The Federal Bilingual Education Program

Christine H. Rossell

Brookings Papers on Education Policy, 2000, pp. 215-243 (Article)

Published by Brookings Institution Press
DOI: https://doi.org/10.1353/pep.2000.0014

The Federal Bilingual Education Program

CHRISTINE H. ROSSELL

SINCE THE 1960s, the U.S. Department of Education has enthusiastically embraced bilingual education. At the time this love affair began, no research evidence supported bilingual education as the best means for limited-English proficient (LEP) children to learn English and other subjects that a child will be tested on in English. Nor did any agreement exist on the definition of the target population or even on what bilingual education is. Some thirty years later, there still is no consistent evidence available to support bilingual education as the best means for LEP children to learn English and other subjects that they will be tested on in English, nor any agreement on the definition of the target population or bilingual education. Yet, the federal government's enthusiasm for bilingual education seems undiminished.

What is bilingual education? There are currently three different basic instructional programs for LEP students: (1) native tongue instruction, characterized by learning to read and write in the native tongue and learning subject matter in the native tongue and eventually transitioning to English, (2) structured immersion—all-English instruction in a self-contained classroom containing English language learners, and (3) regular classroom instruction with English as a Second Language (ESL) instruction in a pullout setting—all being implemented in school districts around the country and all called "bilingual education" by federal, state, and local administrators, legislators, reporters, and educators.[1] Only the Spanish speakers, however, are receiving bilingual education through native tongue instruction, according to the theory (the program described above), and not even all of them are. Thus, there is no single treatment

called "bilingual education" that is implemented in the same way and understood to be the same thing by everyone.

The Origin of Title VII

Bilingual education began as a part of the civil rights movement of the 1960s. Senator Ralph W. Yarborough (D-Texas) catalyzed Hispanic support for federal intervention on behalf of linguistic minority students by painting a picture of Hispanics as disadvantaged minorities who had been the victims of discrimination. In the 1960s, an almost 40 percentage point gap in high school completion rates existed between whites and Hispanics, favoring whites, and an almost 10 percentage point gap existed between blacks and Hispanics, favoring blacks. Given that Hispanic immigrant children were all being taught in English, the problem seemed obvious to their advocates. Instruction in English was the cause of their low achievement, and the loss of their native tongue was a civil rights violation that could be remedied by bilingual education.

Congress embraced this logic by passing three statutes that addressed the educational and civil rights of linguistic minority students: (1) the Bilingual Education Act of 1968, also known as Title VII of the Elementary and Secondary Education Act (ESEA), (2) Title VI of the Civil Rights Act of 1964, and (3) the Equal Educational Opportunities Act (EEOA) of 1970. Administrative agencies, such as the Office for Civil Rights (OCR) and the Office of Bilingual Education and Minority Language Affairs (OBEMLA), were created as a result of this legislation and directed to implement these provisions.[2] The federal courts also reviewed several lawsuits brought by language minorities claiming that school districts had violated these federal laws by ignoring the language barrier.[3]

The 1968 Bilingual Education Act was the first and most important piece of federal legislation devoted exclusively to the needs of language minority students. Although the 1964 Civil Rights Act and the 1965 Elementary and Secondary Education Act dealt with the problems of children disadvantaged by segregation and poverty, Hispanics and their advocates felt they were uniquely disadvantaged by language and culture and that this was not addressed in programs for poor children and children of color.

Hispanic advocates had high hopes for the 1968 Bilingual Education Act, but in the end, it was nothing more than a modest grant-in-aid pro-

gram designed to promote research and experimentation in bilingual education. The policy declared:

> In recognition of the special educational needs of the large numbers of children of limited English-speaking ability in the United States, Congress hereby declares it to be the policy of the United States to provide *financial* [emphasis added] assistance to local educational agencies to develop and carry out new and imaginative elementary and secondary school programs designed to meet these special educational needs. (Section 702)

School districts that did not want federal funds did not have to implement bilingual education or any other program for limited-English proficient children. Nor were the federal funds at that time sufficient to be much of a carrot, because the amount of money authorized for the fiscal year ending June 30, 1968, was a mere $15 million, increasing to $30 million in fiscal 1969 and $40 million in fiscal 1970.

Defining the Eligible Population

The original bill filed by Senator Yarborough identified the eligible population as Spanish speakers. This met with considerable opposition, and so the final bill encompassed all language groups:[4]

> For the purposes of this title, "children of limited English-speaking ability" means children who come from environments where the dominant language is other than English. (Section 702)

Because some children who come from non-English-speaking families are fluent in English and high in achievement, this definition broadened the target group considerably.

Title VII was reauthorized in 1974, 1978, 1984, 1988, and 1994 and will be reauthorized again in 1999. The 1974 amendments (section 703) narrowed the definition of the eligible population to those who were of "limited-English-speaking ability," not just from a non-English-speaking family. The 1978 amendments changed the term to "limited English proficiency" (section 7003) and added American Indians and Alaskan Natives to the eligible population of students. The 1978 amendments also added reading and writing to the difficulties children might have in English that would deny them the opportunity to learn successfully in classrooms where the language of instruction is English.

The 1984, 1988, and 1994 reauthorizations did not change this definition in any important way, and the amount of funds to be allocated has continued to depend on the number of LEP students claimed by a school district. The general presumption of Title VII has been that determining who is LEP is a rational process. The 1994 reauthorization requires no specific identification or assessment procedures, and the Clinton administration's 1999 proposal to reauthorize the Elementary and Secondary Education Act is similarly silent.[5] In their grant applications, school districts simply have to present data on the number of children and youth of limited-English proficiency in the school or school district to be served and their characteristics, such as language spoken, dropout rates, proficiency in English and the native language, academic standing in relation to the English-proficient peers of such children and youth, and, where applicable, the recency of immigration.

The only problem identified in the legislation is the consistency of national estimates, and this only surfaced in the last reauthorization. The 1994 act specified in section 7132 (b) (3) that the secretary of education may establish (through the National Center for Education Statistics in consultation with experts in bilingual education, second language acquisition, and English as a Second Language) a common definition of limited-English proficient students for purposes of national data collection.

Neither the 1994 legislation nor the Clinton administration's 1999 proposal recognized that, because school districts receive more money for students if they are identified as LEP, a fiscal incentive exists to overidentify. The Clinton proposal dutifully recites a Council of Chief State School Officers' report that, between 1990 and 1997, the number of LEP students increased by 57 percent to roughly 3.4 million and the LEP population more than doubled in eighteen states.[6] These statistics are cited as unimpeachable proof of the need for federal aid to school districts with LEP students.

The fiscal incentive to overidentify comes not only from federal funds, which are about 7 percent of funding for public schools, but also from state funds tied to the number of LEP students.[7] State funds are almost half of school district revenues. Because a below-average student brings in more state and federal revenues if he or she is identified as LEP, a rational actor will want to identify such a student as LEP.[8]

But the problem of identifying LEP students is bigger than simply a fiscal incentive to overidentify. The procedures used by school districts to

identify students as limited-English proficient are inherently illogical. And they cannot be made logical.

The selection procedures used by school districts are reducible to two steps: (1) a home language survey is administered to all students to identify the pool of potential LEP students; and (2) the students identified in the home language survey are tested on several measures of academic performance and are classified accordingly.

The decision to exit a student from bilingual education or special language assistance involves procedures similar to those used to determine eligibility. Students are reclassified as fluent-English proficient (FEP) if they score at or above a certain percentile on an English language test that has been normed on an English-speaking population. The decision to reclassify is tempered by staff judgment, either the child's classroom teacher or a team of professionals employed by the school district.

The Home Language Identification Survey

The home language survey is the first step in the process of identification of LEP students. The Boston home language survey is typical of those used by school districts around the country. It asks parents to fill in basic demographic data such as their country of birth and highest grade completed and to indicate the language they understand best. Parents are asked to respond to the following questions about the home environment:

1. What language(s) are spoken and/or understood by people living in your home?
2. What was the first language your child spoke?
3. What language does your child use when speaking with you?
4. What language does your child use when speaking with brothers and sisters?
5. What language does your child use when speaking with other family members?
6. What language does your child use when speaking with friends in the neighborhood?

If a parent's answer is a language other than English for any one of these questions, the child is considered potentially LEP and referred for testing. In other words, if a parent answers English and Cantonese to question 1, but English only to all other questions, the child is referred for testing even though he or she speaks only English.

The New York City home language survey is similar but has two additional questions about the language the child reads and the language the child writes in. The questions about language use are a little more discriminating in that they are modified by the clause "most of the time." The scoring process is a little more complicated as well. But the outcome is the same. A child who reads and writes only in English can be classified as potentially LEP simply because English and Cantonese are spoken in the child's home most of the time. In short, the home language survey does not try to determine if the child in question is fluent in English. The wording of the questions are intentionally broad because their goal is to identify children who come from language minority backgrounds, not children who are limited in English.

Norm-Referenced Tests

The overinclusiveness of the home language survey would not be a problem if the subsequent steps accurately identified who was not fluent in English. Unfortunately, they do not. However, were it not for the home language survey, many fluent-English-speaking children would be classified as LEP by the tests that are used.

Children identified by the home language survey must take a standardized test normed on an English-speaking population. The first norm-referenced test they take is an English proficiency test. If they pass the English proficiency test, they may take a standardized achievement test of reading, language, and math in English. If they are Spanish speakers, they may also take these tests in Spanish. These are the same tests LEP students will take later when being evaluated for reclassification as fluent-English speakers.

A point on this normal curve—typically between the 20th and the 50th percentile—is selected as the point at which a student is defined as limited-English proficient. Given that the tests are normed on an English-speaking population, it is not possible for all students to achieve the score that classifies them as fluent-English proficient even if they know no language other than English. If the classification criterion is the 36th percentile, at a minimum 36 percent of the children who take the norm-referenced test will always be classified as LEP even if they are fluent in English and they will never be reclassified no matter how good the program is. This is a mathematical principle.

An important question is why do so many school administrators establish criteria for limited-English proficient students that cannot be met even by all of the English-speaking norming population? One reason is ignorance. Educators seem to have been misled by the constant criticism they receive from intellectuals, policymakers, and reporters who castigate them for such sins as having "only half their students at grade level." In my discussions with school personnel, I have found them to be almost universally ignorant of the fact that nationally it is only possible to have half the student population at grade level.[9]

Another reason that people adopt a standard for LEP students that cannot be met by 36 percent of the students in their school district is confusion. Educators apparently believe that children who score below average—any score below the 50th percentile—are children who are in academic difficulty. Because the home language survey identifies those who are from a home where a language other than English is spoken, many educators believe that setting a standard such as the 36th or 40th or 50th percentile identifies children who are academically in trouble because they come from a home where a language other than English is spoken.

This, however, is wrong. The 40th percentile is that point at which 40 percent of the population scores—no more and no less. All of the students, including those scoring below the 40th percentile, could be extremely smart and highly knowledgeable (for example, by comparison with previous generations). Conversely, all the students, including those scoring above the 99th percentile, could be stupid and ignorant (for example, by comparison with previous generations). Percentiles, or any score computed to differentiate children, simple are not relevatory. They are rank orders, not absolute standards, a fact that is usually not known or, if known, forgotten.

Oral Proficiency Tests

In virtually every school district in the country, students identified by the home language survey as potentially LEP have to take an oral proficiency test and, if they are older (for example, assumed to be literate), a written English proficiency test. Typically a kindergarten and often a first-grade student will take only an oral proficiency test.

On the face of it, oral English proficiency tests would seem to be better than a written test at determining whether a child knows enough

English to function in a regular classroom because the child does not have to know how to read to take an oral proficiency test. Unfortunately, oral English proficiency tests are no better than written English proficiency and standardized achievement tests, and for many of the same reasons.

Oral proficiency tests are known to be unreliable (the same outcome is not reached in subsequent tests of the same child) and invalid (they do not accurately determine who is LEP).[10] Like standardized achievement tests administered to the English-speaking student body and written English proficiency tests administered only to the LEP students, oral proficiency tests cannot tell the difference between a student who does not know English and a student who does not know the answer. They are normed on an English-speaking group and the same arbitrary cutoff points are used.

Several experiments have been conducted in which oral proficiency tests have been administered to English monolingual students. Between 40 and 50 percent of these children who know no language other than English received a score that classified them as limited-English proficient.[11] Other studies have found that the tests classify students as limited in their native language, as well as in English.[12] In addition, the tests do not agree with each other. A student can be classified as limited-English proficient by one test, but not by another.[13]

An experiment in Chicago suggests that even above-average students are not immune from being classified as limited-English proficient by an oral proficiency test. The Chicago Board of Education administered the Language Assessment Scales (LAS) to students who spoke only English and were above the citywide Iowa Test of Basic Skills (ITBS) norms in reading. Almost half of these monolingual, above-average, English-speaking children were misclassified as non- or limited-English speaking. Moreover, there is a developmental trend. Seventy-eight percent of the English monolingual five-year-olds, but only 25 percent of the fourteen-year-olds, were classified as LEP.[14]

Teachers are better than tests in determining whether a child is proficient in English, but even they make mistakes and for the same reasons.[15] Like the tests, teachers can become confused as to whether a child does not understand English or does not know the answer, particularly if the teacher does not know the child very well.

Some school districts conduct dual language testing for Spanish speakers. In New York City, for example, as a result of the ruling in *Aspira of*

New York, Inc., et al. v. *Board of Education of the City of New York, et al.*, students were classified as LEP if they scored below the 20th percentile in English and also scored higher in Spanish than in English. Dual language testing reduces error, but it does not eliminate it because tests in two different languages are not equivalent. The 40th percentile on a Spanish proficiency test is not the same ability level as the 40th percentile on an English proficiency test. For one thing, the tests are normed on different populations—Spanish speakers in the case of a Spanish proficiency test and English speakers in the case of an English proficiency test—and for another it is not yet known how to make questions equally difficult in two languages.

Even if it were, few educators would be able to resist concluding that a language minority student who scores at the 10th percentile in Spanish and the 11th percentile in English is limited-English proficient. Educators are as confused as the general public as to what tests mean and most of them appear to believe that a low score has some absolute meaning.

Even if a language minority student is accurately identified as LEP upon entering the school system, a classification criterion of the 40th percentile guarantees that at a minimum 40 percent of the students will never get reclassified as FEP no matter how good the program is and no matter how proficient they are in English.

Moreover, the cutoff point can be manipulated to produce more or fewer LEP students. If a school district or state changes its criterion from the 20th to the 40th percentile, it can in one fell swoop double the number of limited-English proficient children.

The experience of New York City and New York state illustrate this. The city and state identification standards for LEP students, which were different for Hispanic and non-Hispanic students, changed three times from 1975 to 1999. The initial criterion for Hispanic students was determined in 1975 as a compromise between the defendant city school board that wanted fewer students identified as LEP and the Hispanic plaintiff group that wanted more students identified as LEP. The *Aspira* consent decree established the 20th percentile on the English Language Assessment Battery (LAB) to identify Spanish-speaking students as LEP, but only if the student scored higher in Spanish than in English.[16]

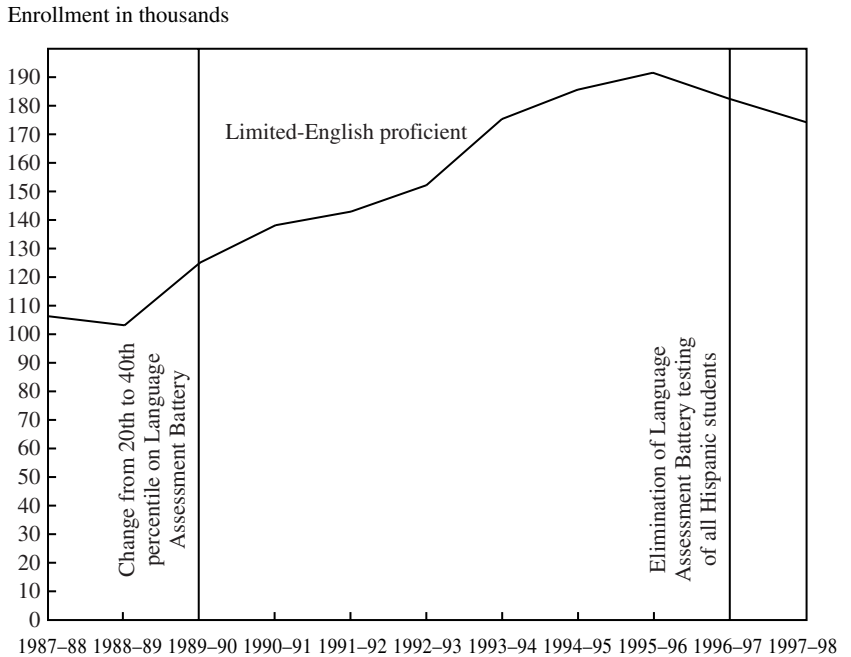
Three years later the city established the 20th percentile for non-Hispanic language minority students in an agreement with OCR. At the

same time, the state of New York was recommending to school districts that they use the 23rd percentile to identify LEP students.

In 1989 the state changed the criterion to the 40th percentile on the recommendation of administrators in the state bilingual education department who believed that the 23rd percentile was too low. They argued that students from language minority families who scored between the 23rd and 40th percentile were having difficulty in English and should be helped. Because 17 percent of the English-speaking population scores between the 23rd and 40th percentile, they, too, would be having trouble in English.

Figure 1 shows that the number of LEP students in the New York City public schools increased by about thirty-five thousand from 1987–88 to 1990–91 as a result of the 1989 change in the LEP standard from the 20th to the 40th percentile. Then the opposite occurred in 1996—the LEP enrollment declined—when the city decided to start using the home language survey as a screening device for Hispanic LEP students. Although

Figure 1. Limited-English Proficient Enrollment Trends in the New York City Public Schools, 1987–88 to 1997–98



the original *Aspira* consent decree required dual language testing (but not a home language survey), sometime during the 1980s the city had stopped using the results of the Spanish proficiency test to classify LEP students. As a result, whether Spanish-speaking students who scored low on the LAB even spoke Spanish was not known. Because LEP enrollment declined by about seventeen thousand students when the home language survey was instituted as a screening device for Hispanic students, at a minimum, about seventeen thousand Hispanic students who had been classified as LEP because of low test scores presumably came from English-speaking families.

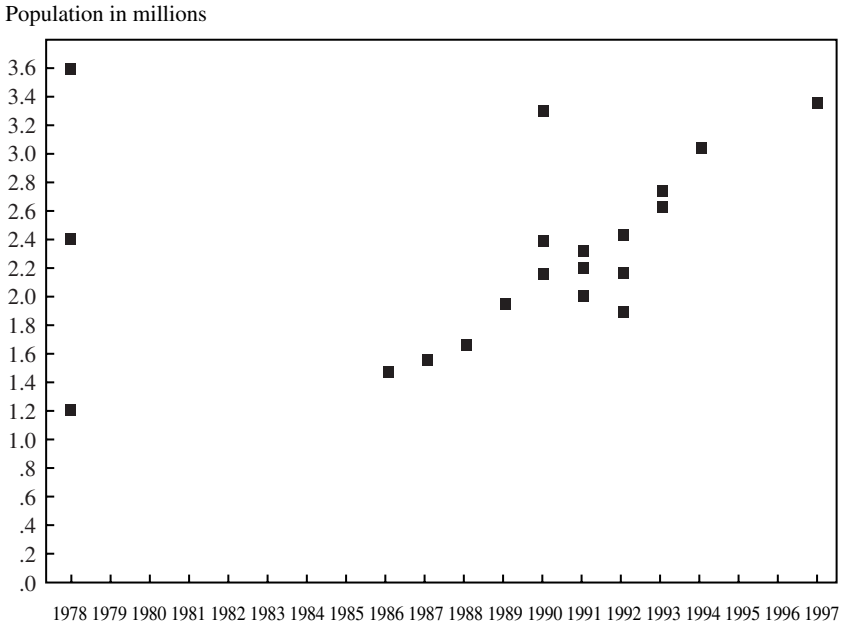
This example illustrates how school districts and states can increase or decrease the number of LEP students simply by changing the standard. Moreover, they do this on a regular basis without any rationale other than the desire to help students who are below average by classifying them as LEP. Almost no one seems to appreciate that being below average, and below grade level, is an affliction of half the student population in the United States and that this is not a reflection of the quality of education or the quality of the students.

In short, the procedures and criteria used by every school district in the United States identify more children as LEP than there are because they cannot tell the difference between a child who does not speak English and a child who has not learned the subject matter and thus does not know the answer. Second, the criterion used—for example, the 40th percentile normed on an English-speaking population—guarantees that at a minimum 40 percent of the students who are administered an English proficiency test will be classified as LEP no matter how fluent they are in English. Unfortunately, there seems to be little or no understanding of this.

The determination of LEP status is made at the local level, but the states and the federal government also make estimates of the number of LEP students. The state governments estimate the LEP population in their state from the numbers given them by their school districts, which are typically based on local standards, although many states recommend standards for school districts to use if they want to receive state funds.

The federal government and researchers also estimate the LEP population. They do this in two ways: first, by aggregating state statistics that typically are based on local statistics, and second, by using decennial census counts. The census counts are based on responses to two questions:

Figure 2. Limited-English Proficient Estimates of Population in the United States, Fiscal 1978–97



languages used in the home and reported ability to speak English. Respondents are asked to rate whether household members speak English "very well," "well," "not very well," and "not at all."

As shown in figure 2, the national estimates that have been generated over the last two decades vary wildly because the standard used varies.[17] If the highest estimate generated in 1978—3.6 million in the Children's English and Services Study—is compared with the lowest estimate in that year, there is a 2.4 million-student disagreement. In later years, there seems to be less disagreement, but this is misleading. All of the estimates are inaccurate because they are all based on the notion that a student from a language minority family who is below average is LEP. Thus, unquestionably, the true number of LEP students is much smaller than the national, state, and local school district estimates, and the only uncertainty is exactly how much smaller. Unfortunately, critics of the process do not understand this.

Kris Anstrom criticizes the census counts because they do not take into account reading and writing in English.[18] But taking into account read-

ing and writing in English would decrease, not increase, the reliability of the designation because household members would have even greater problems estimating proficiency in reading and writing than in speaking. It is easier to determine how household members speak English than how they read or write it. Moreover, what does it mean to not read or write English very well? My fellow professors would probably designate half of their English monolingual college students as not reading or writing English very well. In short, this is an area where adding English proficiency skills would make the evaluation less, not more, reliable. But what is particularly troubling is the unquestioning belief, common in the social sciences, that greater specificity and higher standards make the determination more reliable. They do not.

Oona M. Cheung, Barbara S. Clements, and Y. Carol Miu recommend a national standard for limited-English proficiency that includes a family language other than English and whether the student is below that of an academically successful peer with an English language background.[19] This guarantees that at a minimum half of all language minority students will be classified as LEP because "academically successful" is typically defined as being at grade level, and only half of all students can be at grade level.

Cheung, Clements, and Miu further recommend a definition for fluent-English speaking students that is basically reading, listening, writing, and speaking at the age and "grade appropriate level." Once again, this guarantees that half of all language minority children will be classified as LEP and never reclassified given that only half of all children in the United States can be at a "grade appropriate level." Thus, Anstrom is wrong in her belief that a national standard would provide a more accurate estimation of the LEP population. The only thing a national standard might possibly do is provide a consistent inaccurate estimate for national programs.

The Recommended Program

From the beginning, Title VII has favored bilingual education. In the original 1968 Bilingual Education Act, bilingual education was one of several programs that a school district could adopt to help LEP students, although it was favored among second language programs. By 1974 bilin-

gual education was not just a favored option, but also almost a requirement for receiving Title VII funds. The Reagan administration hardly made a dent in this. The 1984 reauthorization under the Reagan administration added only that in some school districts bilingual education might be impractical because of small numbers of LEP students of the same language or because of unqualified staff. The 1984 reauthorization also specifically declared it to be the policy of Congress "to encourage the establishment of special alternative instructional programs" defined as

> programs of instruction designed for children of limited English proficiency in elementary and secondary schools. Such programs are not transitional or developmental bilingual education programs, but have specially designed curricula and are appropriate for the particular linguistic and instructional needs of the children enrolled.[20]

The act allowed up to 4 percent of overall funds (or up to 10 percent if more than $140 million was appropriated in a single fiscal year) to go to these special alternative instructional programs that did not use the native tongue. However, 75 percent of funds for instructional programs (Part A) were still allocated to transitional bilingual education programs and 21 percent to maintenance bilingual education under the Reagan administration.

The 1988 reauthorization under the Reagan administration continued to favor bilingual education but included further provisions to allow school districts to use different approaches to the education of LEP children. Part A authorized 75 percent of total grant funds to school districts for transitional bilingual education but increased to 25 percent the amount of grant funds that could go to special alternative instructional programs that did not use the native tongue. In addition, a three-year limit was placed on a student's participation in a transitional bilingual education program or in alternative instructional programs, although under special circumstances, a student could continue in a program for up to two additional years. Thus, the Reagan administration only slowed the bilingual education juggernaut.

The 1994 reauthorization under the Clinton administration earmarked $215 million for fiscal year 1995, gave funding priority to programs that provide for the development of bilingual proficiency both in English and another language, but kept the 25 percent maximum allocation for programs that did not use the native tongue. What is different from the

reauthorizations under the Reagan administration, however, is that the Democrats added the requirement that alternative educational programs must be justified either by small numbers of language minority students (although no minimum is mentioned) or a lack of teachers with native language skills (after demonstrating an effort to obtain such).

The 1999 Clinton administration proposal continues to favor bilingual education and even cites the 1998 National Research Council review to support its position.[21] But it also opened a small crack in the bilingual education armor by its many references to the goal of learning English and achieving to high state standards and by its proposal to change the name of the National Clearinghouse for Bilingual Education to the National Clearinghouse on the Education of Children and Youth with Limited English Proficiency.

Although Title VII favors bilingual education, not an ounce of evidence in 1968, and no consistent evidence in the thirty years since, proves its superiority. Moreover, only Spanish speakers seem to receive it, although no recognition is made of this troubling fact in either the educational literature, the legislation, or the regulations.

The theory underlying bilingual education is the facilitation theory, developed by Jim Cummins, which argues that children will be more cognitively developed if they learn to read and write in their native tongue before they learn English. This theory was created a decade after bilingual education was implemented in the 1960s. It is a limited theory, however, because it ignores the issue of the great variation in written language. In particular, it is silent on how to teach Asian children to read and write in their native tongue and why doing so is desirable. The majority of Asian languages use an ideographic system of writing, instead of an alphabet, and have no similarity to English in appearance, thus reducing the number of skills that are transferable, such as sight recognition of words, sounding out of words, and so forth. These languages also take much longer to master than English. In other words, learning to read in the native language, if it is ideographic, may be harder than learning to read and write in the second language, if the latter is English or another phonetic, alphabetic language. As a result, I have not yet found any nonalphabetic bilingual education programs that teach initial literacy in the native language, although many are taught in self-contained classrooms, are called bilingual education, and receive bilingual education funding.

I also have not found any non-Roman alphabet bilingual education programs, even if the alphabet is phonetic (for example, Hebrew, Arabic, the Indian dialects, Russian, and Khmer), that teach initial literacy in the native language. Educators apparently believe that it is too difficult and confusing to teach initial literacy, particularly to young children, in a language with a completely different alphabet from English. For one thing, the transferability of the skill of sight recognition of words is diminished. For another, some of these dialects have more than one alphabet and the rules regarding the combination of consonants and implied or missing vowels are different from English and other Roman alphabet languages so that teaching such languages in an already full day is too formidable a task.

None of the federal and state laws nor bilingual education theory recognizes this. Title VII notes the impracticality of bilingual education for some language groups, but only "due to the presence of small numbers of students of a particular native language or because personnel who are qualified to provide bilingual instructional services are unavailable."[22] To my knowledge, I am the only academic to question the applicability of the facilitation theory to the non-Roman alphabet languages. I suspect this is because to acknowledge that the most academically successful of the language minority children—the Asian students—are instructed almost completely in English would raise too many questions regarding the efficacy of native tongue instruction.

What little research evidence there is indicates that bilingual education, defined as native tongue instruction and native tongue literacy transitioning to English, is, on average, the least effective approach to educating limited English children if one's goal is English language achievement and subject matter knowledge that a student will be tested on in English. Table 1 shows outcomes reported in seventy-two studies that met the standards for a scientific study; that is, they had a treatment and control group and, if there was no random assignment, controlled for pretreatment differences.[23] The effect of transitional bilingual education is compared with "submersion" or doing nothing, English as a Second Language, structured immersion, and maintenance bilingual education—on second language (usually English) reading, language, and mathematics.[24]

Table 1 also shows the effect of structured immersion compared with ESL pullout. Studies are repeated in more than one category of outcome if they had different outcomes at different grade levels or for different

Table 1. Methodologically Acceptable Studies Demonstrating Program Superiority, Equality, or Inferiority by Achievement Test

Percent unless otherwise indicated

Outcome	*Reading*[a]	*Language*	*Math*
TBE versus submersion (do nothing)			
TBE better	22	7	9
No difference	45	29	56
TBE worse	33	64	35
Total number of studies	60	14	34
TBE versus ESL			
TBE better	0	0	25
No difference	71	67	50
TBE worse	29	33	25
Total number of studies	7	3	4
TBE versus submersion/ESL			
TBE better	19	6	11
No difference	48	35	55
TBE worse	33	59	34
Total number of studies	67	17	38
TBE versus structured immersion			
TBE better	0	0	0
No difference	17	100	63
TBE worse	83	0	38
Total number of studies	12	1	8
Structured immersion versus ESL			
Immersion better	100	0	0
No difference	0	0	0
Total number of studies	3	0	0
TBE versus maintenance bilingual education			
TBE better	100	0	0
Total number of studies	1	0	0

Source: Christine H. Rossell and Keith Baker, "The Educational Effectiveness of Bilingual Education," *Research in the Teaching of English*, vol. 30, no. 1 (February 1996), pp. 1–74.

Note: Studies are listed in more than one category if there were different effects for different grades or cohorts. There were 72 methodologically acceptable studies. TBE = transitional bilingual education; ESL = English as a Second Language.

a. Oral English achievement for preschool programs.

cohorts.[25] Those not in the table are excluded because they did not assess alternative second language learning programs or they did not meet the methodological criteria.[26]

Table 1 indicates the percentage of studies showing a program to be better than the alternative it is compared with, the percentage showing no difference, and the percentage showing the program to be worse than the alternative it is compared with. This is repeated for each achievement test—

reading, language, and math. The total number of studies assessing the particular achievement test for each category of comparisons is shown below the percentages.[27]

TBE VERSUS SUBMERSION. Table 1 indicates that for second language reading, 22 percent of the studies show transitional bilingual education (TBE) to be superior, 33 percent show it to be inferior, and 45 percent show it to be no different from submersion (doing nothing).[28] Altogether, 78 percent of the studies show TBE to be no different from or worse than the supposedly discredited submersion technique.

In a standardized achievement test of language, a test of a student's understanding of grammatical rules, transitional bilingual education does even worse than it does in reading. Seven percent of the studies show transitional bilingual education to be superior, 64 percent show it to be inferior, and 29 percent show it to be no different from submersion. Altogether, 93 percent of the studies show TBE to be no different from or worse than doing nothing at all.

These more negative findings for language than for reading suggest that a child is less dependent on school for many of the skills learned in reading—decoding, vocabulary, and understanding concepts—than for grammar. The fine rules of grammar apparently are learned mostly in school and, because they are more complex, are more influenced by school time on task.

In math, 9 percent of the studies show TBE to be superior, 35 percent show it to be inferior, and 56 percent show it to be no different from TBE. Altogether, 91 percent of the studies show it to be no different from or worse than the supposedly discredited submersion technique in developing math proficiency.

TBE VERSUS ESL. Although many so-called submersion situations probably have an ESL program in which students are pulled out of the regular classroom and taught English in small groups for a period a day or a few times a week, it is generally not specified in the evaluations. Nevertheless, many of the studies classified above as submersion could include an ESL pullout component. In seven studies, transitional bilingual education is specifically compared with reading achievement in the regular classroom with ESL pullout. None of these studies shows TBE to be better than ESL pullout in reading. Five studies (71 percent) show no difference between transitional bilingual education and ESL in reading, and two studies (29 percent) show TBE to be worse than the regular class-

room with ESL pullout. Of the three studies that examined language achievement, none showed TBE to be superior, two showed no difference between TBE and ESL, and one showed TBE to be worse. Of the four studies that examined math, one showed TBE to be superior, two showed no difference, and one showed TBE to be worse.

TBE VERSUS SUBMERSION/ESL. Because the suspicion is that many, if not most, of the so-called submersion alternatives had an ESL component, also shown in table 1 are the outcomes for a category that combines submersion and ESL studies. Because of the small number of studies that specifically examine ESL pullout, virtually no difference turned up in the findings—81 percent of the studies show TBE to be no different from or worse than submersion or ESL in reading, 94 percent show TBE to be no different from or worse than submersion or ESL in language, and 89 percent show TBE to be no different from or worse than submersion or ESL in math.

TBE VERSUS STRUCTURED IMMERSION. Table 1 also compares TBE with structured immersion. Most of these studies are of the Canadian immersion programs that come in several carefully documented types—early immersion (late bilingual), delayed immersion (early bilingual), dual immersion, and so forth. In many cases, the programs had to be translated using U.S. terminology. Twelve studies had reading outcomes, one study had language outcomes, and eight studies had math outcomes. No studies showed TBE to be superior to structured immersion in reading, language, or math. In reading, 83 percent of the studies showed TBE to be worse than structured immersion and 17 percent showed no difference. In language, the one study showed no difference. In math, five studies showed no difference and three studies showed TBE to be worse than immersion.

STRUCTURED IMMERSION VERSUS ESL. Three studies compared structured immersion with ESL specifically. These studies all showed structured immersion to be superior to ESL in reading.

TBE VERSUS MAINTENANCE BILINGUAL EDUCATION. The final category in table 1 compares transitional bilingual education with maintenance bilingual education. This study by Marcella Medina and Kathy Escamilla showed that transitional bilingual education produced significantly higher English reading achievement than maintenance bilingual education, although the authors did not acknowledge it.[29]

Confronted with the kind of evidence presented in table 1, the advocates of bilingual education have sometimes contended that the issue is

learning in a language, not learning a language. These data, however, do not show TBE to be superior in either learning a language or learning in a language—in this case, math. Moreover, no research evidence exists on the effects of TBE on learning other subjects such as geography, social studies, and history because standardized achievement tests are not given in these content areas. Thus, any assertion regarding the superiority of TBE in these areas is anecdotal. Moreover, the math findings for TBE suggest an important problem: Subject matter is taught in the native tongue, but the student is tested on his or her understanding of that subject in English. For many students the difficulty of having to translate what was learned in another language could be great enough that the subject matter lost in the translation may equal or surpass what is lost in submersion before the second language is mastered enough to understand subject content. However, the solution is not to test LEP children in their native tongue because the primary goal of Title VII is for students to reach the highest level in a subject in English that they are capable of.

An important limitation of these studies is that they are short term—even the longest one is only five years and the average is one to two years. If the superiority of bilingual education is evident only after six or seven years, these studies are not a good test of the efficacy of bilingual education. However, because the two scientific longitudinal studies—the J. David Ramirez and others study and the El Paso Independent School District studies—did not find bilingual education to be superior and the latter found it to be inferior, it is safe to say there is not yet any evidence that bilingual education is superior even after many years.[30]

But it is important to point out that some scientific studies find bilingual education to be superior. These superior bilingual education programs seem to have two things in common—the native tongue instruction does not reduce the time on task in English (either because the school day or year is expanded or because nonacademic subjects are eliminated to accommodate the native tongue instruction) and the students make a quick transition to English reading and writing. Because learning to read in one's native tongue is easier than in a foreign language, programs that include native tongue reading instruction can apparently be very good programs so long as there is no reduction or only a brief reduction in English language instruction. Moreover, the positive effect of native tongue reading instruction is most likely if there is no time-consuming native tongue writing instruction. Unfortunately, bilingual education programs

typically reduce English language instruction to include copious amounts of time on native tongue literacy. Students can get stuck trying to master native tongue literacy and spend their entire elementary career in a bilingual education program.

What Is the Goal?

When confronted with the evidence that bilingual education is on average less effective in teaching English and subject matter that one will be tested on in English, supporters of bilingual education have sometimes countered that the goal of bilingual education is bilingualism, not English language development. However, the theory supporting bilingual education does not have bilingualism as a goal. The argument is that learning to read and write in one's native tongue produces the greatest cognitive development and that will be evidenced in the second language (that is, English). Bilingualism may be a by-product of that process or even a means of achieving that goal, but it is not the goal.

The federal legislation has been somewhat conflicted on this issue. Knowing more than one language has been a goal of Title VII, although not the primary goal, from the beginning. The 1994 reauthorization states that one of the purposes of Title VII is "developing bilingual skills and multicultural understanding," and the 1999 Clinton administration proposal continues this goal.[31] To achieve this, Title VII includes a program called Foreign Language Assistance, which funds programs of foreign language study for all elementary and secondary school students. The problem is that this program dilutes Title VII's limited resources and changes its purpose from a compensatory education program for children who do not speak English to an enrichment program that might benefit affluent, upper-middle-class children who want to learn another language. This seems to be an unfortunate waste of scarce resources.

Conclusions

I have four general conclusions about the major problems with Title VII and four recommendations for changing the statute.

The Target Group

Accurately identifying who is LEP is not possible.

As the research and statistics on the wildly fluctuating LEP population in the United States suggest, the category is unreliable and the percentages are a function of the criterion used. Different school districts use different criteria, and no logic exists to any of them, which is probably why there is no agreement. The same school districts change the criteria over time and thus are able to increase or decrease at will the number of LEP children they have. Nor is it possible to establish a criterion that would prevent this.

Because the underlying intent of Title VII is to help children who enter school from a home where there is less English than in other homes because the family speaks a language other than English, why not make this the criterion for eligibility? Determining language minority status is easier than ascertaining whether a child is limited-English proficient, because the former can be revealed by a home language survey. The home language survey could be administered in two stages:

1. Administer a short home language survey in English to all entrants into the school system to find children who come from a family that speaks a language other than English (families who cannot respond to the survey in English are automatically classified as language minority).

2. Administer a longer home language survey to language minority families to determine how limited the parents feel their child is in English and how proficient they think their child is in the non-English language (families who cannot respond to the survey in English are automatically classified as language minority).

Children are identified as language minority background from the parents' inability to respond to the survey in English or from the parents' responses to the survey. The primary source of identification would be the first-stage survey, whereas the second-stage survey would be used for clarification. Children would never be reclassified because their identification would not be as limited-English proficient, but as language minority, a classification that is not dependent on misleading test scores. They would have this identification all their school career, thus avoiding the impossible task of deciding when a child is, or is not, LEP.

Some error will occur in determining who is language minority, but far less than in determining who is LEP, because the former would not

depend on tests. Moreover, if children are assigned to short-term structured immersion classes or regular classrooms with extra help as is happening in California with Proposition 227, the negative effect of misclassification would be minimal.

For those who think that to stop collecting LEP statistics and reclassification rates is extremely dangerous, I can only emphasize how dangerously misleading these statistics are. The number of LEP students is a function of the criterion used, which differs from school district to school district. In every school district in the country, the criteria include in the LEP category large numbers of fluent-English speakers and exclude smart children who need extra help with their English. Furthermore, the reclassification rates that school districts and programs are held accountable for are not achievable by all students no matter how good the program is. Moreover, because they are annual rates, they are only a snapshot in time, which makes them even more meaningless. What is the point of collecting and making policy decisions on statistics that are meaningless? Using these statistics is dangerous because it lulls school districts, states, and the federal government into thinking they are making rational decisions based on sound data when they are not. Because the enterprise cannot be rationalized, it should be abandoned, and a simpler, more reasonable one adopted.

Title VII funds would go to school districts based on the number of language minority children they have as determined by the home language survey. School districts would be eligible for additional funds under Title I if these children were also poor and low achieving.

Recommendation 1. The phrase "limited-English proficient" or "limited-English speaking," and any other phrase referring to the ability of a child in English, should be stricken from Title VII and the rest of the Elementary and Secondary Education Act, and the term "language minority" should be substituted.

The Program

No consistent research evidence supports bilingual education as a superior alternative for educating supposedly LEP children. Moreover, it is an especially risky program given how difficult accurately classifying a child as LEP is. A child who does not speak Spanish could be placed in a Spanish bilingual program and taught in a language he or she does not

understand or speak, supposedly to improve his or her English. More commonly, a fluent-English speaking child could remain in a Spanish bilingual program his or her entire elementary career because he or she cannot score above the exit criterion. Even if the facilitation theory is correct, being taught in Spanish cannot help a child's English language achievement if Spanish is not his or her native tongue. Nor is being taught in Spanish helping the English language achievement of a Spanish-speaking child who already knows English.

A prudent approach to the conflicting research and controversy would be to eliminate program recommendations from Title VII. A particular program need not be specified to help language minority children, and it would be prudent to refrain from doing so, given how little research backs up any particular one, how controversial the research and its conclusions are, and how quickly the sands of social science research can shift. Authorizing money for children who might need extra help is one thing; that is based on common sense and the heart. However, specifying exactly what kind of extra help this should be is foolhardy when social science research findings change and school districts can modify any program any way into whatever makes sense for them.[32] Moreover, Title VII seems to be the only part of the Elementary and Secondary Education Act that calls for a specific, and controversial, educational program that is opposed by most Americans, most language minority parents, and most Hispanics.[33]

Recommendation 2. Eliminate all references to bilingual education programs and to developing bilingual skills. Title VII should refer only to programs that provide high-quality instruction to language minority children to meet the performance standards expected of all children.

The Legislation

Title VII is badly written. It is far too long and full of redundant and fragmented programs and agencies as well as self-evident, repetitive statements and assertions such as

> the assistance provided under the application will contribute toward building the capacity of the applicant to provide a program on a regular basis, similar to that proposed for assistance, which will be of sufficient size, scope, and quality to promise significant improvement in the education of students of limited-English-proficiency, and that the application will have

> the resources and commitment to continue the program when assistance under this subpart is reduced or no longer available. (Section 7116 (h) (5))

This could be reduced to one sentence: The local educational agencies (LEAs) will be expected to continue systemwide enhancements funded under Title VII. Everything else in that paragraph has already been stated a number of times before and after that in the text of the legislation.

Title VII is also full of false statements about the superiority of bilingual education. In many cases, assertions are made for which absolutely no supporting research exists. Many statements are puzzling and why they are in the legislation is a mystery, unless the goal of each administration is simply to add text.

The Clinton administration's 1999 proposal continues this tradition. As occurs often in journalism and in educational policy, the opinions of intellectuals and academics are confused with research findings. Not only are many of the same spurious claims made, but new ones have been added. The Clinton administration included the statement that educational technology has the "potential" to improve the education of language minority and limited-English proficient students and therefore the federal government should foster development of that technology. But many things have the potential to help many students. What is the point of including this statement in Title VII?

Foreign language training is another one of these issues that does not belong in Title VII. The original goal of Title VII was to help immigrant children gain proficiency in English. Foreign language training may not even belong in ESEA, which, after all, is a compensatory education bill. The Emergency Immigrant Aid Act program is also redundant if the definition of the target population is changed to language minority. Title VII needs to be simplified, not further complicated. Unfortunately, the Clinton administration has not made any progress in this; it has made matters worse.

Recommendation 3. Reduce Title VII to a few pages stating that:

—Money is to be allocated to school districts based on the number of children from language minority families, and this money should be spent on educational programs, additional staff, and in-service teacher training that have a high likelihood of improving their education.

—A certain percentage of this money can be spent on schoolwide programs and districtwide programs.

—State and local funds are not to be reduced if a Title VII award is made.

—The school districts should be required to show the number of language minority children that they have and that they spent the money on extra staff, in-service training, and programs that have a likelihood of improving the education of language minority children.

—Title VII will fund research and evaluation conducted under the auspices of the federal government, but school districts are not required to do it themselves, although they must collect and provide the data to be used in the evaluations.

—Title VII will fund the National Clearinghouse on Bilingual Education (or the proposed National Clearinghouse on the Education of Children and Youth with Limited English Proficiency).

The Research

The quality of research in this field is terrible. Not all of this is the fault of the school districts that receive funding, however. Doing good research is extremely difficult and expensive, and few people are trained in it. Title VII, for example, includes a research and evaluation component—a minimum of 5 percent of the funds are reserved for program evaluation by school districts. Unfortunately, the 1994 reauthorization appears to advocate an invalid research design when it suggests that school districts compare the academic achievement of LEP students with that of non-LEP students. This is the wrong design and the wrong comparison because LEP children are defined by their low achievement. If they are high achieving, they are not classified as LEP, even if they come from a language minority family. Simple examinations of the two would compare low-achieving children to high-achieving children and be meaningless.

Other misleading statements are in Title VII. Section 7115 states that the secretary shall terminate grants to eligible entities under this section if the secretary determines that "the program evaluation required by section 7123 indicates that students in the program are not being taught to and are not making adequate progress toward achieving challenging State content standards and challenging State student performance standards; or in the case of a program to promote dual language facility, such program is not promoting such facility." The problem is that only an expensive and sophisticated evaluation would be able to do this, and few

individuals have those skills. The statement also implies that this is a simple process of looking at progress over time. It is not.

Because few people in federal, state, or local government understand research design, federal program legislation should not include statements about the characteristics of program evaluations because they are almost always wrong.

But school districts should not be required to evaluate their own programs because they do not know how to do it. Millions of dollars are wasted every year on local evaluations that are not scientific and cannot determine program effectiveness. The federal government should require only that school districts that accept Title VII funds keep and be willing to turn over anonymous student-level data upon request for program evaluation by trained and experienced researchers. The legislation should not specify more than that because it is too often wrong or misleading.

Title VII could offer extra funds to school districts willing to conduct experiments that would enable researchers to understand the effects of different programs. Such an experiment would begin with random assignment of Hispanic language minority children to different instructional environments such as a mainstream classroom with ESL pullout, a transitional bilingual education program, and a structured immersion program. Random assignment would eliminate the need for a pretest (a measure of academic achievement or intelligence before enrollment in the program). A pretest is extremely difficult to obtain because most children are tested in the spring, after they have spent a year or more in a program or school. With random assignment, a pretest need not control for the ability differences of students that might cause them to be chosen for various programs because they are not selected for their characteristics. With random assignment, any differences in outcomes will be the result of the program instead of the characteristics of the children that caused them to be assigned to the program.

For those who worry about the issue of parental consent for random assignment, the California experience with Proposition 227 indicates that most parents look to the schools for guidance and will do what they say. If the school administration tells them their child should be in a traditional bilingual education program, most Hispanic parents will agree. If, a day later, the school administration finds that it does not have enough students assigned to the traditional bilingual education to have a whole class, the same Hispanic parents will agree to have their

child taught in all-English classes. Thus, random assignment is likely to be accepted by Hispanic parents so long as it has the support of the school administration.

But random assignment is not a solution to determining program effectiveness if the children in the programs are tested at different rates. If there are differential testing rates, the program with the lower testing rate will always, all other things being equal, have an unfair advantage over the program with the higher testing rates because the students who would score the lowest are deemed "not ready" to take the test and are excluded from testing.

A bias currently exists in the program evaluations because the students in all-English programs are tested at higher rates than the children in bilingual education programs. Even if there is random assignment to the two programs, evaluations comparing all-English programs with bilingual education programs will be biased against the all-English programs. Therefore, a valid research design must include not only random assignment to the two programs, but also universal testing. To know the true effects of programs, all children must be tested with standardized achievement tests even if they do not know any English.

The Clinton administration's proposal for testing in English after three years of education in the United States is an improvement over the 1994 reauthorization, which had no timetable for testing in English. It is still, however, naïve in failing to understand that, if no testing is done at the beginning of a program, the ability of a statistical evaluation to determine the effect of a program will be severely limited. This is particularly true because of much lower testing rates for students in bilingual education than in all-English programs in their first three years.

Although the Clinton administration's 1999 proposal is also an improvement over the 1994 reauthorization by explicitly stating that the goal of Title VII is to learn English and to achieve to challenging state content and performance standards and by requiring annual testing, it reduces the positive effect of this by encouraging native tongue testing if that would "more accurately reveal" what a child knows. The Clinton proposal specifically singles out Spanish-speaking LEP students when it specifies there should be "tests written in Spanish for Spanish-speaking students with limited English proficiency, if those tests are more likely than tests written in English to yield accurate and reliable information

on what those students know, and can do, in subjects other than English."[34] The problem again is that the LEP students in bilingual education programs will be taking tests in Spanish, if that is their native tongue, and the LEP students in all-English programs will be taking them in English, thus biasing any program evaluation in favor of bilingual education. Given that no one believes that achievement in Spanish is the ultimate goal, valid program evaluations must be of English language achievement with equal percentages of students in each program taking tests in English.

Recommendation 4. School districts should be required to collect basic statistics for individual language minority students and to provide it on demand to federal researchers, but they should not do the program evaluations themselves. The federal government should provide a fiscal incentive to school districts to do random assignment to programs and universal testing. The school districts would be required to provide achievement and other data on individual students to the federal government so that experienced researchers could do statistical analysis to determine program effects.

Given the inconsistency in the research and the lack of strong effects in any direction, assigning students randomly to alternative treatments would not be unethical. Random assignment and universal testing would go a long way toward producing a research study that could do this, and the federal government could make a huge contribution by including funding for it in Title VII.[35]

Title VII can be improved by changing the target group from limited-English proficient students to language minority students; eliminating the support for a specific, and controversial, program—bilingual education—that is only implemented for Spanish speakers; reducing the redundancy and turgid prose of the legislation itself; encouraging school districts to do random assignment and universal testing by awarding additional funds for these experiments; and funding national research studies. In general, Title VII should be simplified—it has grown from six pages to thirty-five pages in thirty years and has become only more fragmented and irrational. Unfortunately, the Clinton administration's proposal worsens this problem by adding verbiage. The bill as it stands now could be cut to less than ten pages and lose nothing of substance.

Comment by Catherine E. Snow

In these comments on Christine H. Rossell's paper about the federal bilingual education program, I will respond specifically to her principal conclusions and her four recommendations. I acknowledge the line of reasoning that underlies those recommendations, and in some cases I agree with both the reasoning and the conclusions. Who could disagree, after all, that the language of federal legislation is often characterized by poor writing, redundancy, and irrelevancies? However, Rossell and I differ in basic orientation to the issues and in the knowledge bases we bring to this topic. Furthermore, Rossell makes some basic misstatements in her paper, misstatements that are in some cases so egregious that it would be remiss of any discussant to let them pass unchallenged.

Political Science versus Developmental Psychology

Rossell is a political scientist whose major focus in her discussions of bilingual education seems to be to identify obstacles to the rational implementation of policies and unpredictabilities in the process by which policies have been decided upon. I am a developmental psychologist, interested in the conditions under which children develop to their full capacities, and in particular the conditions under which the largest possible percentage of children will successfully learn to read. Teaching children to read is, in my view, the most important task schools face, because failure in reading will doom children in most other school-related domains. Furthermore, learning to read is the first serious task children face in school, and success or failure in this task has a long-term impact on children's views of themselves as students and learners.

Thus, my primary interest is how to ensure success in reading—or, to put the issue in the public health perspective adopted by the Committee on the Prevention of Reading Difficulties in Young Children, which I chaired, how to minimize risk of reading failure. Learning to read is, in a sense, a natural act for a normally developing child in a literate culture with good schooling. Factors that enhance the risk that children will not follow this normal developmental pathway, as reviewed in the committee's report, include living in poverty, arriving at school not speaking English, attending schools with generally poor performance, and having

poorer than average language and preliteracy skills on school entry.[36] The superordinate prevention strategies identified by the committee were: (1) ensuring that all children have access to excellent, language- and literacy-rich, preschool environments, and (2) ensuring that all children have access to excellent reading instruction.

What constitutes excellent reading instruction? The answer the committee gave to that question derives from its definition of reading: Reading is the act of constructing meaning from print. In other words, good readers of English use the alphabetic principle—that letters map in regular ways onto sounds—fluently and automatically to access word meaning and construct textual meaning. Excellent reading instruction must give children opportunities to grasp the alphabetic principle, in the context of a continual focus on reading for meaning, and must ensure that children have enough practice in reading to become automatic at the identification of words and to become skilled users of strategies for comprehension.

So what is the relevance of this view of reading to policies for bilingual education? The committee report concluded that the relevance was great—that this view of reading dictated that instruction that involves teaching children to read in a language they do not understand cannot constitute good instruction. Getting to the meaning is the whole point of reading, and children (particularly children from homes with low levels of parental education and little literacy exposure) can be expected to become confused and discouraged if expected to learn to read meaningless words. Furthermore, recognizing the words one starts out reading is an enormous support to the acquisition of the details of the alphabetic principle for most children. Children who cannot use meaning to support their reading during the initial stages of mastering the alphabetic principle are, inevitably, facing a much harder task than children being taught to read meaningful words.

At slightly more advanced stages, as well, knowing the language in which one is reading is crucial to developing sophisticated comprehension strategies and to spending enough time on task to develop fluency and automaticity with word recognition and the implementation of comprehension strategies. Enthusiasm for the practice of reading is a major input to developing fluency and success. Young children are unlikely to be enthusiastic or to engage in the practice they need if they are reading mostly incomprehensible texts.

Thus, the report on preventing reading difficulties made the following recommendation concerning reading instruction for children who arrive at school not knowing English and not knowing how to read: that such children be taught to read in their native language if that is feasible (that is, if instructional materials and qualified teachers are available, and if such children are sufficient in number to justify native language reading instruction), while also receiving instruction focused on building oral proficiency in English, and that transition to English reading instruction take place only after a reasonable level of oral proficiency in English had been achieved. For children for whom these conditions do not hold, the committee recommended that formal reading instruction be postponed until some reasonable level of proficiency in English has been achieved and that the schools design programs to provide rich English language and preliteracy environments for such children, to ensure the most rapid possible acquisition. Furthermore, research should be undertaken to address the question of what constitutes "sufficient oral proficiency in English" to serve as a safe basis for reading instruction.

Two important points must be made about these recommendations. First, they are designed to reduce risk of reading failure. The committee recognized that children can learn to read in a language in which they have very low oral proficiency. Many millions of children have demonstrated that this is possible, not just in U.S. schools but also in colonial school systems across the world. However, such an approach is inherently more risky—more children will fail under this approach than if taught to read under conditions where meaning can be used as a support. The level of school failure that is normal in third-world countries with universal second language reading instruction simply cannot be tolerated in the United States.

Second, this recommendation applies only to children who cannot read on arrival at school. The committee did not make any recommendation concerning the value of bilingual education for older arrivals who can already read in their native language, nor for precocious readers who enter first grade reading fluently in their native language. However, much research suggests that such children can learn to read in the second language relatively quickly and then use second language literacy as a resource for second language oral acquisition.

Specific Claims That Deserve Response

In responding to quotations from Rossell's paper, I recognize that some are less than central to the arguments she is making but nonetheless feel it is imperative to note when they are misleading or unfounded. In each case, her statement is provided in italics.

The U.S. Department of Education has enthusiastically embraced bilingual education. This claim is based on the distribution of funds available through the Bilingual Education Act and its successors. The largest part of these funds has typically been reserved for bilingual programs. But the total funding available has never been sufficient to serve most of the children who would have qualified for and benefited from bilingual programs. Thus, one could with as much justification argue that the Department of Education has failed to support bilingual education enthusiastically.

There is no treatment called "bilingual education" that is implemented in the same way and understood to be the same thing by everyone. This claim is true. The three program types noted in Rossell's paper fail to reflect the full range of program types or of variation within types. A major problem with the evaluations of bilingual programs, which as she notes are often exemplars of the worst in educational research, is that the variation within program type, which is often as great as that between program types, has not typically been taken into account. Many of the studies Rossell cites in her research synthesis would be rejected as fatally flawed by others who demand at least some information about classroom practices before evaluating a program's outcomes.

A major issue in thinking about these program labels is to realize that program labels do not distinguish educational experiences of children efficiently. The program labels mostly refer to classroom configurations—what mix of languages do the children in the classroom speak as native languages, and what language capacities are expected of the teacher? Thus, a transitional bilingual program is defined by the presence of children dominant in one language and a teacher with at least some competence in that language (and, one hopes, in English as well). Details of pedagogy, of distribution of instructional activities over the two languages, and of curriculum are all left undefined by the program category.

"Structured immersion" is described by Rossell as "all-English" instruction in a self-contained classroom containing (exclusively) English language learners. Here, again, classroom configuration rather than explicit models for what to teach and how to teach it defines the program type. No one knows what should be going on in a structured immersion classroom, only that it should be going on primarily in English. Structured immersion is the program type that has been mandated in California by Proposition 227 and that is being proposed for Massachusetts and other states under Proposition 227-like legislation. No one has been trained in how to teach children in structured immersion settings. All that is prescribed is to restrict the use of the children's native language to a minimum. In many parts of the country, places where all the English language learners come from the same language background, structured immersion may not differ much in practice from programs labeled as transitional bilingual in which, as observational research shows, the children's native language may also be used rather little.

What does this mean? At a minimum, it implies that evaluations that compare programs called structured immersion with programs called transitional bilingual, but that provide no data on pedagogical activities and language use in the classroom, are worthless. More depressingly, all the complaints about the level of academic and English language accomplishments of children in transitional bilingual programs almost certainly will be replicated for structured immersion programs.

Only the Spanish speakers . . . are receiving bilingual education through native tongue instruction, according to the theory. This claim reflects Rossell's beliefs that children who arrive at school speaking languages other than Spanish do not receive reading instruction in their native languages and that native language reading instruction is the key defining element of bilingual education. Both these claims could be challenged—on factual and theoretical grounds, respectively.

The majority of children receiving "bilingual education according to the theory" in the United States are Spanish speaking. Also, the majority of language minority children eligible for bilingual education but not receiving it "according to the theory" are Spanish speaking. Because 85 percent of non-English-speaking children in the United States speak Spanish, bilingual education not surprisingly has been developed largely with their needs in mind. Furthermore, the proximity of Spanish-speaking countries and the availability of Spanish language materials facilitate the

introduction of initial Spanish reading instruction, which is undeniably more widespread than initial instruction in other languages. The force of this observation in Rossell's argument is not, however, entirely clear. Even if only Spanish speakers were benefiting from native language reading instruction within bilingual programs, this does not constitute any indictment of the policy from an educational perspective.

In fact, though, sizable numbers of children who speak languages other than Spanish are receiving traditional bilingual education in which literacy is first introduced in the native language. Rossell and I spoke in the fall of 1998 at a meeting at Simmons College where we heard half a dozen bilingual teachers stand up and say that they were teaching children initial reading in languages ranging from Haitian Creole to Chinese.

The second aspect of this claim is that teaching reading in the native language is the key defining element of bilingual education. While not contesting the importance of native language literacy instruction to effective bilingual programs, I would disagree that this has been the primary justification within the intellectual history of the bilingual education movement. Rossell identifies a single theory as justifying bilingual education—Jim Cummins's theory of facilitation, or a common underlying proficiency. This theory has been widely cited in documents explaining the value of bilingual education, in particular in California where it has been combined with another theory, S. Krashen's notion of "comprehensible input" in informing educational practice.[37]

But facilitation is only one take on the value of bilingual education and, as Rossell's exposition reveals, the theory postdates the movement. The major intellectual input to bilingual education was practical wisdom—observations within the Coral Way program that bilingual education was possible—combined with a certain level of desperation concerning the academic accomplishments of non-English-speaking children. Justifications for bilingual practice have ranged widely, for example:[38]

—Basic humanity. It is neither kind nor nurturant to little children to put them in strange situations where they cannot communicate.

—Cultural continuity. Building on the skills and capacities children bring from home presupposes creating greater connections between home and school in language and in other ways.

—Achievement motivation. Children who are academically successful and well adjusted during their first year of formal schooling continue to show better school outcomes.

—Bilingualism. The children's home language skills constitute resources that should be exploited.

—Sociolinguistic realities. In many immigrant communities, both the ancestral language and English are used widely, and thus if schools are part of those communities, both languages will be used in school, though perhaps for different purposes.

I would argue, perhaps agreeing with Rossell, that the most urgent justification for bilingual programs is the value of native language literacy as an academic protective factor, but it is a misrepresentation to think that this is the only rationale for bilingual programs that has ever been proposed.

The procedures used by school districts to identify students as limited-English proficient are inherently illogical. Rossell has identified a number of cases in which school districts have shifted their judgment of who was limited-English proficient (LEP) and some practices of identification and of decisionmaking that are less than exemplary. The "technology" of language proficiency assessment is relatively new, and her worst cases (from twenty years ago) reflect that. It is wrong, though, to suggest that all of the variation in identification rates she notes reflect incoherent identification procedures. The rise in the LEP population between 1987 and 1995 noted in her figure 1 is what would be predicted from immigration rates during those years.

Rossell is right that determinations must be made of whom bilingual programs are meant to serve and how best to identify those children. She proposes, in effect, a secondary prevention strategy—cast a wide net that requires little in the way of individual testing or screening, and provide prevention services to all within those risk groups, because enhanced education cannot hurt those who do not need it. Much can be said for this model as a mechanism for distributing money. Less can be said for it if it becomes a mechanism for ignoring the specific educational needs of language minority children. Furthermore, within the language minority population, different children may need access to different prevention strategies. If teaching children to read occurs optimally in a language they speak well, then before assigning language minority children to English-medium classrooms where formal reading instruction will be offered, whether they speak English well enough must be ascertained. How can this be done? By judicious use of some of the tests that Rossell discusses. She complains about those tests on various grounds. How-

ever, that some native English speakers perform poorly on the same tests does not indict their usefulness as predictors of reading outcomes. Those low-scoring native English speakers no doubt would also profit from special educational treatments—but different ones than the non-native speaker needs.

The basic logic here is simple if one avoids being misled by Rossell's confused psychometric presentation. A six-year-old child who knows only three hundred words in English is going to have trouble in an English-medium first grade where the teacher is focusing on teaching reading. The child with a three-hundred-word vocabulary in English does not have a stable set of phonological distinctions, knows too few words to develop the phonological analysis skills basic to learning letter-sound correspondences, and will encounter many words critical to understanding first-grade texts that he or she does not know the meaning of. If such a child is a native English speaker, I would recommend English language enrichment combined with emergent literacy activities to promote prereading skills. If this child also knows six thousand words in Spanish, then the child should be given formal reading instruction in Spanish so that he or she can master the alphabetic principle and automaticity in letter and syllable recognition and develop the habit of reading, while acquiring more oral skills in English. If this child has become an on-grade-level reader in Spanish by the end of first grade, introducing English reading material at that point is likely to promote oral English development as well.

Nationally it is only possible to have half the student population at grade level. This claim, which Rossell not only makes but also excoriates others for failing to understand, is patently wrong. Rossell fails to display an understanding of the not-very-subtle distinction between norm-referenced and criterion-referenced measures. On norm-referenced measures, some test takers will always score below average. On criterion-referenced tests, however, everyone in a class, a school, or a nation can "pass." Grade level for reading is a criterion—established as a function of curricula and consensual standards. Rossell likely would counter that the curricula and the consensus standards shift to reflect average performance. Standards do shift—in the current period of educational reform, standards are being shifted upward, with the result that more than half the children in the nation are failing to meet many of them. If grade-level standards were averages, as Rossell seems to think they are, such an outcome would also be impossible.

Educators apparently believe that children who score below average . . . are children who are in academic difficulty. Rossell accuses educators of being as confused as the general public as to what tests mean and of appearing to believe that a low score has some absolute meaning. She is right that low scores on norm-referenced tests have no absolute meaning. But they do have power nonetheless, as predictors of academic outcomes and of literacy achievement. One could construct a language test on which first-grade children with excellent language and literacy skills score as low as the 10th percentile (the Scholastic Assessment Test [SAT] for college-bound high school students, for example). But if language minority or English-speaking students are scoring at the 10th percentile on valued age appropriate tests of vocabulary, knowledge of grammar, and oral comprehension in English, they are going to have trouble learning to read and will need adapted programs. All of Rossell's smoke and mirrors about the relativity of norm-referenced tests do not change those basic facts. The shift she notes in New York state's criterion for classification as LEP from the 23rd to 40th percentile may have reflected the reaction of wise practitioners to this finding. Whether the 40th percentile on the Language Assessment Battery (LAB) represents a level of English that is high or low depends on the norming sample. A cutoff at the 60th percentile would be justifiable, if children scoring at the 59th percentile were "having difficulty in English." No absolute meaning is attributable to either high or low scores, but that does not imply that test scores are meaningless. Rossell's university no doubt imposes as an entrance requirement for non-native English speakers a score on the Test of English as a Foreign Language (TOEFL) that is well above the 70th percentile for that test—and then occasionally still admits students whose English proves not to be adequate for the academic demands of a U.S. university.

School districts and states can increase or decrease the number of LEP students simply by changing the standard. Moreover, they do this on a regular basis without any rationale other than the desire to help students. School districts can change incidence of students classified as LEP or any other category by changing cutoff points. Rossell acknowledges that they do this through a desire to help students. While being below average on a universally administered test is the fate of half its takers, scoring below average on a test normed on non-native English speakers is not a likely fate for native speakers of English. Being below grade level could, in principle, afflict all or none of any given population. Again, Rossell con-

fuses norms with criteria, and norms for the language minority population with norms for the entire population.

[Facilitation theory] ignores the issue of the great variation in written language. In particular, it is silent on how to teach Asian children to read and write in their native tongue and why doing so is desirable. Essentially, Rossell is claiming that learning to read Asian languages could hardly develop skills that transfer to English. Rossell is general in her reference to Asian languages, which fall into a number of different language groups and use a variety of orthographies. Vietnamese and Korean, for example, use alphabetic writing systems—in the first case the Roman alphabet; Khmer's writing system is described as "alpha syllabic"; Japanese uses a syllabic system that displays many of the principles of alphabet writing but at a slightly less abstract level; the languages of the Indian subcontinent mostly use consonant-centered alphabetic systems.[39] Thus, Rossell seems to have generalized Chinese morpho-syllabic orthography (which is borrowed for use in sophisticated Japanese and Korean writing, but not for initial literacy instruction) to the entire continent. Nor do the data suggest that it is harder to learn to read in syllabic or morpho-syllabic orthographies. Initial reading in such systems is much easier, though acquiring full literacy (two thousand to twenty thousand Kanji) in morpho-syllabic systems does take some years.

The key issue here, though, is whether transfer of literacy skills from a first to a second language occurs only when the two languages share a writing system. From my own infinitely greater difficulty trying to learn to read Arabic and Hebrew as compared with Spanish and Dutch, I can sympathize with Rossell's intuition-based claim, that transfer to new alphabetic systems is less extensive than transfer within alphabetic systems. Unfortunately a systematic research base does not exist from which to argue either side of this case. However, there clearly are things one learns from knowing how to read in Chinese that could be useful in learning to read English—the analyzability of print, the need to access a phonological representation of words being read, something about the difference between spoken and written language, the value of literacy skills, and so on. And knowing any alphabetic system prepares one for the phonological analysis of words that constitutes the basis for any other alphabetic system—and often is a great barrier to young readers of English because English has a deeper orthography than most alphabetic languages. I must agree with Rossell that it may not be worth the time and

effort to teach children to read first in a language in which they will not have ongoing opportunities to read. Transfer of skills from first to second language literacy may require more than initial literacy accomplishments in the first language. In this case, given the absence of realistic opportunities to become an advanced or fluent reader in some languages because of absence of reading materials or ongoing support and instruction, some children perhaps should postpone formal literacy instruction until after the development of sufficient English proficiency.

Although Title VII favors bilingual education, not an ounce of evidence in 1968, and no consistent evidence in the thirty years since then, proves its superiority. The claim that research fails to confirm the value of bilingual education, central to Rossell's paper and to her work in this area, is demonstrably false. She and I agree that much of the research on this topic is execrable. She has included many of those studies in the overview presented in table 1. A. C. Willig and J. P. Greene reviewed far fewer studies in their meta-analyses because they identified crucial design flaws in many of the seventy-two studies Rossell allows to "vote" in her method.[40] Such flaws ranged from including graduates of transitional bilingual programs in the putative control group to failing to distinguish between Canadian (elite, elective) and U.S. immersion programs. The flaws in Rossell's approach to synthesizing research on bilingual education have been repeatedly discussed in numerous venues. Despite her willingness to stretch her methods so as to disfavor bilingual education as a model, her results do not demonstrate that bilingual programs are bad for children. The scores come out about even for the various programs across the various assessments offered. A fairer presentation of data and a stricter selection limited to well-designed studies make clear that bilingual education is much better for children entering school without literacy skills in English.

Four Recommendations

First, eliminate reference to limited proficiency in English and substitute "language minority" as the criterion for distributing Title VII funding. This proposal would increase the population base for distributing Title VII funding and would streamline the process of estimating the numbers of children eligible. Though it would not solve all problems of identification, it would simplify identification and could reduce error. If

elimination of consideration of children's language skills did not extend to the decisions made about their educational program, but was limited to determining a basis for distribution of funding, it might be sensible and deserves a trial.

Second, eliminate reference to specific educational programs in Title VII. This recommendation is based on a claim with which I categorically disagree, that research cannot be used to justify bilingual approaches to educating language minority children with limited skills in English. Thus, while I endorse the underlying notion that states and districts should be required to use solid research evidence in justifying their decisions about education, I cannot endorse this recommendation. Bilingual programs offer the best education to many children. Their quality needs to be improved, as Rossell points out. But the United States is not responding to the evidence that math and science programs are often ineffective by eliminating them. Instead, efforts are made to improve them. The same approach should be adopted for bilingual education.

Third, rewrite Title VII to eliminate redundancies and irrelevancies. Rossell and I agree entirely that Title VII is badly written. The impact of this on schools and children is not, however, so negative that major efforts to call in the language police need to be launched.

Fourth, Title VII legislation should require school districts to collect and make available outcome data, not to carry out evaluation studies. Like Rossell, I am appalled at the money and time wasted in carrying out evaluations of educational programs, generally with inadequate designs and insufficient resources to do a good job. However, requiring school districts to provide outcome data to outside evaluators will hardly solve the problem of generating good evaluations. A major issue in evaluating educational programs (those for language minority children and all others) is knowing what is going on inside the classroom. Program labels are not enough. Thus, I agree that evaluation efforts need to be rethought but do not agree that the solution offered by Rossell is adequate.

Comment by Charles Glenn

Christine H. Rossell has amply demonstrated what bilingual education guru Jim Cummins once called an "entry and exit fallacy in bilingual education." That is, the methods used to identify which children are

to be placed in separate bilingual education programs do not identify reliably those whose educational needs derive primarily from the dominance of a language other than English. Instead, they identify children whose proficiency in oral and written English is below national norms, whatever the reason. The same arbitrary criterion makes it difficult to return children, once designated as limited-English proficient, to the educational mainstream.

The key to a more sensible policy was suggested by Catherine E. Snow, who pointed out that a continuum exists of need and of language development that includes, at some point, every student in school. Because this is self-evidently true, it would make sense to abandon the labeling and educational segregation of some language minority students on the basis of an arbitrary cutoff point. My study of a dozen nations with large numbers of immigrant children found that only in some American states and in some highly controversial programs in Sweden is it considered appropriate to educate these children separately from the majority after an initial transition period of (in most cases) a year of intensive instruction in the language of the host society.[41]

Educational segregation is more harmful to language minority students than to any of the groups for which serious efforts have been made to integrate—female students, black students, special needs students. A case can be made, I believe, for single-sex schools, for schools with a special focus on the needs of African American youth, for schools concentrating on a particular disability. No convincing case can be made, I submit, for herding together language minority children whose most urgent educational task is to become effectively integrated into U.S. society. Language minority students have a compelling need to be with peers for whom English is the first language if they are to learn the language well.

They also have a compelling need to be held to the same educational standards as other students. Too often they are subjected to what I call "Jim Crow educational standards," which almost guarantee that they will not be able to participate in secondary and higher education on equal terms. Blame for these separate but unequal expectations must be shared by educational progressives and conservatives alike. The progressives have recoiled from holding language minority children to expectations that seem culturally insensitive and threatening to their self-esteem, which has led to bilingual education becoming a sort of comforting cultural bubble-bath for too many students who deserve to be challenged

instead. Conservatives, meanwhile, have sometimes focused so single-mindedly on the acquisition of English that other academic objectives are neglected. Even as a technique for teaching English, this is unwise. Proficiency in a language, beyond an elemental level, is developed by using it for real tasks that matter, such as mastering academic materials, not for artificial exercises.

The fundamental mistake made by both sides in the debate over educating language minority children is to focus on language instead of education as the central issue. As both Rossell and Snow seem to agree, effective education can be provided either through use of the home language or through structured immersion. Either can be done well or badly. A couple of years ago, I served as a reviewer for the National Research Council's (NRC) study of thirty-five years of evidence on the teaching of language minority children.[42] The NRC report concludes that, despite countless research studies and evaluations (costing hundreds of millions of dollars), one approach is not superior to the other. "It is clear," the authors note,

> that many children first learn to read in a second language without serious negative consequences. These include children in early-immersion, two-way, and English as a second language (ESL)-based programs in North America, as well as those in formerly colonial countries that have maintained the official language [of the colonizer] as the medium of instruction, immigrant children in Israel, children whose parents opt for elite international schools, and many others. . . . The high literacy achievement of Spanish-speaking children in English-medium Success for All schools . . . that feature carefully-designed direct literacy instruction suggests that even children from low-literacy homes can learn to read in a second language if the risk associated with poor instruction is eliminated.[43]

Later in the report, the authors candidly conclude, "We do not yet know whether there will be long-term advantages or disadvantages to initial literacy instruction in the primary language versus English, given a very high-quality program of known effectiveness in both cases.[44]

The emphasis should now shift to ensuring that whichever method is chosen in particular circumstances be implemented by competent teachers following a demanding curriculum and with accountability for clear and measurable results. What will that take? Some concrete measures should be reflected in federal and state educational policy for language minority students.

First, the principal and teachers in each school should be responsible for planning and implementing the education of all the students in that school and should have broad discretion about the instructional methods that they use. Development of both oral and written language is a continuous process alongside the other tasks of schooling, and only those directly involved with students should be diagnosing what each needs at a particular time and prescribing the challenges and the support that will best meet those needs. Only those working in the school can develop an effective combination of integration for common tasks and separation for special help.

To make this possible, state and federal programs supporting the education of language minority students should not prescribe teaching methods or the language used but should hold schools accountable for the measurable, steady progress of these students in all required academic subjects.

Second, teachers and school administrators should receive specific training in strategies for language development, including how to diagnose and prescribe for the needs of language minority students. States should make this an important requirement of teacher and administrator certification, and coherent pre-service and in-service training in these skills should be a priority for federal funding. Additional research is not needed to determine what the necessary skills are. Much is already known about good practices in promoting language development; what is not known about, and perhaps will never be known about in view of the complexity and variation of all the factors involved, is what a complete model of good schooling for language minority children would be. Those practices should be taught to every teacher and administrator, not just to those who are preparing to work in separate bilingual programs.

The fine print of the National Research Council report concedes that "we need to move away from thinking about programs in such broad terms and instead see them as containing multiple components—features that are available to meet the differing needs of particular students."[45]

Perhaps some day a general model will emerge for the education of language minority children, though I am skeptical about that. Those who work in pedagogy as an academic discipline have long sought to make the field an exact science comparable to the natural or, more modestly, to the social sciences, with strong and reliable predictive power. If such a general theory of learning "linguistic, social, and cognitive skills" (as

the NRC study wistfully puts it) is ever developed, it should take language minority children into account. May it happen, and soon.

But until that glad day comes, the interests of language minority children will be better served by principled and theory-based experimentation on effective schooling of poor children of whatever ethnic background, taking language into account in how they are assessed and taught, than by putting faith in research on second-language acquisition. A fair amount is already known about how to develop bilingualism among middle-class children, especially if their parents are bilingual, but very little is known about how to overcome the academic underachievement of Latino youth or why they are outperformed by youth from other immigrant groups. Do they have more in common with underachieving African American students?

I would put my money on schools that are effective by other measures to be effective also for language minority students, but only if they are set free to tackle problems for every student without programmatic preconditions.[46]

Third, such a strategy of school-level freedom and accountability requires that language minority students be included in all assessments of academic progress. In some limited instances this will appropriately be done through assessment in their home language, but the great difficulty of making assessments in different languages comparable, and the implicit message that students are not expected to demonstrate proficiency in English, creates a danger of returning to Jim Crow standards. In general, it is preferable to assess language minority students in English, while making allowances in reporting and using the results for the challenges they face.

Because schools will choose different strategies for language minority students, parents should be allowed to choose among schools. For some, the maintenance and development of the home language in school will be much more important than it will be for others. Surveys have found, again and again, that Latino parents tend to want the school to help maintain their children's Spanish (though not at the expense of time devoted to English), while Asian and other language minority parents prefer to do that at home or through after-school community groups. Parents should be able to opt for a school that supports their own educational goals.

Finally, schools would be enriched if they provided elective and supplemental—not transitional—language support in a variety of world

languages to students whose parents speak those languages as well as to students whose parents do not. In place of the touchy-feely multicultural activities in so many schools, it would be much healthier for students of different ethnic backgrounds to tackle together the difficulties and the rewards of a language, and thus to learn from one another.

Notes

1. There are variations on these models, and they form a continuum.

2. The Office of Bilingual Education and Minority Language Affairs (OBEMLA) has undergone several name changes since it was first established as the Office of Bilingual Education in 1974. Rachel F. Moran, "Of Democracy, Devaluation, and Bilingual Education," *Creighton Law Review*, vol. 26 (1993), pp. 255–319.

3. See Christine H. Rossell and J. Michael Ross, "The Social Science Evidence on Bilingual Education," *Journal of Law and Education*, vol. 15 (Fall 1986), pp. 385–419.

4. See Arnold H. Leibowitz, *The Bilingual Education Act: A Legislative Analysis* (Rosslyn, Va.: National Clearinghouse for Bilingual Education, 1980).

5. The complete text of the Clinton administration's proposal to reauthorize the Elementary and Secondary Education Act, called the Educational Excellence for All Children Act of 1999, can be found at http://www.ed.gov/offices/OESE/ESEA/legislation and a summary at http://www.ed.gov/offices/OESE/ESEA/prospectus/overview.html.

6. Council of Chief State School Officers (CCSSO), *State Education Indicators with a Focus on Title I* (Washington, 1998). These statistics are presented state by state for 1989–90 and 1996–97 and must be added to get the total for the nation. The growth rate shown in the CCSSO report after adding up each state is higher than that cited in the Clinton proposal. It is 72 percent.

7. Department of Education, National Center for Education Statistics, *Digest of Education Statistics,* NCES 93–292 (Washington, 1993); and Department of Education, National Center for Education Statistics, *The Condition of Education,* NCES 93–290 (Washington, 1993).

8. Contrary to popular belief, however, typically little fiscal incentive exists to enroll a limited-English proficient (LEP) child in bilingual education. The LEP identification is the source of extra state revenues, not the program that the child is placed in. With regard to federal funds, although three quarters of the federal Title VII funds are earmarked for bilingual education programs, the money received is only a small portion of the cost of these programs. Thus, it is hard to imagine that a rational actor would implement a bilingual education program solely for the money.

9. The concept of grade level and reading below grade level is almost universally misunderstood, not only by laymen, but also by educators. Grade level is simply the average achievement for a particular grade; it has no absolute meaning. It is not possible, for example, for all students in the norming population to be at grade level because it is not possible for all students to be at or above average, only half can be.

10. J. David Ramirez, Sandra D. Yuen, and Dena Ramey, *Second Year Report: Study of Immersion Programs for Language Minority Children* (Alexandria, Va.: Sra Technologies, 1986). See also Keith Baker and Christine Rossell, "An Implementation Problem: Specifying the Target Group for Bilingual Education," *Educational Policy,* vol. 1, no. 2

(1987), pp. 249–70; and Christine Rossell and Keith Baker, "Selecting and Exiting Students in Bilingual Education Programs," *Journal of Law and Education*, vol. 17, no. 4 (1988), pp. 589–623.

11. Robert Berdan, Alvin So, and Angel Sanchez, *Language among the Cherokee: Patterns of Language Use in Northeastern Oklahoma*, part 1: *The Preliminary Report* (Los Alamitos, Calif.: National Center for Bilingual Research, 1982); Bureau of the Census, *1984 Analysis Data for the Office of Planning, Budget, and Evaluation: Decision Resources* (Department of Education, 1984); and C. Perlman and W. Rice Jr., "A Normative Study of a Test of English Language Proficiency," paper presented at the annual meeting of the American Educational Research Association, San Francisco, Calif., 1979.

12. Sharon E. Duncan and Edward A. De Avila, "Relative Language Proficiency and Field Dependence/Independence," paper presented at the annual meeting of Teachers of English for Speakers of Other Languages (TESOL), Boston, 1979.

13. D. Ulibarri, M. Spencer, and G. Rivas, *Comparability of Three Oral Language Proficiency Instruments and Their Relationship to Achievement Variables* (California State Department of Education, 1980); G. Gillmore and A. Dickerson, *The Relationship between Instruments Used for Identifying Children of Limited English Speaking Ability in Texas* (Houston, Texas: Education Service Center, Region IV, 1979); Robert A. Cervantes, *Entry into and Exit from Bilingual Education Programs* (Washington: E. H. White Inc., 1982); and Sol Pelavin and Keith Baker, "Improved Methods of Identifying Who Needs Bilingual Education," paper presented at the annual meeting of the American Research Association, Washington, D.C., 1987.

14. Perlman and Rice, "A Normative Study of a Test of English Language Proficiency."

15. Nancy L. Russell and Alba A. Ortiz, "Assessment of Pragmatic Skills of Kindergarten Limited English Proficient Children in a Dialogue Model of Communication," paper presented at the annual meeting of the American Educational Research Association, San Francisco, Calif., 1989; and Southwest Regional Laboratories for Educational Research and Development (SWRL), *Development of Entry/Exit Criteria and Associated Assessment Procedures for Bilingual Education Projects* (Los Alamitos, Calif., 1981).

16. *Aspira of New York, Inc., et al*. v. *Board of Education of the City of New York, et al.,* F. Supp. 1161 (1975).

17. The sources are as follows: Fiscal 1978 (2.4 million): Michael J. O'Malley, *Children's English and Services Study: Language Minority Children with Limited English Proficiency in the United States* (Rosslyn, Va.: InterAmerica Associates, 1981), cited in Robert E. Barnes and Anne M. Milne, "The Size of the Eligible Language-Minority Population," in Keith A. Baker and Adrianna A. de Kanter, eds., *Bilingual Education* (Lexington, Mass.: D.C. Heath and Company, 1983), p. 13; (3.6 million): O'Malley cited in Barnes and Milne, "The Size of the Eligible Language-Minority Population," p. 11; and (1.2 million): Barnes and Milne, "The Size of the Eligible Language-Minority Population," p. 14, table 1-2. Fiscal 1986 (1,472,042), fiscal 1987 (1,553,918), fiscal 1988 (1,656,180), fiscal 1989 (1,946,107), fiscal 1990 (2,154,781), fiscal 1991 (2,232,500), fiscal 1992 (2,430,712), and fiscal 1993 (2,735,952): Allison Henderson, Brenda Donly, and William Strang, *Summary of the Bilingual Education State Education Agency Program Survey of the States' Limited English Proficient Persons and Available Educational Services 1992–1993* (Westat Inc. and Development Associates, September 1994), p. 9, figure 2-1. Fiscal 1990 (3,307,500): D. August and K. Hakuta, *Federal Education Programs for Limited English Proficient Students: A Blueprint for the Second Generation* (Stanford Working Group, 1993), cited in Kris Anstrom, "Defining the Limited-English Proficient Student Population," *Directions in Language and Education, National Clearinghouse of Bilingual Education*, vol. 1, no. 9

(Summer 1996), http://www.ncbe.gwu.edu/ncbepubs/directions/09.htm; and (2,388,243): figure from Department of Commerce, *1990 Census of Population and Housing Summary Tape File 3C* (Bureau of the Census, Data User Services Division, 1990) cited in Paul J. Hopstock and Bonnie J. Bucaro, *A Review and Analysis of Estimates of the LEP Student Population*, submitted by Special Issues Analysis Center (Arlington, Va.: Development Associates Inc., 1993). Fiscal 1991 (2,314,079): Howard Fleischman and Paul Hopstock, *Descriptive Study of Services to Limited English Proficient Students*, vol. 1: *Summary of Findings and Conclusions* (Arlington, Va.: Development Associates Inc., 1993), p. 3; (1,997,742): Paul Hopstock and others, *Descriptive Study of Services to Limited English Proficient Students*, vol. 2: *Summary of Findings and Conclusions* (Arlington, Va.: Development Associates Inc., 1993) cited in Anstrom, "Defining the Limited-English Proficient Student Population." Fiscal 1991 (2,198,778), fiscal 1992 (2,429,815), fiscal 1993 (2,620,747), and fiscal 1994 (3,037,922): Brenda Donly, Allison Henderson, and William Strang, *Summary of the Bilingual Education State Education Agency Program Survey of the States' Limited English Proficient Persons and Available Educational Services 1993–1994* (Westat Inc. and Development Associates, September 1995), p. 11, table 2-1. Fiscal 1992 (2,160,000): H. Puma and C. Jones, *Prospects: The Congressionally Mandated Study of Educational Growth and Opportunity: Interim Report* (Abt Associates Inc., 1993), cited in Anstrom, "Defining the Limited-English Proficient Student Population"; and (1,892,845): ALEC Foundation and U.S. English, *Bilingual Education in the United States, 1991–92* (Washington, 1994), p. 7. Fiscal 1997 (3,358,163): Council of Chief State School Officers, *State Education Indicators with a Focus on Title I*.

18. Anstrom, "Defining the Limited-English Proficient Student Population."

19. Oona M. Cheung, Barbara S. Clements, and Y. Carol Miu, *The Feasibility of Collecting Comparable National Statistics about Students with Limited English Proficiency* (Washington: Council of Chief State School Officers, September 1994).

20. 1984 Elementary and Secondary Education Act (ESEA), Section 703 (6).

21. Catherine E. Snow, M. Susan Burns, and Peg Griffin, eds., *Preventing Reading Difficulties in Young Children* (Washington: National Academy Press, 1998).

22. 1984 ESEA, Title VII, Section 702 (7).

23. Citation information on the studies in table 1 can be found in Christine H. Rossell and Keith Baker, "The Educational Effectiveness of Bilingual Education," *Research in the Teaching of English*, vol. 30, no. 1 (February 1996), pp. 7–74, especially appendix A and appendix C. For a summary of these results and other issues, see Christine H. Rossell and Keith Baker, *Bilingual Education in Massachusetts: The Emperor Has No Clothes* (Boston, Mass.: Pioneer Institute, 1996).

24. A more detailed discussion of these results can be found in Rossell and Baker, "The Educational Effectiveness of Bilingual Education."

25. A cohort is a group of students who are followed across grades in their progression through school. Thus, a group of students who started kindergarten in 1960 and graduated from high school in 1974 would be one cohort. A second cohort might be a group of students who started kindergarten in 1961 and graduated from high school in 1975.

26. These studies are listed in alphabetical order in complete citation format in Rossell and Baker, "The Educational Effectiveness of Bilingual Education," appendix B and references.

27. The "voting method" is used to evaluate the literature's findings. In the voting method, the percentage of studies showing each program's outcome is calculated. A theoretically preferred method is meta-analysis—a statistical analysis of the effects of bilingual education across all studies. But because of the lack of information in these studies by

which to estimate a common outcome measure, the two meta-analyses that have been conducted have had only a small number of studies. Ann C. Willig's 1985 meta-analysis included only thirteen of the thirty-nine acceptable studies of transitional bilingual education from Keith Baker and Adrian de Kanter's 1981 study. Jay P. Greene's 1997 meta-analysis of Rossell and Baker's 1995 study analyzed only eleven of the seventy-two studies. Both Willig and Greene concluded that bilingual education was superior. See Ann C. Willig, "A Meta-Analysis of Selected Studies on the Effectiveness of Bilingual Education," *Review of Educational Research*, vol. 55, no. 3 (1985), pp. 269–317; Keith Baker and Adrian de Kanter, *The Effectiveness of Bilingual Education Programs: A Review of the Literature, Final Draft Report* (Department of Education, 1981); and Jay P. Greene, "A Meta-Analysis of the Rossell and Baker Review of Bilingual Education Research," *Bilingual Research Journal*, vol. 21, nos. 3 and 4 (1997), pp. 103–21.

28. Oral progress in preschool or kindergarten is included in this category because a reading test for these grades is inappropriate.

29. Marcello Medina and Kathy Escamilla, "Evaluation of Transitional and Maintenance Bilingual Programs," *Urban Education*, vol. 27, no. 3 (1992), pp. 263–90. J. David Ramirez and others also examined maintenance bilingual education (late-exit bilingual education) but unfortunately did not directly compare it with transitional bilingual education (contrary to media reports and their own conclusions). Although their graphs appeared to show that the students in late-exit bilingual education were way behind the students in transitional bilingual education although beginning to catch up, no statistical analysis was performed to determine if they would ever catch up. See J. David Ramirez and others, *Final Report: Longitudinal Study of Structured Immersion Strategy, Early-Exit and Late-Exit Transitional Bilingual Education Programs for Language-Minority Children*, report to the Department of Education (San Mateo, Calif.: Aguirre International, 1991).

30. Ramirez and others, *Final Report*; El Paso Independent School District, *Interim Report of the Five-Year Bilingual Education Pilot 1986–87 School Year* (El Paso, Texas: Office for Research and Evaluation, 1987); El Paso Independent School District, *Bilingual Education Evaluation: The Sixth Year in a Longitudinal Study* (El Paso, Texas: Office for Research and Evaluation, September 1990); and El Paso Independent School District, *Bilingual Education Evaluation* (El Paso, Texas: Office for Research and Evaluation, November 1992).

31. ESEA, Title VII, Section 7102 (b) (2).

32. Descriptions of these varied interpretations can be found in Rossell and Baker, *Bilingual Education in Massachusetts,* chapter 4*;* and Christine Rossell, "Teaching Language Minorities: Theory and Reality," in Diane Ravitch and Joseph Viterriti, eds., *City Schools: Lessons from New York* (Johns Hopkins University Press, forthcoming 1999).

33. See Rossell and Baker, *Bilingual Education in Massachusetts,* chapter 7.

34. Clinton administration proposal, p. VII-7.

35. See also Rossell and Baker, *Bilingual Education in Massachusetts*, for a detailed description of how a research study should be conducted and what variables it should include.

36. C. E. Snow, M. S. Burns, and P. Griffin, eds.*, Preventing Reading Difficulties in Young Children* (Washington: National Academy Press, 1998).

37. J. Cummins, "The Role of Primary Language Development in Promoting Educational Success for Language Minority Students," in *Schooling and Language Minority Students: A Theoretical Framework* (California State University at Los Angeles, National Evaluation, Dissemination and Assessment Center, 1981), pp. 3–49; and S. Krashen, *Language Acquisition and Language Education* (Prentice-Hall International, 1989).

38. C. E. Snow, "Rationales for Native Language Instruction in the Education of Language Minority Children: Evidence from Research," in A. Padilla, H. Fairchild, and C. Valadez, eds., *Bilingual Education: Issues and Strategies* (Newbury Park, Calif.: Sage, 1990).

39. E. Schiller, "Khmer Writing," in P. T. Daniels and W. Bright, eds., *The World's Writing Systems* (New York: Oxford University Press, 1996), pp. 467–73; and W. Bright, "The Devanagan Script," in P. T. Daniels and W. Bright, eds., *The World's Writing Systems* (New York: Oxford University Press, 1996), pp. 384–90.

40. A. C. Willig, "A Meta-Analysis of Selected Studies on the Effectiveness of Bilingual Education," *Review of Educational Research*, vol. 55, no. 3 (1985), pp. 269–317; and J. P. Greene, "A Meta-Analysis of the Effectiveness of Bilingual Education," University of Texas at Austin and Thomas Rivera Policy Institute, 1998.

41. Charles L. Glenn with Ester J. de Jong, *Educating Immigrant Children: Schools and Language Minorities in Twelve Nations* (New York: Garland, 1986).

42. Diane August and Kenji Hakuta, eds., *Improving Schooling for Language-Minority Children: A Research Agenda* (Washington: National Research Council, 1997).

43. August and Hakuta, *Improving Schooling for Language-Minority Children*, p. 60.

44. August and Hakuta, *Improving Schooling for Language-Minority Children*, p. 179.

45. August and Hakuta, *Improving Schooling for Language-Minority Children*, p. 158.

46. Thomas P. Carter and Michael L. Chatfield, "Effective Schools for Language Minority Students," *American Journal of Education*, vol. 97 (1986), pp. 200–33.