

Escamilla, K. (2006)

Monolingual Assessment and Emerging Bilinguals: A Case Study in the US.

In O. Garcia, T. Skutnabb-Kangas, & M. Torres-Guzman (Eds.), *Imagining Multilingual Schools*.
Ch. 9: pp. 184-199. Clevedon, England: Multilingual Matters Ltd.

Posted with the permission of the publisher.

For more information, please go to www.multilingual-matters.com.

LINGUISTIC DIVERSITY AND LANGUAGE RIGHTS

IMAGINING MULTILINGUAL SCHOOLS

*Languages in Education
and Glocalization*



Edited by **Ofelia García, Tove Skutnabb-Kangas
and María E. Torres-Guzmán**



Dedication/Dedicatoria/ Omistuskirjoitus

The three of us, and most of our authors, are teachers/educators ourselves. Over the years we have worked with teachers. We have observed classrooms and interacted with thousands of teachers in all types of imaginable contexts. In many countries, teachers are poorly paid, work in unsatisfactory and often desperately bad conditions, with harsh and often impossible demands. Teachers are made invisible most of the time, but get all the blame when schools are seen 'not to deliver'. Despite this, most teachers persist, 'deliver', and are real change agents.

We admire your hard work, your energies and efficiency, your professionalism, your love and care, your reflection, analysis and openness to change, your devotion, and your imagination in finding ways to show children (and parents) that other worlds are possible!

We want to dedicate this volume to you and to all the world's teachers of multilingual children. We want to thank you for inspiration and for believing and showing through your daily work that change is possible.

Admiramos tu duro trabajo, tu energía y eficiencia, tu profesionalismo, tu amor y cuidado, tu reflexión, análisis y apertura al cambio, tu devoción, y tu imaginación al buscar maneras de enseñarle a los niños (y a sus padres) que hay posibilidad de otros mundos.

Queremos dedicarte este volumen, a ti, y a todas las maestras del mundo de niños multilingües. Queremos agradecerte tu inspiración y tu compromiso, mostrado a través de tu trabajo diario, de que el cambio es posible.

Me ihaillemme teidän kovaa työtänne, tarmoanne ja tehokkuuttanne, ammattitaitoanne, rakkauttanne ja huolenpitoanne, harkitsevaa ja erittelevää ajatteluanne ja muutosalttiuttanne, työllenne omistautumista ja mielikuvitustanne löytäessänne keinoja osoittaa lapsille (ja vanhemmille) että toisenlaisetkin maailmat ovat mahdollisia!

Haluamme omistaa tämän teoksen teille ja kaikille maailman monikielisten lasten opettajille. Haluamme kiittää teitä innostavasta esimerkistä ja siitä että uskotte ja näytätte päivittäisellä työllänne että maailmaa voi muuttaa.

Library of Congress Cataloging in Publication Data

Imagining Multilingual Schools: Languages in Education and Globalization.
Edited by Ofelia García, Tove Skutnabb-Kangas, and María E. Torres-Guzmán.

Linguistic Diversity and Language Rights: 2

Includes bibliographical references and index.

1. Education, Bilingual-Cross-cultural studies 2. Language and education-Cross-cultural studies. 3. Multicultural education-Cross-cultural studies. 4. Multilingualism-Cross-cultural studies.

I. García, Ofelia. II. Skutnabb-Kangas, Tove. III. Torres-Guzmán, María E.

IV. Series: Linguistic Diversity and Language Rights: 2.

LC3715.L45 2006

370.117

2006001484

British Library Cataloguing in Publication Data

A catalogue entry for this book is available from the British Library.

ISBN 1-85359-895-X / EAN 978-1-85359-895-1 (hbk)

ISBN 1-85359-894-1 / EAN 978-1-85359-894-4 (pbk)

Multilingual Matters Ltd

UK: Frankfurt Lodge, Clevedon Hall, Victoria Road, Clevedon BS21 7HH.

USA: UTP, 2250 Military Road, Tonawanda, NY 14150, USA.

Canada: UTP, 5201 Dufferin Street, North York, Ontario M3H 5T8, Canada.

Copyright © 2006 Ofelia García, Tove Skutnabb-Kangas, María E. Torres-Guzmán and the authors of individual chapters.

All rights reserved. No part of this work may be reproduced in any form or by any means without permission in writing from the publisher.

Typeset by Wordworks Ltd.

Printed and bound in Great Britain by the Cromwell Press Ltd.

New York and Trønninge Mose
Ofelia García, Tove Skutnabb-Kangas, María E. Torres-Guzmán

Chapter 9

Monolingual Assessment and Emerging Bilinguals: A Case Study in the US

KATHY ESCAMILLA

When a bilingual individual confronts a monolingual test ... both the test taker and the test are asked to do something they cannot. The bilingual test taker cannot perform like a monolingual. The monolingual test cannot 'measure' in the other language.
Valdés & Figueroa, 1994: 87

Introduction

The number of persons in the United States who are labeled as language minority people is large and growing. The term 'language minority' is commonly defined as people for whom English is a second language. However, included in this definition are people who are bilingual, or whose heritage language is not English. In 2003, the number of language minority people in the US was estimated to be 7.5 million (Ovando *et al.*, 2003). The US Department of Education requires that all schools in the US determine the extent to which students who are language-minority students are also limited in English proficiency. At one time in the US limited English proficient children were labeled LEP, but they are now more commonly referred to by federal, state and local school districts as ELL – English Language Learners (Ovando *et al.*, 2003). The US Department of Education (2003) estimates the total number of ELLs in US public schools to be about 5 million, and further estimates there are about 350 different language groups included in the population of language minority students and ELL students. Many federal and state documents tout the existence of enormous linguistic diversity in US schools. The state of Colorado, for example, has documented more than 102 language groups represented in its public schools (Escamilla *et al.*, 2003b). However, on closer inspection, this statistic is somewhat deceptive for in reality the vast majority (about 70%) of ELLs in the US speak one common language – Spanish.

The study reported in this chapter focuses on Colorado, a state whose demographic situation parallels the national one. In 2003, about 70,000

public school students in Colorado were identified as ELLs. This constituted about 10% of the entire school population. Further, although there were 102 documented language groups in Colorado schools, the vast majority (56,000 students or 80%) spoke Spanish as a first language (Escamilla *et al.*, 2003b).

In spite of national and state data showing that the overwhelming majority of language minority and English Language Learners speak a common language (Spanish), language diversity is frequently used as a reason to promote monolingual English assessment policies, and in many places to require that assessments be conducted only in English. In the case of Colorado, Colorado Department of Education officials frequently argue for assessment and testing to be conducted only in English because, in their view, it would be 'too expensive' and 'not feasible' to develop assessments in 102 languages (Lenhart, 2003). It is important to note that the creation of assessments in Spanish would account for 80% of the linguistic diversity and possibly create a more equitable and accurate assessment system for ELLs, a point that will be expanded on later in this chapter.

Ruiz (1988) proposed that there are three basic language orientations utilized by nations, communities and schools as they engage in the creation of language policies and planning: language as a problem, language as a right and language as a resource. Ruiz posited that ELLs in the United States, particularly those ELLs whose first language is Spanish, have historically been viewed as 'problems' in US schools. Since students' limited ability in English is viewed as a 'problem,' US schools have been charged to create policies and practices to 'fix the language problem' of Spanish-speaking students. Recent educational reform movements such as standards-based education and high-stakes testing have served to exacerbate the notion that speaking a language other than English in the US is a problem that must be remediated by the schools.

Standards-based education in the US began in the early 1990s and was designed to improve academic achievement for all students (McLaughlin & Shepard, 1995). By the year 2000, 49 states had implemented some form of standards-based educational reform (Cunningham, 2000). The standards-based reform movement has two basic components – content and assessment. *Content standards* include the knowledge, skills, and other understandings that schools should teach in order for students to attain high levels of competency in challenging subject matter (McLaughlin & Shepard, 1995). *Assessment standards* (sometimes referred to as performance standards) describe how students should be able to demonstrate that they have acquired the requisite content and/or skills. While assessments can be conducted in many different formats, the predominant assessment format is a standardized test.

The implementation of wide-scale assessments for all students to meet

the requirements of standards-based education reforms has been reinforced and expanded by the passage of the *No Child Left Behind Act* (NCLB, 2002). *No Child Left Behind* is a federal mandate that now requires all states to carry out annual assessments of student progress in reading, writing and various content areas (for more on the consequences of NCLB see Hornberger, Chapter 11 and Shohamy, Chapter 8). Results of this annual testing are now used as measures of school accountability and effectiveness. Test results are used to determine the schools' effectiveness and monitor the children's progress. The use of wide-scale testing is now the predominant means by which schools and school districts demonstrate that they are meeting the mandates of *No Child Left Behind* and the standards-based movement. Wide-scale testing, for this reason, has come to be known as 'high-stakes testing.' Wide-scale testing involves giving the same standardized test to a large population of students (Amrein & Berliner, 2002). Wide-scale tests become 'high-stakes' tests when consequences (either sanctions or rewards), are applied to their results. Amrein and Berliner (2002) document the consequences that have been attached to schools and school districts with the onslaught of wide-scale, high-stakes testing programs. For example, currently, 45 states use the results of wide-scale tests to rank and rate schools – from best to worst (Amrein & Berliner, 2002). Further, punishments are attached to testing outcomes twice as often as rewards are. Schools that receive unsatisfactory ratings over the course of several years are threatened with closure, firing of teachers and administrators, and/or conversion from public schools to private schools. In many cities in the US, conversion from public school to private or charter school status means that the public school has failed. In short, there is enormous pressure on public school teachers and students to do well on high-stakes tests.

Since the inception of the standards-based education movement questions and concerns have been raised about the application of these policies to students who are learning English as a second language (ELLs). Little or no controversy has surfaced with regard to content standards. Indeed, there is widespread agreement that ELLs can, and should, meet challenging content standards. However, significant issues have been raised with regard to assessment standards for ELLs. August and Hakuta (1997) list some of these issues:

- Standardized tests that most states currently employ to meet state and federal mandates were developed for the assessment of native English speakers, not for ELLs.
- For second language learners, paper/pencil content assessments in English often measure students' lack of proficiency in English and NOT their knowledge of the content.

- ELLs are often placed in short-term educational programs with inadequate materials and poorly prepared staff, which limits their opportunities to learn.
- Many state assessment programs are available only in English. They do not provide opportunities for students to demonstrate their knowledge of content in a non-English language.

For the purposes of this chapter, it is important to note that most wide-scale high-stakes tests utilized in the US use native English speakers as the norm (Abedi, 2001; Menken, 2000). Many have argued that it is invalid to use high-stakes monolingual English tests to assess ELLs (Gottlieb, 2001; Abedi, 2001; Menken, 2000). Nonetheless such tests continue to be used to evaluate the academic achievement of ELLs and other students, and to rate and rank the efficacy of their school programs and the competency of their teachers and administrators. In school districts and schools where there are large numbers of ELL students, the English-only assessment policies are particularly problematic, for it is unlikely that students just beginning to learn English will do as well on these assessments as their monolingual English counterparts. While the debate on the equity and validity of high-stakes testing in general continues, as does the debate on including ELL students in wide-scale high-stakes testing programs, federal and state mandates are dictating that all students, including ELL students, be included in such testing programs.

It is important to note that 20 states, as well as the federal government, have created special testing policies with regard to ELL students with the hope of being more equitable to these students (Rivera & Stansfield, 2000). Special policies include the following:

- (1) exemption for ELLs from testing in English for three years;
- (2) special testing accommodations such as allowing ELL students more time to take the test, allowing ELLs to use bilingual dictionaries, allowing students to respond to test questions in their native language with bilingual personnel;
- (3) taking the assessment in the student's native language.

Each state varies with regard to its special testing policies. However, most states utilize exemptions and accommodations, and very few provide high-stakes tests in students' native languages (Rivera & Stansfield, 2000).

The state of Colorado utilizes all three special testing policies – exemption, accommodation, and limited opportunity for native language assessment. Colorado allows students to be exempted from English CSAP (Colorado Student Assessment Program) testing for 3 years; it provides limited accommodations and allows students to take assessments in reading and writing in Spanish at the 3rd and 4th grade levels only.

This chapter presents results of a case study involving Spanish-speaking ELLs and high-stakes tests in the state of Colorado. First, it examines the way in which high-stakes assessment programs have served to perpetuate the 'language as a problem' paradigm in US schools with large numbers of ELLs. Second, it examines the results of Spanish-speaking ELL students who were allowed to take the assessment in Spanish and compares these results to those who took these assessments in English. Finally, the chapter discusses the privileging of English testing data over test data in other languages (e.g. Spanish) and how the privileging of English outcomes mitigates against the development and implementation of more effective school programs for emerging bilingual children.

High-stakes Testing and Spanish-speaking Children in Colorado

The high-stakes test in the state of Colorado is the Colorado Student Assessment Program (CSAP). This test has been given since 1997 and the number and types of CSAP tests have increased over the years. As with other states, CSAP results are used to rate and rank all Colorado schools and school districts (Colorado SB 186, 2000). Results on the CSAP are used to assign one of five labels to Colorado schools: excellent, high, average, low, and unsatisfactory. As with many other states, Colorado has sanctions for schools that are consistently ranked as unsatisfactory. Sanctions include firing teachers and administrators, and state-mandated conversion to charter schools if unsatisfactory ratings persist for three years (SB 186, 2000). Further, several school districts in Colorado have begun to require that teacher raises and salaries be tied to outcomes on the CSAP test.

With regard to English Language Learners, Colorado provides two basic accommodations. ELLs in Colorado can be exempted from the CSAP for three school years while they are learning English. This exemption was intended to avoid unfair and punitive sanctions for schools with large numbers of ELLs. The three-year exemption accommodation is widely used across the US, and is also allowable in *No Child Left Behind* (Rivera & Stansfield, 2000). Colorado also allows for a less common accommodation for Spanish-speaking ELLs – the opportunity to take the CSAP test in reading and writing in Spanish in the 3rd and 4th grades. This accommodation is designed to give schools that are providing some type of bilingual instruction an opportunity to assess children in a manner that matches the schools' instructional practice.

The CSAP results in Spanish and English for the year 2002–2003 provide the data for this chapter. These data illustrate the pervasive view that ELLs, in this case Spanish speaking ELLs, present language problems that schools need to remediate. The data further reveal how high-stakes testing has been

used to reinforce the 'language as a problem' paradigm even in the presence of powerful evidence to the contrary. The data presented below are a part of a larger five-year study that examines the impact of the CSAP on ELLs in Colorado (Escamilla *et al.*, 2000, 2001, 2002, 2003a, 2003b). Findings across the five-year span of this study have yielded similar results.

In July 2003, the Colorado Department of Education released the annual results of the CSAP assessment to the public. Results are placed on the CDE website and are released and reported by the popular media in Colorado. Data in Table 9.1 represent 10 schools in Denver. Scores are reported by the state and individual schools in terms of the percentages of students whose results earned them one of four labels (unsatisfactory, partially proficient, proficient or advanced). The table reports percentages of students in Spanish and English who scored at proficient or advanced. These two categories are considered to be indications that students are performing at grade level. Of the 10 schools, 8 were ranked as low, 1 was ranked as average and 1 was ranked as unsatisfactory. The state called the CSAP results 'alarming' for Denver public schools. The popular media reported that the results demonstrated a 'gap in achievement' between Denver and other schools districts, and district officials and teachers were asked to account for why the rankings at these schools were low.

Among the responses given for the low rankings in press releases and interviews with district officials were the following: (1) poverty, (2) large numbers of Spanish-speaking students, (3) large numbers of students who are ethnically Latino and (4) bilingual education programs. All sample schools are implementing a form of early-exit bilingual program, labeled ELA by the Denver Public Schools. The press, the Colorado Department of Education, school district administrators, and teachers cited common reasons for the perceived under-performance of these schools. The schools and the community attributed perceived low performance to language, ethnicity, poverty, and to the bilingual education program. Data presented on Table 9.1 seem to confirm these perceptions. For example, 8 out of 10 schools have a student population that is over 90% Latino. Six of the schools have a population where over 50% of the students are ELLs and speak Spanish as a first language. The other four schools have a Spanish-speaking ELL population that accounts for one-third of the student body. Finally, in 8 out of the 10 schools 90% of the students live in poverty.

Each of these schools provides a short-term bilingual program to students who enter school speaking Spanish and as ELLs. In these programs, students learn literacy and content area subjects in Spanish first and receive daily instruction in English as a Second Language. At each of these schools, students who were in the early exit bilingual program, who were in 3rd or 4th grades at the time of the study, and who were learning to read and write in Spanish were given the CSAP assessments in Spanish. At

Table 9.1 Denver Public Schools: Demographics, school rankings, and comparison of 3rd grade Spanish and English reading and writing outcomes (Spring 2003)

School	State-rating 2001-2002	School population	% Latino	% ELL Spanish	% Poverty	Spanish CSAP reading% proficient/ advanced	School wide English reading CSAP % proficient/ advanced*	Spanish CSAP writing % proficient/ advanced	School wide English writing CSAP % proficient/ advanced*
Beach Court	low	454	91	37	84	79	35	74	42
Bryant Webster	average	460	97	51	92	91	67	96	60
Cheltenham	low	601	92	43	97	67	42	74	21
Cowell	low	625	92	55	95	82	44	72	27
Eagleton	low	556	94	50	90	70	27	79	19
Fairmont	low	475	89	47	97	76	45	71	15
Gilpin	low	440	63	33	85	75	33	63	16
Knapp	low	661	91	55	91	56	52	42	19
Munroe	low	608	93	54	95	67	39	63	42
Smedley	unsat.	545	95	52	96	59	29	68	21

* For Spring 2003, 51% of the 3rd grade students across Denver Public Schools who took English CSAP in reading earned a rating of proficient or advanced, and 35% of those who took English CSAP in writing earned a rating of proficient or advanced.

each of these schools there are grade levels and strands of students who are not in the early-exit bilingual program. They are in English-only instructional programs. The students in the English-only instructional programs learn to read and write in English and were given the CSAP in English. The data in Table 9.1 also present the reading and writing outcomes for these schools and compares Spanish and English results. From these data, a different pattern emerges.

The CSAP reading and writing data are reported by school and by grade according to the number and percentage of students who fall into one of four categories (advanced, proficient, partially proficient, and unsatisfactory). Numbers presented in Table 9.1 indicate the percentage of children at each school who scored in the proficient and advanced ranges. Proficient and advanced ranges equate to grade level proficiency.

To illustrate these results further, consider the example of Beach Court School. At this school 79% of the children who took the CSAP reading assessment in Spanish scored in the proficient/advanced range, while 35% of the children who took the CSAP in English scored in the proficient or advanced range. The district average for students scoring in the proficient or advanced range is 51%, meaning that the students who took CSAP in Spanish outpaced the district average in English and outpaced the students at their own school who took the test in English. Students at Beach Court who took the test in English performed lower than the district average. Results for each of these 10 schools illustrate that the achievement for students taking CSAP reading in Spanish is higher than that for students taking the CSAP in English. Further, Spanish results in each of these schools are higher than the district-wide average in English. These data raise serious questions with regard to the assumption that the school's overall ranking was a result of having too many poor Spanish-speaking students who were participating in bilingual programs. The 10 schools represented in Table 9.1 reflect 10 elementary schools in Colorado that are the very heavily impacted by Spanish-speaking immigrant students and thus have very high numbers and percentages of ELL students.

Table 9.1 also presents the results on the writing portion of the 3rd grade CSAP and compares results on the Spanish CSAP and English CSAP on the writing portion of the exam. These results mirror the reading portion of the exam and once again illustrate that the achievement levels in Spanish are higher than the achievement levels in English. Again to illustrate the findings, let's take the case of Cheltenham School. At Cheltenham, 74% of the students who took the CSAP Spanish writing exam scored proficient or above, while only 21% of the students who took the CSAP writing exam in English scored proficient or advanced. Students taking the CSAP in Spanish were above the district-wide average of 35% and students who took the test in English were below. Again, students taking the writing test

in Spanish at all 10 schools had better outcomes than those students taking the assessment in English at these same schools and better outcomes than the entire district.

It is important to note that the intent of the standards-based education movement in Colorado and the US is to improve student achievement vis-à-vis content standards. In the data reported above, the content standards relate to reading and writing. The law did not specify that students must demonstrate their emergent knowledge of reading and writing in English, it simply mandates demonstration that students are becoming literate. In this case, it seems that Spanish-speaking students are becoming literate in Spanish at higher levels than peers learning only in English.

The data comparing outcomes between students taking the CSAP in Spanish and those taking the CSAP in English are remarkable for several reasons. First, the results question the conventional wisdom reported by the Colorado Department of Education, the popular press, teachers, and administrators that poverty, speaking Spanish and being Latino are explanations for low school ratings and academic under-achievement. In fact, the converse seems to be the case. It is interesting and noteworthy that the state does not display Spanish results and English results for schools in Colorado side by side. Spanish results and English results are reported on separate areas in the state website, even for students who attend the same schools.

Further, in the 8 years since the CSAP results have been reported by the Department of Education to the popular media in Colorado, outcomes of the Spanish CSAP have been inconsistently and infrequently reported to the public. For example, all of the English CSAP results for the English CSAP are available from the Colorado Department of Education website (<http://www.cde.state.co.us>). Though the Spanish data are also public domain information, accessing the Spanish data is more of a challenge. For example, to access the Spring 1999 data, researchers had to wade through boxes at the Colorado Department of Education that were still as they had been turned in by school districts. In the year 1999, the data were neither summarized by the Department, nor reported in any official documents or other outlets. There are no data for 1999 on the CDE CSAP website. For the Spring 2000 data, the Spanish data were summarized by CDE and placed in notebooks, but were still not placed on the CDE website or available to other outlets. After much pressure, in Spring 2001, data, for the first time, were summarized, and placed on the CDE website. However, data were catalogued in separate areas on the CDE web-site and it was difficult to locate the Spanish results. For example, English results could be located under 3rd grade English reading (<http://www.cde.state.co.us/cdeassess/csap>), while Spanish results were located under '*lectura*' (<http://www.cde.state.co.us/cdeassess/csap>). In order to find the Spanish results and

compare them to English, one had to know that '*lectura*' meant 'reading' in Spanish. The same is true for the writing results that were categorized under 'writing' in English and '*escritura*' in Spanish. Since data for English and Spanish are not reported nor presented side by side, it is difficult to question the pervasive assumptions that Spanish speakers are responsible for low and unsatisfactory school ratings.

There are several other reasons why the Spanish reading and writing outcomes are rather remarkable in the Denver school district. The CSAP testing company annually releases 25% of the test items on all CSAP tests so that teachers can use these items to help children practice for the test (CSAP: 1998). Test items, however, are released only in English, not in Spanish. Reasons for releasing items only for the English version of the test vary. One explanation from the Colorado Department of Education is that it's too expensive to release items in Spanish. Another from CDE claims that too few children take the test in Spanish to justify this exercise. Yet, another claims that the test in Spanish is 'easier' because Spanish is an easier language to learn to read and write and, therefore, no practice is necessary. Whatever the rationale for not allowing the Spanish-speaking children to practice for the CSAP, the fact is that the preparation to take the test is not the same for Spanish and English test takers. English speakers get to practice for the exam and Spanish speakers do not. Still, Spanish speakers over the past five years have consistently outperformed comparable English-speaking peers on the CSAP test (Escamilla *et al.*, 2003b).

The test results are similarly remarkable in that despite claims by the testing company that the CSAP is free of cultural, economic and linguistic bias (CSAP, 1998), there is evidence to the contrary. For example, the third grade writing prompt for the 2004 CSAP asked the children to write about their favorite camping trip. Taking vacations to the mountains and camping are arguably experiences that few poor immigrant children have had, thereby placing them at a disadvantage when having to write a narrative about this kind of experience. Again, in spite of this apparent disadvantage, children taking the CSAP in Spanish did well.

Finally, there is a general tendency to undervalue the results on the Spanish CSAP. As stated above, Spanish results are not regularly reported to the press, and the public has paid scant attention to them over the past 8 years. Further, the state legislators and policy-makers have been ambivalent about the Spanish CSAP since its inception. For example, during one school year (2000–2001), the Colorado Department of Education issued three different policy directives about how the Spanish CSAP would be used and factored into school ratings. One directive stated that children could take the CSAP in Spanish but that it would count negatively toward a school's overall ratings. In this directive, schools were informed that each child taking the test in Spanish would receive a score of minus 0.5 toward the

overall school ratings. In other words, testing in Spanish would result in a negative (below zero) score for a school. Having many children take the test in Spanish would have resulted in the lowering of an overall school ranking. Threats of litigation and accusations that the Colorado State Department of Education had overstepped its legislative authority prompted a second policy directive stating that if students took the CSAP in Spanish, it would not count negatively toward the school ranking, but it would also not count positively. In other words, it would have zero impact. This directive amounted to making the Spanish test results, while not negative, invisible and irrelevant. Further, this directive meant that only English CSAP outcomes would count toward a school's rating. The threat of punitive measures if children took the test in Spanish followed by the threat to discount (zero count) test results in Spanish caused many schools to question whether or not they should be teaching children to read and write in Spanish, given the potential negative consequences associated with teaching and testing in Spanish. It finally took a legislative act (SB 98, 2001) to establish that outcomes on the Spanish CSAP tests would be given equal weight to the outcomes on English CSAP assessments.

It is fortunate for children in Colorado that currently Spanish and English CSAP outcomes must be given equal weight in the calculation of school ratings. Unfortunately, SB 98 was passed in April 2001, one month after children had taken the CSAP tests for the year 2000–2001. Fearful of being punished for teaching in Spanish, many schools decided that it was not worth the risk to teach and assess children in Spanish and switched children from Spanish to English instruction. The schools that decided to stay the course and continued to teach and test in Spanish did so in a negative, even hostile climate. Again, given the environment, the consistent and positive testing outcomes in Spanish are remarkable.

Implications

The data presented above have significant implications for monolingual high-stakes assessment systems. Ironically, even in a state like Colorado where children are allowed to take high-stakes tests in a limited way in their native language (in this case Spanish), the high-stakes testing environment seems to have reinforced the notion that 'language is a problem' in spite of overwhelming evidence to the contrary. In fact, the documentation and presentation of counter-evidence to the mentality that 'language is a problem and causes low CSAP scores' has resulted in political activity designed to limit the number of children who are eligible to take the assessment in Spanish. Using the federal mandates under *No Child Left Behind*, the state has instituted very strict policies requiring that all children be tested in English after 3 years, no matter what the child's level of English proficiency.

This policy does not prevent a school or school district from testing a child in both English and Spanish. Indeed, from a research perspective this may be the most interesting and appropriate means to assess the academic achievements of Spanish-speaking ELL students. However, given the expense associated with the CSAP assessment program and the time the assessments take, there are currently no schools or school districts in Colorado that have the resources or time to carry out double assessments. Thus, for a limit of three years, children are allowed to take assessments in Spanish, but under no circumstances are emerging bilingual children assessed in both of their languages.

An additional layer of complexity was added to the assessment picture in Colorado in 2004–2005. A CSAP assessment test in mathematics was developed for 3rd and 4th grade students and was administered for the first time in Spring 2005. These two math tests were developed in English only. Therefore, teachers who had taught children in Spanish because children could take assessments in Spanish are now faced with a situation where children can take reading and writing tests in Spanish in 3rd and 4th grade, but these same children must take mathematics assessments in English. The emphasis on testing outcomes in English and the development of new assessments in English only are placing increasing pressure on schools and teachers to abandon teaching in Spanish in spite of positive test results.

A significant implication of this study is that conventional wisdom is hard to change. The notions that being poor, and speaking Spanish are explanations for low achievement are so prevalent that they go unquestioned. When counter evidence is presented, this evidence is often ignored. Most alarming, however, is the great potential that exists for educators and policy-makers to prescribe inappropriate educational solutions to ELL students because they have misdiagnosed and misunderstood the nature of perceived underachievement.

Evidence is beginning to emerge in Colorado of ill-advised solutions being applied to misidentified problems. There are new policy initiatives that are designed to curtail the ability of teachers and schools to teach and test in Spanish. For example, the new English-only math test and the 3-year mandate to test in English, together with policies that do not test children in two languages, are reducing the number of children who are being taught and tested in Spanish, and are limiting the time that children may learn in Spanish. If children are not being taught and tested in Spanish, it is not possible for them to demonstrate what they know in Spanish. If sufficient pressure is placed on schools and teachers to teach and test only in English, then there will be no Spanish outcomes to be reported, and therefore, no opportunity to document the potential of emerging bilingual children when assessed in their native language. There will also be no evidence to document the efficacy of bilingual education programs. No native language teaching

and assessment means no opportunity to collect evidence to challenge conventional wisdom, thereby protecting the status quo 'language is a problem' paradigm.

In spite of evidence of the success of teaching and testing ELL children in Spanish, school districts in Colorado continue to propose to solve the 'problem of educating Spanish-speaking students' with programs that devalue Spanish and emphasize English. Many school districts in Colorado are proposing instructional programs that use more English and less Spanish. They are suggesting that more English as a Second Language classes be implemented to take the place of physical education, music and art. They are also proposing more homework in English and more skills-based learning (in English). English classes for parents are also being recommended, so that parents can help their children in English. Finally, they are advising more practice for high-stakes tests in English for all ELL populations, particularly Spanish speakers.

This type of curriculum has been labeled the intensification curriculum by Berliner and Biddle (1995). In this case, since educators and policy makers view the problem of underachievement to be one of too much Spanish and too little English, they propose to reduce teaching and testing in Spanish and to intensify teaching and testing in English. Berliner and Biddle suggest that intensification strategies designed to 'improve English' may not produce the desired academic achievement. More English may not be better.

If high-stakes testing in Spanish is further limited or eliminated altogether in Colorado, the consequences are likely to be negative for Spanish-speaking children. Shohamy (2004a) has stated that a language test is a dangerous weapon, potentially lethal to all children (for more on this, see Shohamy, Chapter 8). Language tests in the form of high-stakes tests are potentially lethal weapons if they are given in English only to children who are learning English as a Second Language. Used in this manner, such tests serve only to legitimize the notion that there is a 'gap' in achievement, that ELLs are underachievers, and that speaking a non-English language causes problems in learning (Black & Valenzuela, 2004). However, the use of high-stakes tests in languages such as Spanish, even in a limited way such as in Colorado schools, provides the opportunity for children to be tested in their native language, thus illuminating a different set of outcomes and realities.

In this study, the outcome of the Spanish CSAP documented that poor Spanish-speaking immigrant children can learn to read and write well, can meet state content standards, and can meet the demands of high-stakes testing when allowed to demonstrate in their native language what they know. The use of high-stakes tests in Spanish in Colorado has allowed us to document a different outcome – a profile of an emergent bilingual who outperforms comparable monolingual English peers in their schools.

Rather than an under-performing English Language Learner, we now have a profile of students who are helping schools be better than they would be if only monolingual students were in attendance. Sadly, such positive results have largely gone unnoticed and unreported in Colorado; and no doubt much more documentation and dissemination of results will need to occur for such positive results to begin to change the current conventional wisdom. Moreover, results such as those presented in this chapter threaten the status quo. Current trends in Colorado evidence greater efforts to maintain and protect the status quo through more intense English programs and fewer opportunities to teach and learn in Spanish.

I end this chapter where I began, and with a note of hope mixed with frustration. As Valdés and Figueroa (1994) have stated, we will never adequately assess the bilingual mind through the use of monolingual tests. In a perfect world, our emerging bilingual students would be assessed in both English and Spanish to document and support their bilingualism. In Colorado, we have an assessment system that enables us to begin to understand how high-stakes tests, rather than being lethal weapons, might be tools to document the linguistic strengths of Spanish-speaking students. Such assessments might be used to counteract arguments that language is a problem and a cause of underachievement. Further, outcomes of these assessments might become tools to argue for the expansion of dual language and other bilingual programs designed to fully develop bilingualism and biliteracy in children who currently carry negative labels such as ELLs.

It is conceivable that these high-stakes tests could be given in both English and Spanish to emerging bilinguals to also document and affirm the value of bilingualism and give added weight to knowledge that can be demonstrated in each language. In such a system, the outcomes in two languages of emerging bilingual children could be used as a positive weight in a school's ranking. In other words, there might be an incentive for students to learn and take assessments in two languages. Optimistically, these potentially lethal tools could be used to validate that two is indeed greater than one. In reality, however, in Colorado, as in other states, high-stakes tests in languages other than English are devalued, results in Spanish are ignored, and policies are based on conventional wisdom rather than on a thorough understanding of the data. To imagine, invent, implement and evaluate multilingual schools will require utilizing a range of assessments and tools to allow children to demonstrate what they know and are learning in all of their languages. The current US policies that emphasize monolingual-English-only assessment mitigate against the potential of multilingual schools.

Summary

This chapter presents a case study examining the implementation of large-scale programs for testing emerging bilingual students. US schools have become increasingly diverse linguistically. However, assessment systems to determine student academic achievement have become more monolingual in their focus. Second language learners are frequently blamed for any perceived under-performance on these exams. In a few states, including Colorado, Spanish/English emerging bilinguals are allowed to take assessments in Spanish. Data collected in this case study compared academic achievement in reading and writing between students who took the Colorado CSAP test in Spanish and English. Findings indicated that students taking the assessments in Spanish outperformed students taking the CSAP in English, and in the schools studied, Spanish outcomes exceeded district-wide English averages. Findings support the need to assess emerging bilinguals in both their languages.

Este capítulo demuestra cómo las pruebas de serias consecuencias afectan a los estudiantes que se están desarrollando como bilingües. Las escuelas estadounidenses cada año tienen más diversidad lingüística. Sin embargo, el sistema de asesoramiento que determina los logros de los estudiantes cada día se enfoca más en el inglés. La culpa de no tener éxito en estos exámenes se le achaca a los estudiantes que están aprendiendo inglés como segundo idioma. En algunos estados como Colorado, los estudiantes que van a ser bilingües tienen la oportunidad de ser asesorados en español. En este estudio comparamos los logros en lectura y escritura de los estudiantes que tomaron el examen CSAP en Colorado en español y en inglés. Los resultados indican que los estudiantes que tomaron la prueba en español tuvieron más alto rendimiento que aquellos que la tomaron en inglés, y que en las escuelas que hemos estudiado, los resultados en español son mucho más altos que los del distrito en inglés. Los resultados apoyan la necesidad de que se le dé los exámenes a estos estudiantes en los dos idiomas. (Spanish)

本文欲呈現單一個案研究：此一個案乃全面性察驗具有雙語能力的學生，將如何面臨未來美國的考試制度。當代美國的校園環境已漸漸呈現多元語言風貌，但是，在考試制度方面卻仍然偏向使用單一語言來測試學生在學科上的表現並藉此判定其學業成就。在面臨考試時，英語為其第二語言的學生多半被視為學業成績低落者。目前在美國境內有幾個州，包括科羅拉多州，已開始實施以下應考政策：具有西班牙語和英語雙語能力的學生，將可選用自己的母語—西班牙語文來應考。此一個案的研究資料，即是搜集學生以西班牙語和英語來應考科羅拉多州的州立考試的資料並加以比較其應試結果。研究結果顯示，以學生本身而言，選擇使用本身西班牙母語來應考的學生，其應試成績乃遠勝於那些不用母語而使用英語去應考的學生之應考成績。同時，研究成果並指出，以學校整體而言，那些選擇採用西班牙語來讓學生應考科州州考的學校學生應試成績，亦高於只選用英語來進行該州州考的學校之學生成績。此個案研究結果，將進一步支持未來以兩種語言來測試具有雙語能力學生之考試需求。

(Chinese)