

Selecting Students For Bilingual Education Under The *Keyes* Agreement

Keith Baker*

The *Keyes* agreement¹ details the procedures for identifying the students eligible for enrollment in the Denver bilingual education program. First, the agreement requires that a home language survey be conducted which asks whether: (1) the student's first language is other than English; and (2) someone speaks a non-English language in the student's home. Secondly, if the answer to either of these questions is "yes," the student is tested for English oral proficiency and, if in grade two or above, the student undergoes testing in academic achievement in language and reading as measured by a standardized achievement test; and all students are tested for English oral proficiency. A student is assigned to the bilingual program either if the oral proficiency test classifies him as limited English proficient or if the student scores below the cutoff score on the CTBS (Comprehensive Test of Basic Skills)—the thirtieth percentile for elementary school or the fortieth percentile for middle and high school.

In this article I will examine whether this procedure can successfully identify students who would benefit from instruction in a non-English language as called for by the *Keyes* agreement. I begin with the problems encountered in trying to determine which students are eligible for bilingual education. Secondly, I discuss the multiple causes of poor performance in school—causes which were not taken into consideration as part of the *Keyes* selection methods. Lastly, I look at the problems associated with the specified methods for transferring students from the bilingual program into the regular English-speaking classroom.

* Social-science analyst, Office of the Deputy Under Secretary for Planning, Budget, and Evaluation, United States Department of Education. The views expressed are those of the author and do not necessarily represent either the policies or opinions of the United States Department of Education.

1. *Keyes v. School Dist. No. 1*, No. C-1499 (D. Colo. Aug. 17, 1984) (order approving consent decree) [hereinafter cited as the Consent Decree]; see Baker & de Kanter, *Assessing the Legal Profession's Contribution to the Education of Bilingual Students*, for a discussion of the *Keyes* agreement.

I.

THE INADEQUACY OF THE SELECTION PROCESS

The first step in the *Keyes* procedures, the home language survey, predisposes the process to failure. Knowing that a child's first spoken language was not English does not guarantee that the child is not now a competent English speaker or that the child speaks a non-English language well enough for it to be an effective language of instruction. The home language survey, moreover, asks if *anyone*, not just the target child, in the home speaks a language other than English. Since there are large intergenerational differences in language use in immigrant groups,² knowing that a child's parents or grandparents speak Spanish, for example, provides no information as to what language the child speaks. A monolingual English-speaking child could answer "yes" to this question. Thus, the first step of the *Keyes* procedures could identify a monolingual English-speaking child as in need of instruction in a non-English language he neither speaks nor understands for the purposes of improving his English.

The second step of the *Keyes* selection process utilizes standardized and language proficiency tests to determine if students whose "home language" is other than English should be placed in bilingual education programs. Problems with both types of these tests have been well documented and will be summarized below.

Standardized achievement tests can easily be misused in determining student eligibility for bilingual education for three reasons. First, the proportion of monolingual English-speaking students equivalent to the percentile score used as a cut-off would be found in need of bilingual instruction. That is, thirty percent of all monolingual English-speaking students in grade school and forty percent in high school will fall below the *Keyes* criteria; yet, monolingual English-speaking students do not need bilingual instruction. Second, the cut-off scores are arbitrary. Lastly, standardized tests do not provide enough information to be able to determine whether a student could benefit from bilingual education or not.

The student takes a standardized achievement test, the CTBS, and is put in the bilingual program if the student scores below the thirtieth percentile (or fortieth percentile in the higher grades). The CTBS, however, is designed so that roughly one out of every three monolingual English-speaking students in the entire elementary school population and forty percent of middle and high school students will fail these criteria. There-

2. Hernandez-Chavez, *Language Maintenance, Bilingual Education, and Philosophies of Bilingualism in the United States*, in INTERNATIONAL DIMENSIONS OF BILINGUAL EDUCATION (J. Alatis ed. 1978).

fore, to the extent the home language survey delivers English-speaking children for testing, the standardized test may erroneously classify a large proportion of them as needing bilingual instruction. The chances that an English-speaking child from a non-English- or partial-English-speaking home background will score below the cut-off score are higher than that for the English-speaking general population, because students from non-English language homes generally are socioeconomically disadvantaged and disadvantaged students do poorly on standardized tests.³

A second problem with standardized tests is that the cut-off score at the thirtieth percentile is arbitrary. Why, for example, did the *Keyes* agreement use the thirtieth percentile? Why did the Texas Title VII projects audited by the Inspector General use the fortieth percentile?⁴ Why did not these programs use the fifteenth, thirty-third or fifty-second percentile? Percentile cut-off scores are arbitrary. Being arbitrary, they differ from place to place because there is no good reason for selecting one over another. Being arbitrary, they are meaningless.

Standardized tests are *not* designed to measure whether a student has a sufficiently good command of English to be able to succeed in school or in society. Standardized tests are designed to differentiate a monolingual population into one hundred different, ordered categories (designated by percentile). Those essential language skills⁵ that all people must possess to function in the classroom or in society are useless during standardized testing. Since everyone has these skills, they do not help differentiate the population, therefore, they are not included in the test. But it is these common skills which are most critical to the language minority students' mastery of English. It is these skills that are not tested by standardized tests.

The third problem which arises in the use of standardized tests is that the entry decision test score will be extremely unreliable since it will

3. A. Rosenthal, K. Baker & A. Ginsburg, *The Effects of Language Background on Achievement Level and Learning Among Elementary School Students*, 56 SOCIOLOGY OF EDUCATION 157-69.

4. DEPARTMENT OF EDUCATION, REVIEW OF FEDERAL BILINGUAL EDUCATION PROGRAMS IN TEXAS 19 (1982) [hereinafter cited as REVIEW] (available at the offices of the LA RAZA LAW JOURNAL, Room 37, Boalt Hall School of Law, University of California, Berkeley, CA 94720).

5. Explaining these essential language skills is beyond the scope of this paper. Briefly, these language skills can be defined as the most basic of the communication skills necessary to communicate in society. These skills are so basic that one with a minimum amount of English, for example, would possess them. For instance, when kindergarteners walk into the classroom for the first time, teachers assume that when they say "Please sit down in a chair," that the meaning of these six words are understood by the pupils. These words are easily understandable and, hence, one with a minimum amount of knowledge of the English language would follow the command. This level of understanding English would be tested in a language proficiency test but not in a standardized test. The latter is geared toward differentiating various levels of knowledge, not a command of language skills. Standardized tests assume the basic command of the language already exists.

be based almost entirely on guessing.⁶ Answers which are guesses provide no information about the reasons why the student guessed. In almost every case, the test score will be meaningless and so will any decision based on the test score.

A single score on a standardized test, moreover, is not sufficient to determine if a student will be unable to successfully learn in a regular English-speaking classroom. This problem is made obvious in the following thought experiment.⁷

Imagine a Vietnamese refugee child, ten years old, who enrolls in the Denver School System, is tested and found to score at the forty-fifth percentile. No special help is available for him. Imagine that somehow, his family managed to bring along his school records from Vietnam and they showed that he performed at the eighty-fifth percentile. Now we know this is clearly a student who belongs in a special language program because he has a lot of English to learn before he can realize his potential in the regular English-speaking classroom. But the *Keyes* procedures will exclude him from any help. Now imagine another student in a similar situation. This time, however, the student tests at the twentieth percentile in English. This student is put into the bilingual program. His home school test records, however, show he performed at the twentieth percentile when he was tested in his home language. This student seems to have learned English to the full level of his potential and does not belong in bilingual education.

As this thought experiment demonstrates, the wrong decision is made in both cases by the *Keyes* methods. Rather than correcting the misclassification errors made in the home language survey, the use of standardized tests adds misclassification errors of its own, compounding the mistakes.

Language proficiency tests, required of all students, exacerbate the problem. Following the passage of state bilingual legislation requiring language proficiency tests, both California and Texas set up blue-ribbon commissions to identify and recommend satisfactory language proficiency tests for use throughout their respective states.⁸ Both commissions concluded that there were *no* psychometrically acceptable language

6. If a student does not understand the language in which the test is written, the student will only be able to guess at the answers. Therefore, the student's score will not evidence the student's knowledge of the tested subject matter but will reflect chance guessing.

7. Readers who question whether the *Keyes* procedures can be critically examined by thought experiments are reminded that Albert Einstein discovered relativity by a thought experiment.

8. J.D. Ramirez, B. Merino, T. Bye & N. Gold, *Assessment of Oral English Proficiency: A Status Report* (1981) (hereinafter referred to as Ramirez) (available at the offices of the LA RAZA LAW JOURNAL, Room 37, Boalt Hall School of Law, University of California, Berkeley, CA 94720); and TEXAS EDUCATION AGENCY, *REPORT OF THE COMMITTEE FOR THE EVALUATION OF LANGUAGE ASSESSMENT INSTRUMENTS* (1979).

proficiency tests in existence.⁹

One of the most important problems with an oral language proficiency test is its lack of predictive validity concerning academic achievement.¹⁰ These tests cannot differentiate between students who will do well in school from those who will do poorly. But this is precisely the issue at the heart of any selection method intended to identify students needing special help. If a test cannot separate those students who will have trouble in school because they temporarily lack English language proficiency from those who will always do poorly, there is no point testing the students. Assigning them to programs based on the results of such tests is unsound educational practice and unfair to the student.

The inadequacy of the current methods used to identify limited English proficiency (LEP) has been well documented. The National Institute of Education developed a special test to measure LEP, the Language Measurement and Assessment Instrument (LM&AI). This test was administered to a stratified sample of students. Statistical estimating procedures were used which estimated there were 3.6 million LEP students in the nation. A re-analysis of this data found that two-thirds, or 2.4 million, of these students spoke English as their primary language.¹¹ In another study, the LM&AI was administered to Cherokee students at the request of the Cherokee Nation which wished to know the extent of the need for Cherokee bilingual education.¹² Through home interviews, the researchers found that eighty-two percent of the Cherokee students were English monolinguals. The LM&AI classified forty-eight percent of

9. Other reviews of language proficiency tests have confirmed their lack of reliability and validity. D. Horst, D. Johnson, H. Nava, D. Douglas, L. Friendly, and A. Roberts, *An Evaluation of Project Information Packages (PIPs) as Used for the Diffusion of Bilingual Projects* (Vol. III): A Prototype Guide to Measuring Achievement Level and Program Impact on Achievement in Bilingual Project (May, 1982) (available on Educational Resources Information Center, hereinafter cited as ERIC, Ed No. 193-955); Center for Bilingual Education, Northwest Regional Educational Laboratory, *Assessment Instruments in Bilingual Education: A Descriptive Catalogue of 342 Oral and Written Tests* (1978) (available on ERIC, *supra*, Ed No. 173-373); B. Pletcher, N. Locks, D. Reynolds and B. Sisson, *A Guide to Assessment Instruments for Limited English Speaking Students* (1978) (available on ERIC, *supra*, Ed No. 165-499); and Dissemination and Assessment Center for Bilingual Education, *Evaluation Instruments for Bilingual Education: An Annotated Bibliography* (1976) (available on ERIC, *supra*, Ed No. 111-182).

10. Canale, *On Some Dimensions of Language Proficiency*, in *ISSUES IN LANGUAGE TESTING RESEARCH* (J. Oller, Jr. ed. 1983); J. Cummins, *Language and Literacy Learning in Bilingual Instruction: Policy Report* (September, 1983) (available on ERIC, *supra* note 9, Ed No. 245-575); J. Cummins, *The Role of Primary Language Development in Promoting Educational Success for Language Minority Students*, in *OFFICE OF BILINGUAL BICULTURAL EDUCATION, CALIFORNIA STATE DEPARTMENT OF EDUCATION, SCHOOLING AND LANGUAGE MINORITY STUDENTS: A THEORETICAL FRAMEWORK* (1981).

11. Barnes, *The Size of the Eligible Language Minority Population*, in *BILINGUAL EDUCATION: A REAPPRAISAL OF FEDERAL POLICY* (1983).

12. R. Berdan, A. So & A. Sanchez, *Language Among the Cherokee: Patterns of Language Use in Northeastern Oklahoma* (Part 1, Preliminary Report) (1982) (available on ERIC, *supra* note 9, Ed No. 234-637).

these monolingual English-speaking children as LEP and, therefore, presumably in need of instruction in Cherokee so they could improve their English. Recently, the United States Department of Education had the LM&AI administered to a nationally representative sample of *monolingual* English-speaking school-aged children.¹³ The test classified forty-two percent of them as LEP. Despite extraordinary efforts to develop a test to measure LEP, the LM&AI has convincingly failed.

The states have fared no better than the federal government in their attempts to develop measures of limited English proficiency. The Office of the Inspector General of the Department of Education said:

Texas districts use LEP identification procedures which are inadequate and may result in an overstatement of LEP students needing bilingual education services . . . For example, 812 (53 percent) of the 1,524 LEP-classified students participating in Edgewood's three Title VII projects were categorized as LEP because they had been designated as under-achievers by the district (scoring below the 40th percentile on the Comprehensive Tests of Basic Skills total battery score). The home language surveys of many of these students showed that the parents spoke both English and Spanish, but did not indicate which language was "most relied upon for communication." We were advised by the principal and teachers at one participating school that the students' low socioeconomic status was the reason for these LEP-classified students' underachievement (difficulty in reading and writing English), rather than exposure to a language other than English.

While priority in funding is to be given to projects serving those LEP students with the greatest need for bilingual education services . . . English was the dominant language for most Title VII project participants, including at least 1,378 or 33 percent of the participating LEP students.

Most LEP-classified preschool students participating in Austin's demonstration project spoke only English, based on the results of oral language proficiency tests. Since these preschool children could not be tested for reading and writing proficiency, they were classified as LEP mostly because home language surveys showed that Spanish was sometimes used at home, but was not "the language normally used by the parents of the child," in accordance with program definitions. In most instances, the language surveys showed either that mostly English was used or that both Spanish and English were equally used by the parents. A project teacher who administered the language proficiency tests said most of the students needed only oral language development (as do many non-LEP preschool students) rather than bilingual education.¹⁴

13. This study is soon to be made publically available.

14. REVIEW, *supra* note 4, at 19-20. The Texas bilingual education program is strikingly similar to the *Keyes* agreement:

The bilingual education program regulations (34 C.F.R. 500.4) state that a student "whose native language is other than English" or "who comes from a home in which a language

A similar study found that large numbers of supposedly Spanish-speaking children selected for participation in California's state bilingual program spoke so little Spanish that the schools were having to teach them Spanish before they could be provided with the legally required Spanish instruction in nonlanguage subjects.¹⁵ These examples illustrate the flaws of the *Keyes* procedures. English-speaking students are included in programs designed to teach in a non-English language and to teach English to supposedly non-English speakers.

The preceding is a general discussion of the problems of language proficiency tests. The *Keyes* agreement specifically requires the use of the Bilingual Syntax Measure (BSM) and the Language Assessment Scale (LAS).¹⁶ The specific problems and shortcomings of these two tests have been detailed previously:

With few exceptions the LAS, BSM . . . do not identify the same percentage of children as (do other tests) nor are particular students reliably classified by these instruments . . . The LAS tends to rate the proficiency level of students lower than either bilingual teacher judgments or ethno-linguistic analyses . . . LAS English has a moderate relationship to linguists' analyses of free speech though it assesses consistently lower . . . Data elicited by the BSM-English appear to be unrelated to data elicited from taped naturalistic conversation of the same individuals.¹⁷

other than English is *most relied upon* for communication" is LEP *only if, as a result of these circumstances*, the child "has sufficient difficulty in understanding, speaking, reading, or writing the English language to deny him or her the opportunity to learn successfully in classrooms in which the language of instruction is English" (emphasis added).

Procedures used by Texas school districts to identify LEP students for participation in bilingual education programs were established in accordance with the definition of limited English proficiency contained in the Bilingual Education Act.

(1) Home language surveys were conducted to determine if a child came from a home in which a language other than English was normally relied on for communication, either by the child or by the child's parents.

(2) When the home language survey indicated the influence of a language other than English, the child was given one of eight English language proficiency tests sanctioned by the state to determine if the child was proficient in *speaking and understanding* English. If the child was not proficient in these areas, that child was considered to be LEP.

(3) The academic level of the students who were found to be proficient in speaking and understanding English was also reviewed based on the results of standardized achievement tests which show proficiency in *reading and writing* English. Those students scoring below the 40th percentile on the achievement test in comparison to national achievement tests were also considered LEP.

Most districts, however, automatically categorized students as LEP even if the home language surveys showed that the child spoke only English and their parents only occasionally spoke Spanish. In addition, most districts did not have a procedure to ensure that students' difficulties in reading and writing English was the result of the influence of a non-English home language.

Id.

15. Dulay & Burt, *Aspects of Bilingual Education for LES/NES Students*, in BILINGUAL PROGRAM, POLICY, AND ASSESSMENT ISSUES (Convocation on Bilingual Program, Policy, and Assessment Issues ed. 1980).

16. Consent Decree, *supra* note 1, ch. 1, § V(A).

17. NATIONAL INSTITUTE OF EDUCATION, REPORT ON THE TESTING AND ASSESSMENT IMPLICATIONS OF THE TITLE VI LANGUAGE MINORITY PROPOSED RULES 75-76 (1981) (hereinafter

Even if the tests specified for use in Denver are the "best" available tests, their problems, nevertheless, are so severe that they should not be used for educational decision-making. Indeed, "[t]he technical capabilities of the field is not yet able to meet the needs or requirements of the proposed procedures."¹⁸

Language proficiency tests, in fact, may be less reliable in predicting whether a student needs to be in a bilingual classroom than are the opinions of those teaching in the bilingual program. A review of studies that compared language proficiency tests to bilingual teacher judgments of the students' proficiency found that teacher judgments did the better job.¹⁹ We must ask what is the point of prescribing elaborate surveys and testing programs to replace teacher judgments when the empirical literature seems to show teacher judgments work better than the tests?

The relatively large proportion of programs that set out to improve reading and math skills but succeed only in making student performance worse than if no special help had been given is surprising.²⁰ This result could well be a consequence of the failure of the selection procedures. If so, since the *Keyes* agreement calls for a flawed selection process, we can predict that the *Keyes* bilingual program is doomed to fail.

II.

MULTIPLE CAUSES OF POOR PERFORMANCE IN SCHOOL

A method of selecting students for bilingual programs must be able to identify students who do poorly in school because their dependency on a non-English language prohibits their successful participation in regular English-speaking classrooms. To do this, the selection method must separate students who do poorly in school because of this language barrier from those who do poorly in school for other reasons, since studies have shown that limited knowledge of English is but one factor affecting classroom performance in language-minority students.

One study addressing this issue looked at a nationally representative sample of 15,000 students in grades one through six.²¹ Using standardized test scores and home language data from parent interviews, the researchers were able to construct an analogue to the procedures included

cited as NATIONAL INSTITUTE OF EDUCATION). The five major research studies discussed by the National Institute of Education included the LAS and/or BSM tests. See also, Ramirez, *supra* note 8.

18. NATIONAL INSTITUTE OF EDUCATION, *supra* note 17, at 77.

19. *Id.* at 54-62.

20. K. Baker & A. de Kanter, *Federal Policy and the Effectiveness of Bilingual Education, BILINGUAL EDUCATION: A REAPPRAISAL OF FEDERAL POLICY* (1983); DEPARTMENT OF EDUCATION, *EFFECTIVENESS OF BILINGUAL EDUCATION: A REVIEW OF THE LITERATURE* (Final Draft Report) (1981) (available on ERIC, *supra* note 9, Ed No. 215-010).

21. Rosenthal, *supra* note 3.

in the proposed *Lau* regulations for identifying students in need of bilingual education.²² Once these students were identified, their achievement level was analyzed to see to what extent their low achievement was a result of their non-English language background or the result of other home background factors.

Although the analysis did find a small home language background effect on achievement, other home background factors were much more prominent as causes of the students' poor performance in school.²³ In short, the procedures prescribed in the proposed *Lau* regulations fail to sort out students who have problems in school because of the language barrier from students who have problems in school for other reasons. That is, the procedures cannot distinguish between students who need bilingual education and those who need some other type of compensatory education. The *Keyes* procedures, in fact, will fare worse than the proposed *Lau* regulations because *Keyes* omits some corrective methods that were included in the proposed *Lau* regulations.

In another study conducted among Hispanic students, the researchers found that once socioeconomic status was controlled, the use of Spanish at home had *no* relationship to achievement as measured by a standardized achievement test.²⁴ In other words, Hispanic students have problems in school because they come from poverty backgrounds. Using Spanish at home had nothing, or very little, to do with their poor educational performance.²⁵

For our purposes, we can take the results of these studies as indicative of how difficult it is to identify the educational effects of a non-English language background. If these sophisticated research studies cannot isolate the effects, there is little likelihood that the crude methods prescribed in *Keyes* can succeed. As the Mexican-American Legal Defense and Education Fund attorney involved in the *Keyes* litigation has noted:

The art of language assessment is not sufficiently sophisticated to identify those children who are bilingual and to trace their underachievement to their linguistic background. Indeed, given the multiple variables that affect academic success, it seems unlikely that there will ever be a single

22. *Id.*

23. *Id.*

24. G. MAYESKE, J. OKADA, W. COHEN, A. BEATON, JR. & C. WISLEER, A STUDY OF THE ACHIEVEMENT OF OUR NATION'S STUDENTS (1973).

25. Other studies, also, have found that both language background and other home background factors are related to the educational attainment of Hispanics. See, e.g., A. So & K. Chan, What Matters? A Study of the Relative Impact of Language Background and Socioeconomic Status on Reading Achievement (1982) (available from the offices of the LA RAZA LAW JOURNAL, *supra* note 8); C. VELTMAN, RELATIVE EDUCATIONAL ATTAINMENT OF HISPANIC-AMERICAN CHILDREN, 1976: A COMPARISON TO BLACK AND WHITE ENGLISH LANGUAGE CHILDREN (1980). This literature suggests that nonlanguage factors may be a stronger determinant of poor school performance in language minority students than are language factors.

test which relates the degree of underachievement to the student's linguistic background.²⁶

III.

RECLASSIFICATION

Having placed students in bilingual education, the *Keyes* agreement next details the method of deciding when these students are ready to make the transition to the regular English-speaking classroom. The *Keyes* agreement keeps students in the bilingual education program until they pass the oral proficiency test and score above the thirtieth or fortieth percentile, for elementary and middle/secondary students, respectively, on a standardized test.²⁷

In addition to the reliability and validity problems encountered in standardized and language proficiency tests discussed above, the reclassification procedures have a further serious flaw: they may permanently keep a child in bilingual education although the student is no longer limited English proficient.

Another thought experiment makes this problem clear. The standardized test score of a LEP student can be thought of as consisting of two parts: (1) a language barrier and (2) other background factors that cause poor performance. For example, assume that a LEP student scored at the twentieth percentile when he entered the program—thirty percentile points below the national norm. The two parts of his poor performance are distributed this way: five percentiles were lost due to the language barrier and twenty-five percentiles were lost because he came from a poverty background. Assume the bilingual program was a complete success and it removed the language barrier. His test score would now increase by five percentiles, the size of the language barrier. After he no longer needed bilingual education, he would score at the twenty-fifth percentile and would be kept in the program. No matter how much more bilingual education he was given, his score would never rise above the twenty-fifth percentile because, once the language barrier effect is eliminated, no further achievement gains will result from additional bilingual education.

Consider another hypothetical problem. Imagine that the errors we know exist in the entry classification methods had produced a student population for the bilingual program that was one hundred percent monolingual English-speaking. Would the reclassification procedures get monolingual English-speaking students out of the bilingual program? The answer is "no." These students would remain in the bilingual pro-

26. P. Roos, *Bilingual Education: The Hispanic Response to Unequal Educational Opportunity*, LAW AND CONTEMP. PROBS., Autumn, 1978, at 119-20.

27. Consent Decree, *supra* note 1, ch. 1, § III; see also Baker & de Kanter, *supra* note 1.

gram if no measurement error appeared in their test scores. Standardized tests are designed to place forty percent of all monolingual English-speaking students at or below the fortieth percentile score at each grade level. If a student scoring at the fortieth percentile in grade ten learns the normal amount over the school year, he will still score at the fortieth percentile a year later. The tests are designed this way. When the entry classification procedures place an English-speaking student in the bilingual program, he will never escape under the *Keyes* reclassification procedures.

These inherent problems in the *Keyes* procedures are, in reality, worse than in the hypothetical examples. The LEP population does poorly in school largely because of nonlanguage home background disadvantages.²⁸ Although only forty percent of the national, monolingual English-speaking population will fail the *Keyes* high school reclassification criterion, a far larger proportion of LEP students whose language barrier has been corrected will fail the criterion because of their greater concentration in disadvantaged home backgrounds.

This analysis has two interesting implications. First, future evaluations of the Denver program will falsely judge the program a failure because students will not pass the reclassification standard. But, as we have seen, this is inevitable even if the program succeeds in removing the language barrier. By looking *only* at the test scores, evaluators can mistakenly conclude that the program is a failure because few students transfer out of the program. Advocates, then, will return to court demanding even more bilingual education. Second, once the language barrier has been corrected, students who continue to score low need other kinds of help, not further bilingual education. But these reclassification procedures will deny them the help they need by needlessly keeping them in the bilingual program.

IV.

CONCLUSION

The procedures specified in the *Keyes* agreement will fail to identify many students who need and can benefit from the bilingual education program prescribed in the *Keyes* agreement. The *Keyes* procedures misclassify students in two ways. First, some students who could benefit from special language help are excluded from the bilingual program. Second, and more importantly because of its more likely occurrence, the *Keyes* procedures misclassify students who will not benefit from bilingual education as needing bilingual education.

Placing misclassified students in a bilingual education program

28. Rosenthal, *supra* note 3; see text accompanying notes 21-26.

when their poor school performance in school is not a consequence of a language barrier is harmful to the students and society for several reasons. First, it wastes scarce resources. Spending money on a special program for students who cannot benefit from that program means that other needy students elsewhere in the school system will be denied the extra help they need. Second, students who do not need bilingual instruction will not benefit from it. Third, providing bilingual instruction to students who do not need it deprives them of a chance of receiving the special help, such as compensatory education, they do need. Fourth, providing bilingual education to students who do not need it is worse than ineffective. It is harmful. To teach an English-speaking student math in Spanish, for example, not only denies him a meaningful math lesson, it denies him the opportunity to learn other subjects. Classroom time is fixed and the addition of a math lesson in Spanish means that some other subject has to be eliminated from the school day.

This analysis has examined whether the *Keyes* procedures for selecting and maintaining students for bilingual services are compatible with the bilingual instruction called for by the *Keyes* agreement. It appears that the two are not compatible. However, it does not necessarily follow from this analysis that the fault lies in the identification procedures. Two separate issues must be kept in mind. First, who are the children entitled to some kind of special help? Second, once this class of students is identified, what kind of special help do they need? If the procedures specified in the *Keyes* agreement are appropriate for identifying students needing help, across-the-board bilingual education for all of them is the wrong program. The *Keyes* procedures identify a group of students with one common characteristic: poor English skills. Some of these students may be English monolinguals, some may be non-English monolinguals. A program of bilingual instruction is not best for all these students. Some would be much better off in an all-English special language program, but that option is ruled out in the *Keyes* agreement. If the intent of *Keyes* is to establish a bilingual program, then changes must be made in the selection procedures so that only those students who can benefit from such a program are selected for it. Otherwise, the *Keyes* agreement will deny many students an appropriate education.