. This increasing emphasis is arguably the most visible manifestation of assessment concerns that have both potential and actual ramifications for Latino/a students. However, there are other areas that, although less visible in current public discussions, have particular importance for Latino/a students, especially those who are English learners. One area has a longer history, yet still continues to be an area of concern: assessment for special education placement. Another area is the use of testing and assessment processes related to entrance into postsecondary educational institutions. This chapter will examine the issues surrounding each of these areas, and try to provide suggestions related to formulating a research agenda that addresses them. First a bit of background on each of these areas will be provided, outlining the issues, and then a discussion of how these translate into areas of concern for a research agenda will be provided. The purpose here is not to provide a definitive review of each area, but to provide a brief review of key issues in each domain and how they hold particular importance for Latinos.

LARGE-SCALE HIGH-STAKES ACCOUNTABILITY MEASURES

One important component of the effort to reform schools is the use of rigorous learning standards for all students (Lachat, 1994). The basic idea is that standards should be public, clear, and hold high expectations for students, serving as a key tool to improve schools and student outcomes (National Council on Education Standards and Testing, 1992; National Education Goals Panel, 1993). A majority of states are tying state-mandated testing programs to new or revised state curriculum standards. Moreover, 48 states have statewide assessment programs,

and almost all report making revisions in their state tests to align them more closely with specific state standards (Education Week, 2000).

The new federal education legislation just passed also includes new requirements for states and school districts in terms of demonstrating student achievement. Annual testing based on state standards in reading and mathematics will be required for all students in grades 3 through 8. Increasingly, these test results are tied to significant consequences. As of 1996, for example, at least 23 states reported connecting explicit consequences to state test results, such as funding cuts, warnings, mandatory assistance from outside experts, loss of accreditation, or state takeover (Bond, Roeber, & Braskamp, 1996). Similarly, the new federal education legislation permits parents to transfer students in failing schools to successful public or charter schools.

In the best of circumstances, these standards should serve as tools of equity and excellence, assuring that every stakeholder concerned with the educational process is working toward the same worthwhile goals while serving as a guide for what should be taught, how progress should be monitored, and what should be assessed. Under these ideal conditions, the tracking and placement of certain groups of students such as poor or non-English speaking students would decrease or disappear. However, there are many problems, both real and potential, that serve to diminish the promise of these otherwise commendable goals. Urban schools, that have disproportionate numbers of students from low SES and diverse cultural and linguistic backgrounds, have disproportionately felt the negative impacts of testing policies (Darling-Hammond, 1994).

As one example, Buly and Valencia (2002) conducted a study in Washington State that looked at students who failed the state reading assessment. She found that whereas English Language Learners represented 11% of the population taking the test, they represented 43% of those students who failed the test. Likewise, although low SES students represented 47% of those taking the test, they represented 67% of those who failed. More important, she did an analysis of the patterns of performance of those who failed, and was able to distinguish clusters of students based on error profiles. For example, "automatic word callers" were able to identify words fluently, but did not understand what they read, whereas "slow comprehenders" could identify words and understood them but were not fluent. As Valencia noted, each of the various subgroups identified would require different instructional interventions to perform better. Valencia concluded that scores on state assessments mask important group and individual differences,

and may be especially problematic with poor and non-English proficient students. Normally, only summative scores are attended to for accountability purposes, while the types of detailed analyses more useful for instructional purposes are ignored.

Some of the criticisms and problems are not directly focused on Latino/a students per se but affect them nevertheless. For example some have argued that many state-mandated tests emphasize obscure knowledge of isolated facts or vocabulary, which then has the effect of forcing teachers to narrow the curriculum and deemphasize higher-order critical thinking skills (McNeil, 2000; Yeh, 2001). Yet as Darling-Hammond has noted, this is particularly apt to take place in urban schools populated by Latino/a and other students of diverse backgrounds (Darling-Hammond, 1995).

Others have noted that tests should not be used to gauge innate ability or make inferences about potential academic success under conditions of unequal opportunity to learn. Many poor Latino/a students and other students of color are exposed to less challenging curricula, have less access to resources, less high-quality instruction, and have a higher likelihood of being in unsafe or unsupportive school environments (Darling-Hammond, 1994; Nettles & Nettles, 1995).

Other areas of concern related to standards and high-stakes tests include the following:

- Standards may not include or address the needs of all students such as English learners
- In the development of tests and assessment indicators, provisions may not be made for English learners or students from diverse cultural backgrounds, so that the tests become de facto language test or test of cultural familiarity
- There may be no alignment between state or district standards and assessment methods, so that assessments do not provide a true picture of what was learned. In high-stakes situations, the consequences for low performance may be significant. In this context, high-stakes situations are those educational contexts where student performance is related to long-term, real-life consequences such as placement in special education, graduation from high school, college entrance, and so on.
- Assessments may be too narrowly focused and restrict curriculum, especially for groups that traditionally score low
- Test results are increasingly tied to rewards and punishments, resulting in a situation where some students or groups (traditional

“low achievers”) are less valued—one potential outcome is that the neediest schools are staffed by the least qualified teachers

- Assessments may be used only for accountability purposes, but not used to improve instruction or in the reallocation of resources
- Political considerations become the driving force in the development and implementation of standards and assessment. (Smith, Heinecke, & Noble, 1999).

It has been convincingly argued that most current assessments are not designed to capture the full range of what English language learners know and much of what they are able to do (LaCelle-Peterson & Rivera, 1994). There have been various attempts to deal with this shortcoming. One approach, the default position, has been to simply exclude students who do not speak English from tests. For those interested in accountability issues and monitoring students' performance over time, however, this is not satisfactory, as it excludes these students from the same expectations and challenges of high standards of other students. As an example, one study in 1994 found that of the 48 states responding, 44 permitted exemptions for English language learners. The reasons varied from recent arrival status (usually from 1 to 3 years), teacher recommendation, and even participation in an ESL program, although this situation has begun to change (Olson & Goldstein, 1997).

Another approach has been to use test accommodations, or modifications in test administration procedures. Butler and Stevens described a taxonomy of test accommodations that fall into two categories: modifications of the test (use of native language, text change in vocabulary, change in linguistic complexity, use of visual supports, glossaries, dictionaries) and modifications of the procedure (extra time, breaks, administration over several sessions, directions given in native language, and so on) (Butler & Stevens, 1997b). They further point out that the use of accommodations is complex and argue the need to match assessment accommodations to the specific characteristics of individual students or subpopulations. Although this body of work is not large, it does suggest caution against blanket statements about the general effectiveness or lack of effectiveness of a particular form of accommodation for any individual or subgroup (Abedi, Lord, & Hofstetter, 1998).

Still another approach is the use of the native language in assessment, either through the development of parallel assessments or translating existing English assessments. Some critics have pointed out that the issue of equivalency of item difficulty is problematic, and this approach

may interact differentially with the language in which the child has been instructed (Figueroa, 1990). One complication is that English-language learner status is not a single entity (LaCelle-Peterson & Rivera, 1994; Valdes & Figueroa, 1994).

Butler and Stevens proposed a comprehensive framework for unpacking some of the variability surrounding English learners, including variables related to the home, community, school, educational background, language factors, and personal characteristics (Butler & Stevens, 1997a). In addition, they stressed the need to consider three critical background variables that impact performance and should be considered in thinking about accommodations: academic English language proficiency (see Bailey, in press, for an excellent discussion of this issue), prior education, and length of time in the United States. At present, there is a lack of information related to high stakes assessment with Latino/a students and this is clearly an important priority given the importance that such assessment has come to play in the national education agenda. Current standardized assessments are heavily loaded with academic English language proficiency but ignore the other dimensions noted by Butler and Stevens.

SPECIAL EDUCATION ASSESSMENT

One consequence of low academic achievement and low test scores, especially at the early grades, is referral for special education placement or remedial programs. There are a whole set of assessment issues that accompany this area, the most prominent being the role of ability testing and the problem of overrepresentation in certain categories for certain subgroups of the general population, in particular African American and Latino/a students (for a comprehensive review of the assessment issues involved, especially involving the use of IQ tests, see Valencia & Suzuki, 2001). This disproportional representation issue will be briefly discussed later.

The issue of disproportionate representation of minority students in special education has haunted the education field for almost 35 years. An early study by Mercer (1973) documented disproportionate representation in classes for students with mild retardation (educable mentally retarded [EMR]) and raised other issues regarding biased assessment and placement practices. These practices included, for example, inappropriate testing in English of students not proficient in

the language, use of on-the-spot translations, and inappropriate norms that did not include representative samples.

A short time later the National Academy of Sciences convened a panel on this issue of overrepresentation of minority students in special education. The ensuing report reaffirmed earlier placement patterns and criticisms of existing assessment and placement practices (Heller, Holtzman, & Messick, 1982). As part of the panel's work, a review of the available Office for Civil Rights (OCR) data found that minority students were overrepresented in classes for the mentally retarded (EMR and trainable mentally retarded [TMR]) as well as in classes for the emotionally disturbed (Finn, 1982). These analyses documented effects for district size and size of minority enrollment, and disproportionate representation rates were found to be related to the presence or absence of bilingual programs. A major recommendation of the panel was to deemphasize individual differences in achievement and to focus more on assessing the validity of instructional settings. Importantly, the panel also recommended greater attention to assessing problems before a child accumulates a long history of failure and before formal testing for special education diagnosis.

There have been many changes since these early reports, but in many ways things have not changed much at all. There have been significant legal battles fought over these issues; there have been significant changes to education law; there have been significant changes in demographics in many school districts and states; there have been significant educational initiatives such as the school reform movement and the current emphasis on reading (Rueda, Artiles, Salazar, & Higareda, 2002).

Unfortunately, there are indications that disproportionate representation, especially in the "judgmental" categories of special education, remains a problem.¹ Because of the continuing nature of this problem, more recently a second panel from the National Academy of Science was convened on this same issue (Donovan & Cross, 2002). This just-released report found that at the national level, whereas 5% of Asian students are identified for special education, the rate for Hispanics is

¹Generally the disability categories involved include mild mental retardation, emotional/behavioral disorders, and specific learning disabilities. Usually, the defining characteristic of students with these labels includes low academic performance and/or atypical social behavior. These categories are called "judgmental" because students appear normal in most arenas outside of classroom and test situations in contrast to more severe disabilities where students manifest identifiable characteristics such as with Down syndrome, autism, deafness, and so on.

11%, for Whites, 12%, for American Indians 13%, and for Blacks over 14%. More specifically, Black students are overrepresented in the category of mental retardation and emotional disturbance, and underrepresented in gifted and talented programs, and

- the greatest disproportion is in MR, where Black children are more than twice as likely as all other students to be assigned. About 2.6% of all Black children are identified as MR.
- about 1.5% of Black students are labeled ED—about twice the rate for all other students.
- although about 6% of all students are labeled gifted and talented, only 3% of Black students are so labeled.

The report also found that while Hispanic students were not overrepresented in the categories of mental retardation or emotional disturbance, they were underrepresented in gifted and talented programs. More specifically, only about 3.5% of Hispanic students are placed in gifted programs, about half the rate for White students.

American Indian/Alaskan Native students were overrepresented in the learning disabilities category—(White, Hispanic, and Black student placement rates hover between 6 and 6.5%, whereas American Indians are at almost 7.5%) and underrepresented in gifted and talented programs, with just under 5% placed. In contrast, Asian students are about half as likely as other students to be labeled as students with mental retardation, and about a third as likely to be labeled as students with emotional disturbance or learning disabilities. Yet almost 10% of Asian students are placed in gifted and talented programs; this is three times the rate for Blacks and Hispanics, and a 30% higher rate than for White students.

Like its predecessor, this report recommended that assessment focus on features of the context rather than simply the individual, including prior educational history, quality of instructional environments, access to resources, teacher quality, and so on. In addition, the report suggested much greater emphasis on early intervention, trying to identify problems way before a formal referral is necessary, and providing ample additional instructional help, especially in early reading. Significantly, the report went on to recommend that special education eligibility should be determined on the basis of resistance to high-quality interventions rather than to performance on arbitrary tests. Specifically, it recommended that an IQ-achievement discrepancy

should be eliminated as a requirement for determining that a child has a learning disability, thus eliminating one of the major uses of IQ tests in the eligibility determination process, and proposed noncategorical classifications that focus on matching a student's specific needs to an intervention strategy, thus removing the need for traditional labels.

Although on the surface these developments may seem positive, there are still significant imbalances in special education placements, as the national data indicate. Although the specific patterns and categories may have changed over time, the major issue is that the consistency and systematic nature of disproportionate representation is reason for concern and these are not random patterns.

Somewhat more troubling is that there are some data that suggest that when more local analyses are done, the problem of disproportionate representation may be more significant than the nationally aggregated data would suggest (Oswald, Coutinho, Best, & Singh, 1999). For example, the size of a school district and of the special education programs as well as the size of the representation of the ethnic group in the district can also shape the problem (Artiles & Trent, 2000; Reschly, 1997). The need for disaggregated data is evidence that one potential result of English-only mandates may be related to increased rates of special education placement for English learner and immigrant students (Artiles, Rueda, Salazar, & Higareda, 2000).

Although the call to focus on intervention (especially reading) rather than just type of special education placement or label is commendable, little of the intervention research in the area of reading has "unpacked" variables such as language proficiency—little is known about reading acquisition for ELL students (Garcia, 2000). The call to focus on reading instruction as an answer may ignore the critical distinction between reading and literacy, focusing on reading without equal concern for the wider issue of literacy (Rueda & McIntyre, 2002). As is noted in this chapter, while the focus on instruction and intervention is a good one, other developments threaten to become problematic issues, including the standardization of instruction and materials, and universality of approaches—all three of these issues translate quite easily into problems which have major implications for the domain of assessment. Although some have pointed to alternative assessment procedures such as performance-based assessment as a potential solution to some of these long-standing issues, some have wisely cautioned that alternative procedures are not inherently equitable, and caution must be taken regarding how and why they are used (Darling-Hammond, 1996).

TESTING AS A GATEKEEPER FOR HIGHER EDUCATION

The effects of problematic assessment procedures do not stop in the early elementary years. In fact, the issues related to tests and assessment may have a significant impact on student careers much further down the academic pipeline. For example, although Latino/a students have made some gains in postsecondary education over the past half century, helped in part by tools such as affirmative action, there are still wide discrepancies among various groups in terms of access to higher education (Garcia, 2001). Just over one million Hispanic students were enrolled in institutions of higher education in 1998, twice as many as a decade earlier (National Center for Educational Statistics, 1998). Still, Latino/a students are half as likely to graduate from college as White students (National Center for Educational Statistics, 1995). Whereas a variety of factors have been examined with respect to this problem, assessment (especially entrance exams such as the SAT taken during students' high school years) has been a key factor. These high-stakes tests are normally taken during the later years of students' K-12 school careers. Entrance exams have become increasingly important with the withering support for affirmative action programs and policies (Bowen & Bok, 1998). Specifically, they play an increasingly large gatekeeping role, particularly in key states such as California and Texas.

Although a complete treatment of the complex issues surrounding affirmative action and the role of standardized entrance exams in relation to access to higher education is beyond the scope of this chapter, a recent study is informative (see Garcia, 2001, for a more comprehensive review of the issues). The study in question was conducted in California, a state that is heavily Latino/a and also the location of the first major rollback of affirmative action in postsecondary education. This was spurred by the passage of Proposition 209 in the state, which eliminated the consideration of race, ethnicity, and gender in public employment, public contracting, and education. The University of California system approved policies consistent with this mandate that went into effect in the spring quarter of 1997-1998.

Koretz, Russell, Shin, Horn, and Shasby recently conducted a study that focused on this major university system to explore how eliminating affirmative action in college admissions affects the diversity of admitted students (Koretz, Russell, Shin, Horn, & Shasby, 2002). The study sought to compare the effects of changing admissions from a period

before the affirmative action ban to the race-neutral admissions policy later put into force. Although the study was complex and addressed a number of related issues, one finding was that admissions processes that were race-neutral and relied solely on SAT and GPA had major effects at the most selective campuses in the University of California system, although the effects were smaller at the less selective campuses. Importantly, none of the alternative admissions models that were analyzed in the study could replicate the composition of the student population that existed in the system prior to the elimination of affirmative action.

As access to higher education becomes increasingly test-dependent, it is critical that considerations of fairness and equity be taken into account and that close monitoring of both access to and graduation from quality institutions of higher education be monitored. Whereas concerns about fairness and equity are commonly cited when considering high-stakes tests for students in the early grades, it is equally important at the other end of the educational continuum. Especially critical is the secondary school period of a student's career, because course-taking, counseling, test preparation, and other factors are related to later outcomes.

CONCLUSION

Although the current focus on education is a positive development, it has often led to a desire for quick fixes that are unrealistic and without promise for meaningful change. With respect to low achievement, for example, some common responses have included

- a market-driven approach with an emphasis on privatization;
- standardization of (and often mandated) curricula, instructional materials, teaching approaches;
- retention;
- remedial programs, either supplementing or replacing ongoing programs;
- accountability, often operationalized by single measure high-stakes standardized tests.

Accountability is one of the key elements of the recently passed federal blueprint for school reform. Yet, the time frame producing results

is increasingly short and the consequences are increasingly punitive. This is of great concern for students who are English learners.

The No Child Left Behind Act is remarkable in its single-minded call that limited English proficient students receive English-only instruction and in the failure to promote bilingual education or primary language development as equally viable instructional approaches to support the development of English language and content skills with limited English proficient students. If students are to be held accountable to challenging standards, they must be given appropriate opportunity to master the standards.

Recently, researchers have begun to call for balance in the context of instruction, particularly in the areas of reading and literacy. Policies that mandate a single approach to instruction have caused concern in this respect. For example, Pearson and Raphael have called for balance in the following areas (Pearson & Raphael, 1999):

- Authenticity—the balance between doing school-like activities versus those with some connection to everyday life
- Classroom Discourse—the balance between adult versus child control over topics and turn taking
- Teachers' Roles—the balance between the amount of teacher control over student activity on a continuum from explicit instruction to co-participating
- Curricular Control—the balance between local control and external control over what is taught

The discussion in this chapter of some of the assessment issues of importance to Latinos suggests that a similar call for balance can be applied to assessment. What are some of the dimensions of assessment that might need to be considered? A preliminary list of such considerations is presented here.

Achievement and Ability Versus Opportunity to Learn

There is some controversy about the distinction between achievement and ability tests. In general, ability tests give information about what one has learned measured at this point in time, whereas achievement tests try to assess the underlying psychological aptitudes that are considered to be more stable and predictive of performance in the future.

An example of the latter would be IQ tests. However, most large-scale standardized tests used for accountability purposes are usually ability tests. Although these may be useful for comparative purposes, that is, to differentiate one student from another along one or more dimensions, they typically do not provide information about opportunity to learn. Opportunity to learn can be defined as a range of variables likely to influence student performance, including access to resources, access to high-quality instructional content and processes, extra-school opportunities, and direct preparation for the test being administered (Herman, Wakai, Wakai, & Heath, 1997). It is increasingly clear that without information on opportunity to learn, it is difficult to interpret the performance of students of diverse backgrounds and make comparisons. An important area to consider in the future will be taking opportunity to learn issues into account in the design and interpretation of assessment instruments and also addressing opportunity to learn as a worthwhile issue in its own right.

Reading Versus Literacy

Consistent with a recent major report on reading from the National Research Council (Snow, Burns, & Griffin, 1998), this chapter points to the need to differentiate reading and literacy. The former focuses on the individual psychological processes involved in decoding and comprehending text, or "... the use of the products and principles of the writing system to get at the meaning of a written text" (Snow et al., 1998, p. 42). Literacy, however, includes reading, but also focuses more broadly not only on the act of reading but on the beliefs, attitudes, and social practices that literate individuals and social groups engage in a variety of settings and situations, including those involving technology (Pearson & Raphael, 1999). Although the underlying individual cognitive processes underlying decoding text may be more universal in nature, the social practices underlying various types of literacy are more culturally bound. It is the combination of both that defines the adult who uses the tools of reading and writing in sophisticated ways. With respect to assessment, it is important to assess both dimensions to assure that not only basic skills are mastered, but that the actual uses of literacy in school and beyond are appropriately developed. The complex social and cognitive processes underlying the uses of literacy in everyday life are just as critical as basic skills and facts, and future assessments should focus on both dimensions.

Participation Versus Accuracy

There is currently much debate over how to treat students of diverse abilities and languages on large-scale high-stakes assessments. As noted earlier in this chapter, early approaches followed the strategy of simply excluding particular groups of students such as non-English speakers. However, because this strategy also reduces the degree to which such students can be monitored in terms of educational outcomes, more focus has shifted to accommodations in testing. Researchers at the Center for Research on Educational Standards and Student Testing (CRESST) at the University of California at Los Angeles have done especially useful work in this regard. However, as some (Shepard, 2001) have noted, there is a great deal of inconsistency in terms of how different states deal with these accommodations. Moreover, as Shepard (2001) notes, there is some question regarding this work in terms of the balance between participation and accuracy. That is, whereas it is important that particular subgroups of students not be excluded (and, thus, schools not be held responsible for their outcomes), there is also a need that assessments provide accurate information. For students who do not speak the language, for example, the assessment becomes a de facto language assessment in spite of what the test purports to measure. This is an issue that will have to be closely monitored as accountability measures become attached to higher and higher consequences.

Accountability Versus Instruction

Whereas the thrust of new federal legislation and various state initiatives focus on accountability issues, there is a trade-off in terms of a focus on assessment as a tool to inform teaching. Large-scale accountability measures are not designed to directly inform day-to-day instruction, and they are not used by teachers in this way (except where they drive teach-to-the-test practices). However, good instruction is informed by constant assessment and monitoring which is used to modify instruction. Teachers need to be supported with professional development activities that inform them of new developments in classroom-based techniques such as performance assessments and other means to improve their teaching. In the rush to focus on large-scale accountability measures, budgets need to be examined to assure that teachers are supported with assessment more closely linked to instructional activities.

Single Versus Multiple Indicators

Whereas public discourse regarding accountability primarily focuses on accountability *measures*, assessment specialists are increasingly talking about accountability *systems*. The idea is that because of various sources of error, no single measure is a good index of performance. Rather, an integrated system of related measures is more likely to provide a more accurate and complete account of performance. This is especially critical for Latino/a populations, where measures are much more likely to occur because of the interplay of language, culture, and socioeconomic factors with specific test characteristics. Especially critical will be the means to hold the accountability systems *themselves* accountable. Baker, Linn, Herman, and Koretz (2002) have developed a comprehensive set of Standards for Educational Accountability Systems that address five domains, including system components, testing standards, stakes, public reporting formats, and evaluation. Among the selected items of concern to this chapter are the following:

- Accountability expectations should be made public and understandable to all participants in the system
- Accountability systems should employ different types of data from multiple sources
- Accountability systems should include the performance of all students, including subgroups that historically have been difficult to assess
- Decisions about individual students should not be made on the basis of a single test
- If tests are to help improve system performance, there should be information provided to document that test results are modifiable by quality instruction and student effort
- If test data are used as a basis of rewards or sanctions, evidence of technical quality of the measures and error rates associated with misclassification of individuals or institutions should be published
- Evidence of test validity for students with different language backgrounds should be made publicly available

What are some previously ignored dimensions that require attention in the future? Past assessment has tended to focus assessments

entirely on the individual student at a single point in time. There is a fundamental rule in assessment that is often overlooked—when students are being compared and inferences made about ability or potential, *comparisons are only valid when they have had equal opportunity to learn the material and acquire the skills covered in the test*. What other aspects are relevant to assessment if we broaden our focus beyond the individual? The following dimensions remain largely unexamined in current assessment work.

Opportunity to Learn

What is the nature of present and past instructional environments and social mediation have students had access to? What is the quality, type, amount, and range of school-like and non-school-like literacy activities and artifacts students have had access to?

Sociocultural Context

What is the social organization of learning settings and features of important social contexts in students lives, including materials, activities, and people?

Sociopolitical Considerations

What are the local dynamics in terms of past and current power (economic, political) relationships among various groups in specific communities and schools? How are these embedded in important social institutions such as schools and classrooms? How are these perceived and experienced by individual students and their families? How do these ultimately impact classroom relationships, attitudes toward schooling, and ultimately school achievement?

Whereas tests and assessment systems often have been narrowly construed and have served as roadblocks and negative factors in the educational and life outcomes of many Latino/a students, it does not have to be that way. When assessment systems are well conceptualized and comprehensively designed, and administered and used in appropriate ways, they can be tools used as powerful tools for accountability as well as instructional purposes. However, there are also many reasons for concern, as noted in the chapter, so that past abuses are not

repeated. The factors and issues outlined in this chapter can form the foundation for an agenda for future research and development.

REFERENCES

- Abedi, J., Lord, C., & Hofstetter, C. (1998). *Impact of selected background variables on students' NAEP math performance* (Center for Research on Evaluation, Standards, and Student Testing). Los Angeles: University of California, Los Angeles.
- American Psychological Association Board of Educational Affairs. (1995, December). *Learner-centered psychological principles: A framework for school redesign and reform*. Washington, DC: American Psychological Association.
- Artiles, A., Rueda, R., Salazar, J., & Higareda, I. (2000, November). *Factors associated with English learner representation in special education*. Presented at the Minority Issues in Special Education Conference. Boston, MA: Law School, Harvard University.
- Artiles, A., & Trent, S. (2000). Representation of culturally/linguistically diverse students. In C. Reynolds & E. Fletcher-Jantzen (Eds.), *Encyclopedia of special education, Vol. 1 (2nd ed.)* (pp. 513–517). New York: John Wiley & Sons.
- Bailey, A. L. (in press). From *Lambie* to *Lambaste*: The conceptualization, operationalization and use of academic language in the assessment of ELL students. In K. Rolstad (Ed.), *Rethinking school language*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Baker, E. L., Linn, R. L., Herman, J. L., & Koretz, D. (2002, Winter). Standards for educational accountability systems. *The CRESST Line*, 1–4.
- Bond, L., Roeber, E., & Braskamp, D. (1996). *Trends in state student assessment programs fall 1996*. Washington, DC: Council of Chief State School Officers.
- Bowen, W., & Bok, D. (1998). *The shape of the river: Long term consequences of considering race in college and university admissions*. Princeton, NJ: Princeton University Press.
- Bransford, J., Brown, A., & Cocking, R. (1999). *How people learn: Brain, mind, experience, and school*. Washington, DC: National Academy Press.
- Buly, M. R., & Valencia, S. (2002). Below the bar: Profiles of students who fail state reading assessments. *Educational Evaluation & Policy Analysis*, 24(3), 219–239.
- Butler, F., & Stevens, R. (1997a). *Considerations in addressing large-scale assessment issues for language minority populations* (National Center for Research on Evaluation, Standards, and Student Testing (CRESST), Center for the Study of Evaluation (CSE), Graduate School of Education and Information Studies.). Los Angeles: University of California, Los Angeles.
- Butler, F., & Stevens, S. (1997b). *Accommodation strategies for English language learners on large-scale assessments: Student characteristics and other considerations* (Center for the Study of Evaluation (CRESST), National Center for Research on Evaluation, Standards, and Student Testing, Graduate School of Education). Los Angeles: University of California, Los Angeles.
- Clare Matsumura, L., Garnier, H. E., Pascal, J., & Valdes, R. (2002, November). *Measuring instructional quality in accountability systems: Classroom assignments and school achievement*. CSE Technical Report 582. Los Angeles: Center for the Study of Evaluation and National Center for the Study of Evaluation, Standards, and Student Testing.
- Darling-Hammond, L. (1994). Performance-based assessment and educational equity. *Harvard Educational Review*, 64, 5–30.
- Darling-Hammond, L. (1995). Inequality and access to knowledge. In J. Banks & C. Banks (Eds.), *Handbook of research on multicultural education* (pp. 465–483). New York: Macmillan.
- Darling-Hammond, L. (1996). Performance-based assessment and educational equity. In E. Hollins (Ed.), *Transforming curriculum for a culturally diverse society* (pp. 245–272). Mahwah, NJ: Lawrence Erlbaum Associates.
- Donovan, M., Bransford, J., & Pellegrino, J. (1999). *How people learn: Bridging research and practice*. Washington, DC: National Academy Press.
- Donovan, M., & Cross, C. (2002). *Minority students in special and gifted education*. Washington, DC: National Academy Press.
- Education Week. (2000). *Quality counts 2000: Who should teach?* Bethesda, MD: Editorial Products in Education, Inc.
- Figueroa, R. (1990). Assessment of linguistic minority group children. In C. Reynolds & R. Kamphaus (Eds.), *Handbook of psychological and educational assessment of children: Intelligence and achievement*. New York: Guilford Press.
- Finn, J. (1982). Patterns in special education placement as revealed by the OCR surveys. In K. Heller, W. Holtzman, & S. Mesick (Eds.), *Placing children in special education: A strategy for equity* (pp. 322–381). Washington, DC: National Academy Press.
- Garcia, E. (2001). *Hispanic education in the United States: Raices y Alas*. Lanham, MD: Rowman & Littlefield.
- Garcia, G. (2000). Bilingual children's reading, Vol 3. In M. Kamil, P. Mosenthal, P. Pearson, & R. Barr (Eds.), *Handbook of reading research* (pp. 813–834). Mahwah, NJ: Lawrence Erlbaum Associates.
- Heller, K., Holtzman, W., & Messick, S. (1982). *Placing children in special education: A strategy for equity*. Washington, DC: National Academy Press.
- Herman, J. T., Wakai, D. C. D., Wakai, S. T., & Heath, T. (1997). *Assessing equity in alternative assessment: An illustration of opportunity-to-learn issues*. Technical Report No. 440. Los Angeles: UCLA, Center for the Study of Evaluation.
- Koretz, D., Russell, M., Shin, C., Horn, C., & Shasby, K. (2002). Testing and diversity in postsecondary education: The case of California. *Education Policy Analysis Archives*, 10(1), 1–34.
- LaCelle-Peterson, M., & Rivera, C. (1994). Is it real for all kids? A framework for equitable assessment policies for English language learners. *Harvard Educational Review*, 64, 55–75.
- Lachat, M. (1994). *High standards for all students: Opportunities and challenges*. South Hampton, NH: Center for Resource Management.
- Lambert, N., & McCombs, B. (1998). *How students learn: Reforming schools through learner-centered education*. Washington, DC: American Psychological Association.
- Linn, R. L. (2000). Assessments and accountability. *Educational Researcher*, 29(2), 4–16.
- Linn, R. L., Baker, E. L., & Betebenner, D. W. (2002). *Accountability systems: Implications of requirements of the No Child Left Behind Act of 2001*. CSE Technical Report No. 567. Los Angeles, CA: Center for the Study of Evaluation and National Center on Evaluation, Standards, and Student Testing.
- McNeil, L. (2000). *Contradictions of school reform: Educational costs of standardized testing*. New York: Routledge.
- Mercer, J. (1973). *Labeling the mentally retarded: Clinical and social system perspectives on mental retardation*. Berkeley, CA: University of California Press.
- National Center for Educational Statistics. (1995). *The educational progress of Hispanic students: The condition of education, 1995*. Washington, DC: U.S. Department of Education.
- National Center for Educational Statistics. (1998). *Annual report, 1998*. Washington, DC: U.S. Department of Education.
- National Council on Education Standards and Testing. (1992). *Raising standards for American education*. Chestnut Hill, MA: Author.

- National Education Goals Panel. (1993). *Setting standards, becoming the best* (Chapter 1, Vol. 1). Washington, DC: Author.
- Nettles, M. T., & Nettles, A. L. (1995). *Equity and excellence in educational testing and assessment*. Boston: Kluwer Academic Publishers.
- No Child Left Behind Act of 2001, Pub. L. No. 107-110, 115 Stat. 1425 (2002).
- Olson, J., & Goldstein, A. (1997). *The inclusion of students with disabilities and limited English proficient students in large-scale assessments: A summary of recent progress*. Washington, DC: National Center for Education Statistics, U.S. Department of Education.
- Oswald, D., Coutinho, M., Best, A., & Singh, N. (1999). Ethnic representation in special education: The influence of school-related economic and demographic variables. *Journal of Special Education, 32*, 194-206.
- Pearson, P., & Raphael, T. (1999). Toward a more complex view of balance in the literacy curriculum. In W. D. Hammond & T. E. Raphael (Ed.), *Early literacy instruction for the new millennium* (pp. 1-21). Grand Rapids: Michigan Reading Association.
- Reschly, D. (1997). *Disproportionate minority representation in general and special education: Patterns, issues, and alternatives*. Des Moines: Iowa Department of Education.
- Rueda, R., Artiles, A., Salazar, J., & Higuera, I. (2002). An analysis of special education as a response to the diminished academic achievement of Chicano/Latino students: An update. In R. Valencia (Ed.), *Chicano school failure and success: Research and policy agendas for the 1990's* (pp. 310-332). London: Falmer Press.
- Rueda, R., & McIntyre, E. (2002). Toward universal literacy. In S. Stringfield & D. Land (Eds.), *Educating at risk students: One Hundred-first Yearbook of the National Society for the Study of Education* (pp. 189-209). Chicago, III: The University of Chicago Press.
- Shepard, L. A. (2001). The role of classroom assessment in teaching and learning. In V. Richardson (Ed.), *The handbook of research on teaching*, 4th Ed. Washington, DC: American Educational Research Association.
- Smith, M., Heinecke, W., & Noble, A. (1999, Winter). Assessment policy and political spectacle. *Teachers College Record, 101*(2), 157-191.
- Snow, C. E., Burns, M. S., & Griffin, P. (1998). *Preventing reading difficulties in young children*. Washington, DC: National Academy Press.
- Valdes, G., & Figueroa, R. (1994). *Bilingualism and testing: A special case of bias*. Norwood, NJ: Ablex Publishing Corporation.
- Valencia, R., & Suzuki, L. (2001). Race/ethnicity, intelligence, and special education. In R. Valencia & L. Suzuki (Eds.), *Intelligence testing and minority students: Foundations, performance factors, and assessment issues* (pp. 187-207). Thousand Oaks, CA: Sage Publications.
- Yeh, S. (2001). Tests worth teaching to: Constructing state-mandated tests that emphasize critical thinking. *Educational Researcher, 30*(9), 12-17.

9

California's Standards Movement: How English Learners Have Been Left Out of the Equation for Success

Teresa I. Márquez-López
University of California, Riverside

INTRODUCTION

This chapter examines how the California Educational Reform Movement of the late 1990s, and recent policy changes, have had a detrimental effect on English learners who comprise more than one fourth of the state's school-age population. The BTSA-ELL Project was developed for use throughout the state. After 3 years of successful implementation, the BTSA-ELL Project was not adopted by the state for English learners and their teachers. This article cites the need to reconsider how California's Educational Reform Movement addresses the needs of English learners.

The Beginning Teacher Support and Assessment Program (BTSA) is California's, 2-year teacher induction program that is designed to be part of the teacher credentialing process. The BTSA-ELL Project was initiated in 1999, to prepare new teachers to meet the educational needs of English learners through professional development. At a time