4-1988

The Efficacy of Bilingual Education in Relation to Student Linguistic Type, Program Type, and Program Setting

Carol Stubbs Rice
Western Michigan University

Follow this and additional works at: https://scholarworks.wmich.edu/dissertations

Part of the Bilingual, Multilingual, and Multicultural Education Commons

Recommended Citation
Rice, Carol Stubbs, "The Efficacy of Bilingual Education in Relation to Student Linguistic Type, Program Type, and Program Setting" (1988). Dissertations. 2161.
https://scholarworks.wmich.edu/dissertations/2161

This Dissertation-Open Access is brought to you for free and open access by the Graduate College at ScholarWorks at WMU. It has been accepted for inclusion in Dissertations by an authorized administrator of ScholarWorks at WMU. For more information, please contact maira.bundza@wmich.edu.
THE EFFICACY OF BILINGUAL EDUCATION IN
RELATION TO STUDENT LINGUISTIC TYPE,
PROGRAM TYPE, AND PROGRAM SETTING

by

Carol Stubbs Rice

A Dissertation
Submitted to the
Faculty of The Graduate College
in partial fulfillment of the
requirements for the
Degree of Doctor of Education
Department of Educational Leadership

Western Michigan University
Kalamazoo, Michigan
April 1988
The focus of this study was the differential effectiveness of bilingual education models as determined through the examination of Title VII project evaluations. The evaluation data were subject to a meta-analysis and the results compared to national norms.

The study examined the efficacy of bilingual education in relation to student type, program type and program setting, using data produced by projects from 1983-1986. The study design was based on the gap-reduction design, as modified for Title VII evaluations, with group means used as the unit of analysis. The study sample contained 2,500 monolingual and limited English proficient students from the Midwest. Analyses were conducted using scale scores calculated according to Title VII guidelines, with the final growth percentage expressed as the Relative Growth Index.

Study results indicated that participation in bilingual education programs produced small to moderate, though statistically significant, differences favoring bilingual education. These differences were found for reading,
language and mathematics. Additionally, results of this meta-analysis suggest that regardless of program setting, all three models studied have similar positive impact, whereas, program type research revealed the self-contained classroom as the most effective.

Based on the results of this study, there is no reason to assume that bilingual project students aren't doing as well as the national norms. Results consistently showed project participants keeping up with or catching up with non-participants.
INFORMATION TO USERS

The most advanced technology has been used to photograph and reproduce this manuscript from the microfilm master. UMI films the original text directly from the copy submitted. Thus, some dissertation copies are in typewriter face, while others may be from a computer printer.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyrighted material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each oversize page is available as one exposure on a standard 35 mm slide or as a 17" × 23" black and white photographic print for an additional charge.

Photographs included in the original manuscript have been reproduced xerographically in this copy. 35 mm slides or 6" × 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI
Accessing the World's Information since 1938

300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
The efficacy of bilingual education in relation to student linguistic type, program type, and program setting

Rice, Carol Stubbs, Ed.D.

Western Michigan University, 1988

Copyright ©1988 by Rice, Carol Stubbs. All rights reserved.
TABLE OF CONTENTS

LIST OF TABLES ........................................ iv
LIST OF FIGURES ......................................... v

CHAPTER

I. INTRODUCTION ........................................... 1
   The Problem ............................................. 4
   Operational Definitions ................................. 4
   Research Objectives .................................... 8
   Conceptual Framework .................................... 8

II. REVIEW OF THE RELATED LITERATURE ....................... 11
   Legislative History and Federal Policy of Bilingual Education ............. 11
   Bilingual Education Models ................................ 17
   Effectiveness of Bilingual Education ......................... 24
   Meta-Analysis and Title VII Project Evaluations ......................... 29

III. DESIGN AND METHODOLOGY ................................ 35
   Procedures ................................................. 35
   Subjects, Population, and Sampling .......................... 36
   Data Analysis .............................................. 37

IV. FINDINGS .................................................. 40
   Overview ................................................. 40
   Descriptive Information ................................... 40
   Project Description ....................................... 42
   Achievement Results ...................................... 55

ii

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Table of Contents—Continued

CHAPTER

Summary of Results.......................... 63

V. CONCLUSIONS AND RECOMMENDATIONS.......... 65

Do Different Models Make a Difference
With Different Types of Students?............. 66

Do Different Models Make a Difference
With Different Types of Programs?............. 68

Do Different Models Make a Difference
In Different Settings?.......................... 69

Limitations and Implications for
Further Research............................... 73

APPENDICES.............................................. 76

A. Conceptual Framework.......................... 77

B. Steps for the Implementation of the Gap-
Reduction Design and Relative Growth
Index............................................. 79

BIBLIOGRAPHY.......................................... 82
LIST OF TABLES

1. Status of Reports................................. 41
2. Description of Program by Year and State........ 41
3. Demographic Information.......................... 43
4. Language Groups and Grade Levels Served........ 44
5. Complexity of the Project........................ 45
6. Program Types...................................... 46
7. Emphasis of the Project.......................... 47
8. Program Models (Sequence of Language Use)...... 49
9. Literacy Goals..................................... 51
10. Entry, Reclassification, Exit, Follow-up........ 52
11. Percent of Teaching Staff Bilingual in English/Primary Language......................... 53
12. Type and Emphasis of Staff Training............. 54
13. Description of the Sample........................ 56
14. Test Characteristics................................ 57
15. Gains in Core Area by Grade Level............... 59
16. Effect of Program Area by Grade Level......... 60
17. Effect of Community Type on Relative Growth Index........................................ 61
18. Effect of Student Linguistic Characteristics on Relative Growth Index.................... 62
19. Effect of Program Type on Relative Growth Index........................................ 63
LIST OF FIGURES

1. Relative Growth Index ..................................... 39
2. Program Models/Types of Student .................... 67
3. Program Models/Types of Program .................... 69
4. Program Models/Settings .................................. 70
5. Comprehensive View of Academic Achievement
   Project by Project ........................................ 72
CHAPTER I

INTRODUCTION

Bilingual education, in a universal sense, means teaching and learning in two languages (Paulston, 1980). The term as used in the United States is quite different, connoting special programs for non- and limited-English-proficient, language-minority students (Tallmadge, Lam, & Camarena, 1985). There are generally two primary objectives of bilingual education: (1) to develop English language skills, and (2) to prevent language-minority students from falling behind their English-proficient peers in other content areas (Tallmadge et al., 1985). The student's native language is not taught, rather it is used as the medium of instruction for students whose proficiency in English is too limited for them to benefit from instruction presented in English. When the native language is taught, the rationale usually applies to early elementary students whose oral language is not English and who have not learned to read in their native language.

Formal federal involvement in bilingual education began with the Civil Rights Act of 1964 and was extended by the Bilingual Education Act of 1968 (Title VII) (cited in Crawford, 1987). When the Bilingual Education Act was launched, it was undertaken with little research to
support the concept and was literally thrust upon the public school system (Tallmadge et al., 1985). By 1974, the Supreme Court's decision, resulting from the *Lau v. Nichols* (1984) landmark case, made it clear that something other than regular school services had to be provided (U.S. Commission on Civil Rights, 1975). In spite of the lack of research, no provisions for funding of research were included in the original legislation. Not until 1977 were sufficient funds appropriated to bilingual education research to make a serious attempt at evaluating program effectiveness (National Institute of Education, 1976). But, even that decision left it up to the state educational agencies to decide what services were appropriate. Consequently, there is still a great deal of controversy over both federal and state policies regarding these programs, despite the fact that federally funded bilingual education in the public schools is almost two decades old.

Diverse conclusions from the few existing reviews of literature on the effectiveness of bilingual education have provided no ready answers for policy makers and have mainly fueled the argument both supporting and opposing bilingual education. The one large-scale national evaluation of the Title VII bilingual programs (Danoff, Coles, McLaughlin, and Reynolds, 1977a, 1977b, 1978) also produced more controversy than answers. (Willig, 1985, p. 269)

The controversy revolves around the quality and legitimacy of bilingual program evaluation as conducted by local programs. The issue of ethics arises as all limited-English-proficient students are supposed to be served
by bilingual programs, thus, an equivalent control group becomes an impossibility in experimental design. Additionally, "When systematic reviews of the bilingual education evaluation literature were conducted (e.g., Baker & de Kanter, 1983b; Dulay & Burt, 1978b; Okada, Besel, Glass, Montoya-Tannatt, & Bachelor, 1982, 1983) only a few evaluations could be identified that met minimal standards of methodological adequacy" (Tallmadge et al., 1985, p. viii).

In order to measure bilingual program effectiveness and gain credibility in the educational community, it is imperative that the quality of research and evaluation be upgraded. Willig (1985) suggested that a solution to the evaluation dilemma would be research that emphasized a comparison of program models under different kinds of conditions without trying to compare their effects to non-existent control groups as has been done in the past. Similarly, Gonzalez (1979) wrote that a wider range of program models should be tested, and that a research question should not be, "Does bilingual education work?" but, rather, "What type of bilingual schooling works best, with what populations, and under which circumstances?" (Gonzalez, 1979, p. 6). Okada et al. (1983) again struck the cord of differential effectiveness of specific bilingual education approaches, citing specifically the lack of research. Most recently, Tallmadge et al. (1985), in
their development of a proposed evaluation system for bilingual education, had two specific objectives, the second of which was meant to address the questions of what kinds of treatments are most effective for what kinds of students. The improvement of the quality of Title VII project evaluations was the ultimate objective of Tallmadge's et al. proposal, and in so doing hoped that ultimately the question of, "How effective are different bilingual education treatments for different types of students in different settings" (p. 87) would be answered.

The Problem

This study addresses a major gap in our knowledge about bilingual education by addressing the question, Is there differential effectiveness among extant models for bilingual education by type of student being served, type of program, and setting in which the program is offered? Okada et al. (1983) pointed out the absolute void of research and evaluation regarding the appropriateness of the various models. They cited this lack of research as a major problem for researchers and program developers in knowing what types of programs work for the many diverse linguistic student populations.

Operational Definitions

The independent variables in this study are the
various bilingual program models, settings, and student types, with the dependent variable of effectiveness as measured by several achievement instruments. The following are program model titles, program types, program settings and student characteristics with brief descriptions.

**Bilingual Program Models**

**Early-Exit Transitional:** Native language instruction is used, but only until students are proficient enough in English to benefit from all English instruction. The goal is to get to an all English curriculum as soon as possible (Tallmadge et al., 1985). This is the most frequently implemented model in the United States (Gonzalez, 1979).

**Late-Exit Transitional:** Instruction in both the native language and English is used for the duration of the program. The goal is proficiency and competency in both languages (Tallmadge et al., 1985).

**English as a Second Language (ESL):** Language minority students are placed in regular instruction for most of the day. During some part of the day extra help in English is given based on a special curriculum designed to teach ESL. The native language may or may not be used (Ovando & Collier, 1985).

**Immersion:** All instruction is in English although structured so that it does not assume prior knowledge of
English. Students may ask questions in the native language, but the teacher answers in English (Tallmadge et al., 1985).

**Structured Immersion:** The same as immersion with one important difference: native language arts classes are included (Ovando & Collier, 1985).

**Submersion:** The absence of any bilingual education intervention with language minority students. This is not allowed under the Lau ruling (Walton, 1984). This is commonly referred to as the "sink or swim" model.

**Maintenance:** The goal is to obtain proficiency in both languages with students maintaining and improving their capacity to speak their first language. Native English speakers are frequently included in these programs (Walton, 1984).

**Bilingual Program Types**

**Self-contained:** The regular classroom teacher provides the bilingual instruction.

**Pull-out:** Students leave the regular classroom for bilingual and/or ESL instruction provided by a specially trained teacher.

**Team Teaching:** Teachers who use English in instruction are paired with teachers who use the primary language of the student.

**Language Proficiency Grouping:** Grouped for
instruction in the classroom based on proficiency.

The preceding program types were described by Okada et al., 1982.

Bilingual Program Settings

Urban: Over 2,500 population
Rural: Population of 2,500 or less (King, 1986)

Bilingual Program Student Characteristics

Language Category: (a) Monolingual in the native language, (b) Dominant in a language other than English, (c) Bilingual in both English and native language, (d) Dominant in English, (e) Monolingual in English (Michigan Department of Education, 1986).

Low Income: An annual income for a family which does not exceed the poverty level determined pursuant to section 111(c)(2) of Title I of the Elementary and Secondary Education Act of 1965 (Federal Register, 1985)

Effectiveness: A measure of English language arts and content area achievement in English as measured on nationally normed achievement tests. Model A, the norm-referenced model of the RMC Research Corporation, is used in determining the difference between predicted posttest performance level and actual posttest performance level.
Research Objectives

The basic objectives of this study are:

1. To determine the effectiveness of different models of bilingual education as they are presently being practiced in the Midwest (Michigan, Iowa, South Dakota, North Dakota, Minnesota, Illinois, Indiana, and Wisconsin).

2. To investigate the possibility of bilingual education models having differential effects on different student types.

3. To investigate the possibility of bilingual education models having differential effects in different settings.

4. To investigate the possibility of bilingual education models having differential effects in different types of program.

Conceptual Framework

Most scholars, according to Paulston (1980), seem to agree on the basic phenomena behind Title VII bilingual education programs in the United States. That is, that there are a number of children from a background of low socioeconomic status who speak no or poor English and who encounter massive school failure and limited participation in the economic life of the country. The major basic assumption underlying the programs is that of unequal
opportunity and the belief that bilingual education helps equalize opportunity. The problem as spelled out in the Bilingual Education Act (cited in Crawford, 1987) reads as follows:

The Congress declares it to be the policy of the United States, in order to establish equal educational opportunity for all children (A) to encourage the establishment and operation, where appropriate, of educational programs using bilingual educational practices, techniques and methods, and (B) ... to provide financial assistance to... educational agencies ... in order to ... develop and carry out such programs ... which are designed to meet the educational needs of such children; and of demonstrating effective ways of providing, for children of limited English speaking ability, instruction designed to enable, while using their native language, to achieve competence in their English language. (Geffert, Harper, Sarmiento, & Schember, 1975, p. 13)

Thus, the immediate objective of bilingual education programs is to equalize opportunity for children from limited English speaking families by compensatory training in English.

The literature makes a compelling case that special programs in school can improve the achievement of language minority children. There is no conclusive evidence, however, that any specific program is more effective than any other (Baker & de Kanter, 1983b). Research evidence points to the fact that a determination needs to be made as to which programs are most effective with which types of children in which locations.

There are many ways to teach language minority
students. This study examined selected models and their effectiveness with specific types of students and in specific settings.

A review of the literature related to bilingual education models and effectiveness is found in Chapter II, followed by the research design of this study in Chapter III. The findings of the study are enumerated in Chapter IV, with final conclusions and recommendations discussed in Chapter V.
CHAPTER II

REVIEW OF THE RELATED LITERATURE

The purpose of Chapter II is to review the literature related to the effectiveness of bilingual education models. The review is divided into four areas of study: (1) legislative history and federal policy of bilingual education, (2) bilingual education models, (3) effectiveness of bilingual education, and (4) meta-analysis and Title VII project evaluations.

Legislative History and Federal Policy of Bilingual Education

Legislation and language minority education have almost reached the century mark. The intense immigration of the late 1800s of poor, illiterate Catholics from Southern Europe, resulted in the passage of an "English only" educational policy in 1890 (Lewis, 1980). Lewis pointed out that this policy was embraced throughout the first half of this century with the prejudice engendered by World War I encouraging it, and not until the early 60s was interest in a national language policy revived. The rash of social programs (Crawford, 1987) emanating from this period included a constitutionally guaranteed right of all resident children to free and equal educational
opportunity, from which the seed for bilingual education programs was planted (Tallmadge et al., 1985). Federal involvement in bilingual education began as a direct result of the 1960s and the civil rights movement, and of the interest of ethnic groups in maintaining their language and culture (Rotberg, 1982). Paulston (1980) wrote that it was the massive school failures of ethnic children which finally forced the authorities to acknowledge the existence of multilingualism and eventually to legislate into effect bilingual education programs.

Legislative enactments and related court ordered decisions at the federal and state levels have had a significant impact on bilingual education programs in recent years (Arista-Salado, 1978). The legislative history of these programs, as summarized from Tallmadge et al., (1985) follows:

The Civil Rights Act (cited in Crawford, 1987): This 1964 Act provided for appropriate instructional services to language-minority students, but did not address the language issue, nor provide any funds.

The Bilingual Education Act of 1968 (Title VII of the Elementary and Secondary Education Act; cited in Crawford, 1987): This Act provided funds for staff training, educational materials, and the implementation of special programs. Although the Act supported a transitional approach as a program model, school districts enjoyed a choice for
implementation. Since the program definition was vague, no specific evaluation criteria for determining effectiveness were provided. Paulston (1980), in discussing bilingual program history, pointed to a 1975 National Institute of Education (NIE) report that called the 1968 Bilingual Act a misnomer, as the long range goal was not bilingualism, rather English proficiency.

Lau v. Nichols (cited in Crawford, 1987): This 1975 landmark case of the U.S. Supreme Court was the result of a Chinese parent taking the San Francisco school board to court. The court ruled that equality of educational opportunity did include students who didn't understand English. Rotberg (1982), in her federal policy research, wrote that the court did not specify how equality should be accomplished. Rather it recognized several alternatives: "Teaching English to the students of Chinese ancestry who do not speak the language is one choice. Giving instructions to this group in Chinese is another. There may be others" (Rotberg, 1982, p. 31). Some of the confusion about federal legislation on bilingual education stems from this decision (Paulston, 1980). The impact of the Lau decision was most significant, according to Tikunoff and Vazquez-Faria (1982), because for the first time language rights were recognized as a civil right.

The 1974 amendments (cited in Crawford, 1987): These amendments to the Bilingual Education Act specified, in
great detail, the following:

1. Policies and procedures local and state educational agencies were expected to follow.

2. Direction to the Commissioner of Education to develop and disseminate bilingual program models.

3. The provision of funds for three landmark court decisions mandating bilingual instruction.

**Lau Remedies 1975** *(cited in Crawford, 1987):* These were not discussed by Tallmadge, but rather were obtained from Tikunoff and Vazquez-Faria (1982). The Lau Remedies were issued by the Office of Civil Rights to provide informal guidelines to the Lau ruling. The Remedies rejected the sole use of English as a Second Language (ESL) at the elementary level as an instructional approach with children who spoke little or no English. They implied that districts must establish bilingual instructional programs and/or equally acceptable alternatives. Although the legal status of the Lau Remedies is not certain, according to Paulston (1980), school districts judged out of compliance stand the risk of losing all their federal funding.

**The 1977 bilingual education regulations** *(cited in Crawford, 1987):* These regulations required that programs funded on a multi-year basis submit evaluation reports twice annually. Pre- and posttest reading scores for program and nonprogram students were to be included.
The 1978 amendments and 1980 regulations (cited in Crawford, 1987): The guidelines were altered to expand the definition of eligible participants, to permit some students of native English to participate, and to require more parental participation. The regulations emphasized institutionalization and provided for demonstration projects. Additionally, evaluation plans were upgraded and were used in making continuation awards.

EDGAR (Education Department's General Administration Regulations; cited in Crawford, 1987): The primary goal of these regulations was to increase the accountability of federally funded programs. EDGAR established criteria for judging the evaluation component in funding proposals.

The 1984 Amendments (cited in Crawford, 1987): The Bilingual Education Act was reauthorized in 1984, adding two significant new provisions. First, the school districts were required to inform parents of program availability and that they were voluntary; second, the authorization of funding for "special alternative" programs that did not require the use of native-language instruction. This amendment opened the door to alternative instructional strategies.

The preceding history has spanned two decades of language policy legislation. The growth of bilingual education programs in number and magnitude over the past decade is due in large measure to legislative and judicial
actions of the past 18 years according to Tikunoff and Vazquez-Faria (1982). They also wrote that many consider bilingual education primarily a political maneuver in that it recognizes and legitimizes the variety of cultures in our society. Vazquez (1978) pointed out that what is often overlooked is that bilingual education confirms the right of every child to a meaningful education. In summary, Tikunoff and Vazquez-Faria (1982) suggested a reasoning behind the confusion and controversy that has come to symbolize the federal involvement in bilingual education.

More than other educational processes, bilingual education operates in a context charged with controversy over the meaning of educational equity, minority group isolation, and whether societal institutions—the government and the school system—are obligated legally to foster and serve linguistic and cultural heterogeneity. (p. 236)

Navarro (1985) attempted to legitimize the case for legislation in the process. He saw bilingual education as a new approach to teaching English growing out of the compensatory efforts in the name of equal educational opportunity. "A common assumption of the enacting legislation is that children with limited English proficiency from low-income backgrounds can be helped out of poverty by overcoming their language barrier through compensatory programs" (Navarro, 1985, p. 291).
Bilingual Education Models

For more than a decade studies have indicated that the Spanish-dominant bilingual student is handicapped in a classroom where English is the only language of instruction (Coleman, 1966). Studies have also shown that in classrooms where Spanish and English are used in teaching, the performance of Spanish-dominant bilingual students improves (Weffer, 1972). Federal legislation and policy have created controversy and confusion around how best to approach language minority teaching. The two major issues pointed out by Tikunoff and Vazquez-Faria (1982) are: First, a consensus on how to provide bilingual education still does not exist, and second, no clear directive has been provided to establish the most effective approaches. Rotberg (1982) saw this confusion in her research on federal policy in the lack of goals and appropriate strategies for achieving goals for language minority children.

Program model refers to the basic broad classes of bilingual programs (e.g., transitional, maintenance) (Cohen, 1980). The Bilingual Education Act of 1968 (cited in Crawford, 1987) supported a transitional approach as a program model, meaning, the native language is used only until students are proficient enough in English to benefit from all English instruction (Tallmadge et al., 1985). Writing 11 years after the 1968 Act, Legarreta (1979) lamented the dearth of research on effective curriculum
models for language minority education, and wrote that current funded programs were using English as Second Language (ESL) models, that is teaching English with or without the use of the native language. Nowhere had federal policy endorsed this model, in that there was a lack of stated goals with appropriate suggested strategies (Carter, 1970) and that had recently been criticized by linguists (Saville-Troike, 1977; Legarreta, 1979). The Lau Remedies and Title VII favored transitional bilingual education, but suggested an ESL component be included in programs (Rotberg, 1982). Zirkel (1972) found that the bilingual model, native language and ESL, had generally positive results when compared to solely ESL, following the rationale of the Lau Remedies. Conversely, Moore and Parr (1978) compared four models: maintenance (teaching and maintaining native language and English), transitional (minimal native language), and immersion (nonbilingual) in one school district. Their results indicated that the immersion model was significantly more successful academically than any of the others.

A slightly different approach was undertaken by Sancho (1980) in his three-year longitudinal evaluation of maintenance and transitional models in a Title VII program. He concluded that neither model was sufficient as a sole approach, and that, "The effects of either a maintenance or transitional treatment are not as significant on
achievement as the degree of linguistic competence which the child initially brings to the school setting" (p. 3936A). In this same period, Legarreta's (1979) longitudinal study included a comparison of five models: immersion (English only), immersion plus ESL, and three bilingual models using different combinations of native language and English. Her findings showed significant differences favoring balanced bilingual treatment (half-day English and half-day native language instruction) most facilitative for both Spanish and English acquisition by children. This finding is reiterated by Cazden (1984) in her study of effective instructional practices in bilingual education that took her to 58 bilingual and monolingual classrooms. She stated that, "Evaluation research studies continue to find that in the most powerful educational environments for producing bilingualism both languages are used as the medium of instruction, but separately" (p. 16).

A large scale government multi-year contract study begun in 1980 by Cardenas et al., paralleled the proceeding private studies. After sampling 524 Title VII Basic Projects, 40% were found to use an ESL pull out (removed from the regular classroom for special English instruction) model. Although the majority were found to emphasize English instruction, many projects were making heavy use of the native language through a more transitional
model (Cardenas et al., 1983). At the same time bilingual legislation was still reflecting a more transitional approach, those programs receiving Title VII funds seemed to be reflecting an ESL approach. Cardenas noted that although few studies document bilingual education successes, there are many that show the effect of past immersion programs. In such programs language minorities dropped out before high school completion at a rate of 50 to 100% (Rotberg, 1982).

By 1980, the controversy surrounding the effectiveness of the one model proposed by the Lau Remedies, transitional bilingual education, had reached the White House Regulatory Analysis and Review Group (Baker & de Kanter, 1983b). As a result, the Department of Education undertook an extensive research program to answer the omnipresent question: Why is only one model of bilingual education being mandated by the federal government? The Final Draft Report submitted by Baker and de Kanter (1981) answered this question amid a whirl of controversy. They found that the literature supported the fact that special programs in school can improve the achievement of language minority children. They found no evidence that any specific model should be either legislated or preferred by the federal government. "The literature on the effectiveness of transitional bilingual education does not justify such heavy reliance on this one model of instruction" (Baker &
also pointed to the nationally representative Danoff et al. (1978) studies that found negative effects for TBE (Transitional Bilingual Education) in English. They concluded that a structured immersion (English instruction with minimal native language) approach may be beneficial for language minority students in the U.S. and may be superior to the transitional bilingual model used in most programs (Gersten, 1984). In effect, they stated that the case for bilingual education was extremely weak (Willig, 1984).

Navarro's (1985) article dealing with problems of language, education, and society suggested that, "Lawmakers do not want to make the politically difficult decision of clarifying the purposes of the law to the extent that some groups may feel that their interests have been excluded" (p. 289). He referred to the "divergent objectives" that exist between maintenance and immersion models as most prevalent, but also focused on the recent debate surrounding the differences between transitional bilingual education and immersion (Baker & de Kanter, 1983b). Finally, Navarro (1985) put forth the three most dominant options for educating the language minority child: maintenance bilingual education, transitional bilingual education, and immersion.

Presently, the U.S. Department of Education is in its
third year of a large scale four-year longitudinal study of three models: immersion, transitional and maintenance (Crawford, 1986). Crawford (1986) reported that on the basis of first year test scores limited English proficient students in bilingual programs outperformed immersion strategy students. He added that these results were the opposite of what the contracting agency had predicted. The agency cautioned that until study completion in 1988 "It would be totally unprofessional and reprehensible" for educators or researchers to draw any conclusions (Crawford, 1986, p. 10).

As the debate continues over which model is more effective for teaching language minority students, some researchers have said that the model is irrelevant when considered with other program characteristics. Paulston (1978) argued that certain societal factors and other program variables are considerably more important in determining academic achievement than the instructional model. Rotberg (1982) pointed out that few studies show one theoretical teaching technique to be clearly superior to another, and that program characteristics generally associated with quality, play a more important role in student achievement than does the initial language of instruction. She suggested that:

There is no educational basis for selecting an optimum instructional model for a country as large and diverse as the United States and that current findings do not indicate that the
transitional bilingual-bicultural approach advocated by the Lau Remedies and Title VII is better on average than other models. Analyses of different educational models, however, suggest that it may be possible to identify factors in certain communities which would favor one educational approach over another. (Rotberg, 1982, p. 34).

A Maryland county went as far as developing a policy statement to the effect that no single model could meet the needs of all language minority children and encourages a variety of models (Rotberg, 1982).

This review has revealed that federal policy and bilingual legislation have shown a strong preference for transitional bilingual models over alternative models. As research studies discussed herein reveal, there are various models with merit and no clear-cut evidence to suggest situations where student achievement is better or worse off because of the model used. Rotberg (1982) maintained that there was no research basis for the federal government to require a specific educational approach, and that past research suggested that federal government intervention in local decisions about instructional models was not productive. One of the highest programmatic priorities, according to Baker and de Kanter (1983b), should be to identify, discontinue and change harmful programs and to find out what differentiates programs achieving positive effects from those producing negative effects. A review of selected literature on bilingual program effectiveness follows.
Effectiveness of Bilingual Education

The most recent review of literature pertaining to bilingual program effectiveness (Tallmadge et al., 1985) revealed that little is known of the benefits of the 1.7 billion federal Title VII dollars and the unknown state and local dollars spent on bilingual projects, although four large scale governmental studies have been undertaken since 1972 as well as several private studies. Willig (1984) characterized the effectiveness studies as providing no ready answers and producing more controversy due to inconclusiveness and widely discrepant findings. Lack of research funds, according to Troke (1978), was a major deterrent to improvement of program quality, thus more conclusive studies.

Because of this lack of research on the differential effectiveness of specific bilingual education approaches, researchers and program developers find themselves, 14 years after the implementation of Title VII bilingual education, with very little sense of what types of programs or approaches work for or match the needs of the many diverse linguistic student populations. (Okada et al., 1982, p. 4)

The four large scale studies reviewed by Tallmadge et al. (1985) revealed only one that dealt with effectiveness in terms of student outcomes. The 1977 study by the American Institute for Research found that the Title VII students did not perform in English as well as their counterparts (Tallmadge et al., 1985). This study was widely criticized because of flaws in the design and
methodology of the research (e.g., Cardenas, 1977; Gray, 1977; O'Malley, 1978). But, Dulay and Burt (1979) maintained that although the results could not be used to judge effectiveness of bilingual education, they did reveal important data on Title VII programs. Of the other three studies, the 1972 Development Associates study explored policy, program management and operation, the 1973 General Accounting Office study examined Title VII evaluation reports and found them of little use, and the 1980 Significant Bilingual Instructional Features study defined successful bilingual treatments (Tallmadge et al., 1985).

A review of bilingual education research and evaluation studies through 1976 by Dulay and Burt (1979) revealed 50 findings from 12 studies relating to student performance. They found that 58% showed significantly positive effects for bilingual children who were dominant in their home language. In this same period, Troike (1978) concluded, after studying 12 programs, that quality bilingual programs could be effective for language minority children. Although Dulay and Burt (1979) were optimistic about the potential of bilingual education in the U.S., they saw bilingual program evaluation quality as a major problem and recommended improvement. Baker and de Kanter (1981) criticized Troike on this very point, and in turn they received Willig's (1982) criticism on their 1981
study for similar reasons.

Aside from the major governmental studies, several attempts have been made to assess program effectiveness using data from local evaluation reports. The first of these was the Zappert and Cruz 1977 study that reviewed 180 studies and evaluations and rejected 172 of them for lack of sound research design (Reyna, 1984). They concluded, according to Reyna, that bilingual education is an effective medium of instruction. Baker and de Kanter (1981) similarly reviewed 176 program evaluations and rejected all but 39 on the same grounds. But their conclusions ruled bilingual education as an ineffective medium of instruction (Baker & de Kanter, 1983b). During 1980 and 1981 four states, California, Michigan, Colorado, and Texas, conducted bilingual effectiveness studies using local evaluation data. The California and Colorado studies spanned three years. These studies, involving thousands of children and hundreds of bilingual education programs, have all found bilingual education to be an effective instructional approach (Reyna, 1984).

Although local evaluation reports recently have been highly criticized for their flaws in design and methodology, (e.g., Baker & de Kanter, 1981; Okada et al., 1982; Tallmadge et al., 1985; Willig, 1982) six doctoral dissertations of significance, based on similar reports, were written between the years 1971 and 1981. The majority of
the results reported from various regions of the U.S. reflected positive effects due to bilingual intervention with language minority students (Dimas, 1981; Olesini, 1971; Oteiza, 1981; Rivera, 1973; Romero, 1977; Zirkel, 1972). The preceding studies measured student outcomes based on achievement instruments and involved controls on which to compare those outcomes.

The bilingual effectiveness literature discussed to this point has involved bilingual studies performed in the United States and influenced by our federal policies. There is an extensive international literature which compares the effectiveness of various types of educational programs for language-minority children (Rotberg, 1982). Paulston (1978), in a comprehensive overview concluded that, "At the world level, the field of research on bilingual education is characterized by disparate findings and inconclusive results" (p. 187). She found that just as the U.S. controversy supports studies that satisfy virtually every possible opinion, so do other studies throughout the world. Rotberg (1982) reported that this conclusion was supported by numerous researchers of international studies, all of whom agreed that it is not possible to select an optimum educational approach for all situations.

Given the diverse opinions and focus of the various research presented in this review, it is appropriate that
Willig (1984) saw the major dilemma in bilingual education centering on the "effectiveness of bilingual programs in regard to school achievement and the diversity of research findings regarding such effectiveness" (p. 7). Sanders (1981) might have deduced a reasoning behind this mass of evidence on effectiveness in his understanding that, the purpose of gathering information on the bilingual programs was "to provide evidence of impact across funded projects to congressional decision makers" (p. 3).

Orum (1982), in her examination of the effects of the federal bilingual programs, asked if the effectiveness studies were truly to improve programs, then "bilingual programs must be examined using a broad-based student and institutional criteria which reflect the variety of ways that these programs may be changing the quality of education available to limited English proficient children" (p. 13). Swain (1979) had also pointed this out when she wrote of the necessity of taking into consideration differences in programs, children, and in communities. In considering the examination of the multiple differences across programs, Willig (1984) noted the impossibility of accounting for all these differences using the "traditional narrative review process" (p. 7). In sum, Swain (1979) pointed out, what is obviously still the case today, that "attempting to come to grips with all the literature, and all the contradictory conclusions reached
in the various research and evaluation studies, quite simply, boggles the mind!" (p. 23).

Meta-analysis and Title VII Project Evaluations

Evidence has shown that a great variation in the designs and reporting of bilingual program evaluations has limited the use of evaluation results in making generalizable conclusions about program implementation and the effects of bilingual education. Martin (1981) pointed out that information is needed in the areas of effectiveness and impact of program design and the success of different instructional models, but that there is an ambivalence in the commitment and interest in program evaluation. One of the contributing factors to this dilemma, according to Sancho (1980), stems from the political pressure under which bilingual education was undertaken with virtually no existing research upon which to build.

A major study was mounted by the National Center for Bilingual Research (NCBR) in 1982 to focus on the Title VII evaluations, in an attempt to synthesize the information in hopes of arriving at a more definitive conclusion as to the effectiveness of bilingual education programs (Cardoza, 1983). The method of synthesis used by the NCBR was meta-analysis developed by Glass in 1976. "Glass and his associates have applied his methods of meta-analysis to a variety of, heretofore, confused research literature.
In almost every case, the research literature has been clarified and general principles have been established" (Schmidt, 1985, p. 7). Cardoza (1983) acknowledged the promise of this technique in enhancing the knowledge of bilingual education because it provided a means of interpreting evaluation reports. In 1981 Martin pointed to the fact that several large scale meta-analysis studies had been completed regarding Title I for compensatory education, but as yet a national study had not been conducted for Title VII. Since 1981 two large scale meta-analyses have been conducted, the NCBR study which rejected more than 75% of the reports because of incomplete information (Tallmadge et al., 1985), and Willig's (1984) meta-analysis that experienced similar results, but who concluded that bilingual education was more effective than not.

Meta-analysis was a new approach to the problem of research integration, according to Hedges (1982), and has become an important supplement to the traditional methods of reviewing research. Okada et al. (1982) wrote that 30 years ago the traditional narrative summary was probably adequate, but due to the burgeoning research literature since the 1960s, "chronologically arranged verbal descriptions of research failed to portray the accumulated knowledge" (p. 6). "As a quantitative review, meta-analysis is distinguishable from primary research, secondary research, and narrative review" (Glass, 1977, p. 6).
Whereas primary research deals with treatment effects on a sample, secondary research re-analyzes these results, a narrative review informally evaluates critically a group of studies, a meta-analysis is distinguished from the preceding in its sample selection, data collections, and analysis (Bangert-Drowns, 1984). "Meta-analysis is nothing more than the attitude of data analysis applied to quantitative summaries of individual experiments . . . it is not a technique; rather it is a perspective that uses many techniques of measurement and statistical analysis" (Glass, McGaw, & Smith, 1981, p. 21).

The value of meta-analysis for summarizing the enormous quantity of Title VII program evaluations was acknowledged by Okada et al. in 1982. They knew that the information derived, "would estimate not only the overall magnitude of effects for each valued outcome, but the differential effectiveness of the programs in various contexts, for distinctive variations of the program and distinctive target populations" (p. 6). Meta-analysis is a method of summative evaluation, reiterated Willig (1984), that also answers the questions of how the characteristics of the various aspects of the studies and of the treatments influence outcomes. Slavin (1984), in his examination of how meta-analysis had been used in education, noted that a growing consensus indicated, that if properly used, meta-analysis could be a useful tool in
synthesizing research literature.

Criticisms of meta-analyses have primarily revolved around the issue of "combining apples and oranges" (Slavin, 1984). Slavin further explained that "combining the results of different studies runs the risk of producing an amalgam that makes no conceptual sense (e.g., Gallo, 1978; Presby, 1978; Wortman, 1983)" (p. 7). Another criticism, somewhat related, points to the tendency of emphasizing main effects, when the interactions often are most important (e.g., Cook & Leviton, 1980; McGlynn, 1982) (Slavin, 1984). A final criticism involves the method of choosing studies. Eysenck (1978) and Mansfield and Busse (1977) criticized the methodology for including poorly designed studies along with well-designed ones. Slavin (1984) judged that many criticisms of the meta-analytic method were more of a theoretical rather than a practical nature. He thought that the method could make an important contribution to research integration if used as a supplement to traditional reviews.

From the advent of the meta-analytic method in 1976, J. A. Kulik (1984) estimated that 320 studies had been conducted to 1984. Some of the areas of study enumerated by Kulik include: The effects of psychotherapy (Landman & Dawes, 1982; Shapiro & Shapiro, 1982; Smith, Glass, & Miller, 1980); the effects of coaching for the Scholastic Aptitude test (DerSimonian & Laird, 1983; J. A. Kulik,
Bangert, & C. L. Kulik, 1984; Messick & Jungeblut, 1981; Slack & Porter, 1980); the effects of desegregation (Crain & Mahard, 1982; Krol, 1979; Wortman & Bryant, 1984); and the effects of computer-based instruction (Burns & Bozeman, 1981; Hartley, 1978; C. L. Kulik, J. A. Kulik, & Bangert-Drowns, 1984; J. A. Kulik, Bangert, & Williams, 1983; J. A. Kulik, C. L. Kulik, & Cohen, 1980). Some meta-analyses in the area of education were listed by Slavin (1984). A summary of these include: The effects of class size on instruction (Glass, Cahen, Smith, & Filby, 1982; Hedges & Stock, 1983); special versus regular class placement of exceptional children (Carlberg & Kavale, 1980); ability grouping in secondary schools (C. L. Kulik & J. A. Kulik, 1982); cross-age tutoring (Cohen, Kulik, & Kulik, 1982); cooperative learning (Johnson, Johnson, & Maruyama, 1983; Johnson, Maruyama, Johnson, Nelson, & Skon, 1981) cues, participation, and corrective feedback (Lysakowski, & Walberg, 1982); and reading instruction (Pflaum, Walberg, Karegianes, & Rasher, 1980). J. A. Kulik (1984) acknowledged that these problems wouldn't be solved through meta-analysis alone, but hoped that increased reliability and dependability of conclusions would be enhanced. In sum, Glass et al. (1981), in writing about what would be gained in the process of meta-analysis, stated that "it is not uniformity in research reviewing and integrating that is desirable, rather it is
clarity, explicitness, and openness—those properties that are characteristics of the scientific method more generally and which impart to inquiry its 'objectivity' and trustworthiness" (p. 22). Bangert-Drowns (1984) concluded that meta-analysis is not a fad and reiterated the fact that "it is rooted in the fundamental values of the scientific enterprise: replicability, quantification, causal and correlational analysis" (p. 57). In terms of the effectiveness of bilingual education, in general, and Title VII projects, in particular, Okada et al. (1982) reasoned that meta-analysis, through the use of evaluation reports that have typically been excluded from reviews of research, would permit a different picture of bilingual effectiveness and put forth "reasoned and open discussion of possible causes" (p. 9).

The potential economic losses due to the obscurity of such data scattered in individual reports and the inability of social scientists to deliver generalizable answers to basic questions of policy are too great to allow us to treat research integration lightly. The potential benefits of meta-analytic method seem enormous. (Bangert-Drowns, 1984, p. 57)
CHAPTER III

RESEARCH DESIGN

Procedures

The independent variables in this study are the bilingual program models, settings, student types, and program types. The dependent variable is effectiveness as measured by standardized, norm-referenced achievement instruments. A meta-analytic approach was used to evaluate 1983-1986 Title VII basic program evaluation reports in an effort to arrive at a summary evaluation of the effectiveness of bilingual education program models.

Meta-analysis techniques, as originally proposed by Glass (1977) and modified for bilingual education by Okada et al. (1982), were used to integrate research findings and Title VII final evaluation reports. Steps involved in synthesizing the information include: defining the domain, locating and obtaining the evaluations, identifying the variables to be coded, coding the evaluations, planning the calculation of Effect Sizes, and analyzing the coded data. Each evaluation was reviewed and treated as an individual case, and the characteristics or variables were coded. Once all information was coded, statistical analysis was employed to investigate the relationship of
the study outcomes to the coded characteristics (Willig, 1984). Okada et al. (1982) developed a methodology of meta-analysis for testing its applicability to the existing body of literature on the effectiveness of bilingual education projects. The methodology was applicable to this study.

Subjects, Population, and Sampling

The domain of documents under consideration for this meta-analysis was limited to bilingual education projects funded through Title VII for the years 1983-1986, serving students in the grades kindergarten through grade 12. The study was limited to the 1983-1986 time period in order to obtain a representative sample of projects completing third year funding at various times, and to maintain a manageable study. This study focused on projects that had completed their third funding year, were located in the Midwest (Iowa, Illinois, Indiana, Michigan, Minnesota, North Dakota, South Dakota and Wisconsin), and concentrated on the evidence of effectiveness. If the project had not completed three years, then first and/or second year data was used. The study was limited to the Midwest, as the Federal Bilingual Administration has divided the United States into various multifunctional resource areas. The states comprising the Midwest area were examined in this study.
Data Analysis

In order to utilize the outcomes of diverse evaluations, a common scale was used. Meta-analysis provides a mechanism for converting these outcomes to a common scale through the application of various formulas (Glass et al., 1981). The common metric produced by these formulas is called an Effect Size (ES), a standardized difference between two mean scores, those of an experimental group and a control, or comparison group ($X_e - X_c / s_c$).

Since Effect Size refers to treatment-related growth and in most bilingual situations it's not possible to obtain valid no-treatment expectations, Tallmadge (1985) suggested using gap-reduction or group criteria-master designs for an evaluation strategy. A similar strategy, Model A, as developed by the RMC Research Corporation, was used with Title I programs. The analysis divides the project group's standardized growth estimate by the comparison group's standardized growth estimate; expresses the result as a percent, with the result being that a Relative Growth Index (RGI) of 100% would correspond to exactly keeping up. Anything above 100% would signify gap-reduction, and anything below 100% would mean that the students were falling behind (Tallmadge et al., 1985).

The gap-reduction design:

1. Is a systematic procedure for measuring the cognitive growth of project students and comparing it to
the growth of nonproject comparison groups.

2. Quantifies student growth relative to a nonproject comparison group and leaves the question of project impact to the interpretation of the findings.

3. Has easy-to-meet requirements for a comparison group (national, state, or local norms; mainstream classmates of the project participants; comparable groups in other schools, etc.).

The gap being measured is a gap between the mean achievement level of the project group and the mean achievement level of the nonproject comparison group. It is expected (or at least hoped) that the gap will be reduced as a result of the bilingual education intervention and will be smaller at posttest time than it was at pretest time.

The comparison group's pre- and posttest standard deviations are pooled and growth estimates for both groups are projected. Finally, the growth of the project group is expressed as a percentage of the growth of the comparison group, thus providing an easy-to-interpret Relative Growth Index (RGI) as illustrated by Figure 1.
Meta-analysis techniques were used to integrate the information contained in the Title VII bilingual education evaluation reports. The characteristics of each report were quantified in an effort to assess the implementation and results of the program. Traditional descriptive statistics (e.g., mean, median and standard deviation) were used to associate the characteristics and outcomes reported in the studies. The results are presented in two forms: (1) a description of all the documents reviewed; and (2) the meta-analyses techniques in chart form.
CHAPTER IV

FINDINGS

Overview

Results reported in this section are based on several different phases of the meta-analysis. In the first section is described the results contained in the descriptive information of the coding form. The second set of results presented identifies those aspects of the projects which correlate with learning through project description. Third, in order of presentation, is the description of the sample of student groups and the analyses of student achievement data.

Descriptive Information

A total of 47 documents were reviewed with 19 of the documents rating inclusion in this study. The remaining 28 were excluded with no usable data, either because of their use of non-standardized tests or no achievement outcome data at all. Results are shown in Table 1.
Table 1

Status of Reports

<table>
<thead>
<tr>
<th></th>
<th>Frequency</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exclude: Primary Criteria</td>
<td>28</td>
<td>60</td>
</tr>
<tr>
<td>Include</td>
<td>19</td>
<td>40</td>
</tr>
<tr>
<td>Total</td>
<td>47</td>
<td>100</td>
</tr>
</tbody>
</table>

Presented in Table 2 are data regarding the number of years the projects had been operative through the 1985-1986 school year, and the states from which the reports originated. Virtually 90% of the reports came from projects in their second or third year of existence.

Table 2

Description of Program by Year and State

<table>
<thead>
<tr>
<th>Project Year</th>
<th>Frequency</th>
<th>%</th>
<th>State</th>
<th>Frequency</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>First</td>
<td>2</td>
<td>10</td>
<td>IA</td>
<td>3</td>
<td>15</td>
</tr>
<tr>
<td>Second</td>
<td>10</td>
<td>53</td>
<td>IL</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>Third</td>
<td>7</td>
<td>37</td>
<td>IN*</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>MI</td>
<td>8</td>
<td>42</td>
</tr>
<tr>
<td></td>
<td>100%</td>
<td></td>
<td>MN</td>
<td>4</td>
<td>21</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>ND</td>
<td>2</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SD</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>WI*</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

* No Title VII Projects
Project Description

Shown in Table 3 is the demographic information. Tables 4 through 10 contain information describing the complexity, organization and emphasis of the projects, as well as the categories of students who were served, their primary languages, and criteria used for their entry into and exit from the projects. Also shown are proportions of type of evidence of project implementation. Staff characteristics are found in Tables 11 and 12.

All projects provided information regarding student demographics as shown in Table 3. Urban areas were represented in 74% of the 19 projects, and rural representation comprised 16%.

Table 4 portrays the language groups served by the projects, at each grade level. Projects serving students whose primary languages are Spanish and Arabic are the most frequently represented, with six having other first languages.

As can be seen in Table 5, the complexity of the projects was very well reported. Eighty-four percent of the projects reported that they operated in more than one school in a single school district. Eleven percent functioned in a single school only.
<table>
<thead>
<tr>
<th>Demographic Information</th>
<th>Frequency</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Type of Community</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urban</td>
<td>14</td>
<td>74</td>
</tr>
<tr>
<td>Rural</td>
<td>3</td>
<td>16</td>
</tr>
<tr>
<td>Mixed</td>
<td>2</td>
<td>10</td>
</tr>
<tr>
<td>Unmentioned</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>SES of Project Students</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Middle</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>Low</td>
<td>16</td>
<td>84</td>
</tr>
<tr>
<td>Unmentioned</td>
<td>2</td>
<td>10</td>
</tr>
<tr>
<td><strong>Immigrant Status</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Over 75% immigrants</td>
<td>10</td>
<td>53</td>
</tr>
<tr>
<td>50-75% immigrants</td>
<td></td>
<td></td>
</tr>
<tr>
<td>25-50% immigrants</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0-25% immigrants</td>
<td>8</td>
<td>42</td>
</tr>
<tr>
<td>Insufficient information</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>Language</td>
<td>Frequencies by Grade Level</td>
<td>Jr. High</td>
</tr>
<tr>
<td>--------------------------</td>
<td>----------------------------</td>
<td>---------</td>
</tr>
<tr>
<td></td>
<td>K 1 2 3 4 5 6</td>
<td></td>
</tr>
<tr>
<td>Spanish</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mexican &amp; Puerto Rican</td>
<td>6 91 93 72 87 61 69 56</td>
<td>246</td>
</tr>
<tr>
<td>Asian Pacific</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cambodian</td>
<td>- 12 7 8 3 6 9 40 120</td>
<td></td>
</tr>
<tr>
<td>Hmong</td>
<td>- - 48 48 - - - - 72</td>
<td></td>
</tr>
<tr>
<td>Lao</td>
<td>- 11 7 6 5 4 21 31 68</td>
<td></td>
</tr>
<tr>
<td>Vietnamese</td>
<td>- 13 4 10 - 2 8 31 71</td>
<td></td>
</tr>
<tr>
<td>Middle Eastern</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arabic</td>
<td>- 86 182 95 97 97 94 34 98</td>
<td></td>
</tr>
<tr>
<td>Native American</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arikora</td>
<td>- 10 14 12 13 6 9 26 -</td>
<td></td>
</tr>
<tr>
<td>Lakota</td>
<td>20 17 14 12 18 14 11 53 19</td>
<td></td>
</tr>
</tbody>
</table>
Table 5

Project Characteristics: Complexity of the Project

<table>
<thead>
<tr>
<th></th>
<th>Frequency</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single School</td>
<td>2</td>
<td>11</td>
</tr>
<tr>
<td>Single school district, more than one school</td>
<td>16</td>
<td>84</td>
</tr>
<tr>
<td>Single school district and private schools</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Consortium of school districts</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>Unable to classify</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>No information</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>19</strong></td>
<td><strong>100</strong></td>
</tr>
</tbody>
</table>

The bilingual program types are presented in Table 6. Although, more than one-half of the evaluation reports reported using self-contained classrooms, they also revealed the use of more than one type. For example, 14% used pull-out in conjunction with individualized learning plans.
<table>
<thead>
<tr>
<th>Program Type</th>
<th>Implicit or Mentioned</th>
<th>Emphasized</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Frequency</td>
<td>Percent</td>
</tr>
<tr>
<td>Self-contained</td>
<td>14</td>
<td>49</td>
</tr>
<tr>
<td>Pull-out</td>
<td>4</td>
<td>14</td>
</tr>
<tr>
<td>Team-teaching</td>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td>Language Proficiency</td>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td>Grouping</td>
<td>3</td>
<td>14</td>
</tr>
<tr>
<td>Individualize Learning</td>
<td>4</td>
<td>14</td>
</tr>
</tbody>
</table>

Provided in Table 7 is information regarding the emphasis placed upon various components in the projects, as described in the evaluation reports. Objectives were more frequently stated than not. In a few instances, no objective was stated, but material was presented indicating that achievement was evaluated. Improvement of program management and of evaluation were the less frequently mentioned non-instructional components. Use of primary language for subject matter (other than language and reading) was frequently not reported. The project evaluator usually assessed the achievement (or not) of the non-instructional, and the culture and history components, while objective measures were utilized to evaluate those...
components for which English and/or primary language instruction was provided. Results show that ESL/English language arts and math received the most emphasis.

Table 7
Project Characteristics: Emphasis of the Project

<table>
<thead>
<tr>
<th>Mentioned</th>
<th>Objective Stated</th>
<th>Objective Not Assessed</th>
<th>Objective Data Included</th>
</tr>
</thead>
<tbody>
<tr>
<td>F % F % F %</td>
<td>F %</td>
<td>F %</td>
<td>F %</td>
</tr>
</tbody>
</table>

Non-Instructional Components

- Development or acquisition of curriculum and instructional materials: 2 2 1 5 5 4.5
- Community/parent participation: 3 3 3 3 10 8.5
- Staff training: - 4 4 13 11
- Improvement of project management evaluation: - - 2 2

Instructional Components

- Culture & history: 3 3 7 6 2 2
- ESL/English language arts: - - 14 12 15 13
- English reading: - - 17 15
- Primary language reading: 1 .5 3 3 1 .5
- English language mathematics: - - 11 9
In Table 8 is presented documentation of the program models (sequence of language use) in which each category of student (NEP/LEP) participated. From the percentages shown for all projects, it would appear that NEP students are presented all English, while LEP students receive equal time in their primary languages in all categories, oral language, literacy and introduction of subject matter. The strategy for English reading percentages revealed a similar pattern for both NEP and LEP students. Just as many delayed teaching English reading as did not delay the process.

Depicted in Table 9 are the literacy goals of the projects. While sequence of language use may differ, a full 73% have English only as the literacy goal, with 27% including some measure of the primary language with English.

In Table 10 is found a description of the methods used by the projects to admit students into the bilingual programs, of reclassifying their English proficiency status, and of exiting them into the mainstream of American education. Selection criteria were specified in 63% of the reports, of those, 10% were unclear in reporting the actual criteria used. The other 53% explained both that objective measures were used and the criteria that students met for entrance into the program. Reclassification criteria were described in 68% of the documents,
Table 8

Project Characteristics: Service Provided to Students by Category

<table>
<thead>
<tr>
<th></th>
<th>ESL &amp; Immersion English Only</th>
<th>ESL &amp; Structured Immersion-English First, Primary Later</th>
<th>Early-exit Transitional-Primary First, English Later</th>
<th>Late-exit Transitional &amp; Maintenance. Both languages simultaneously</th>
<th>Unable to Classify or no Information</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oral Language</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NEP</td>
<td>5 26</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>LEP</td>
<td>3 16</td>
<td>-</td>
<td>1 5</td>
<td>10 53</td>
<td>-</td>
</tr>
<tr>
<td>Development of Literacy Skills</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NEP</td>
<td>5 26</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>LEP</td>
<td>4 20</td>
<td>-</td>
<td>3 16</td>
<td>7 38</td>
<td>-</td>
</tr>
<tr>
<td>Introduction of Subject Matter</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NEP</td>
<td>5 26</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>LEP</td>
<td>5 26</td>
<td>-</td>
<td>6 32</td>
<td>3 16</td>
<td>-</td>
</tr>
<tr>
<td>Strategy for Delayed English Reading</td>
<td>Not Delayed</td>
<td>Delayed Until English Oral</td>
<td>Delayed Until English Oral and Primary Reading</td>
<td>Unable to Classify or no Information</td>
<td></td>
</tr>
<tr>
<td>-------------------------------------</td>
<td>------------</td>
<td>----------------------------</td>
<td>-----------------------------------------------</td>
<td>-----------------------------------</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>F %</td>
<td></td>
</tr>
<tr>
<td>NEP</td>
<td>2 11</td>
<td>3 16</td>
<td>- -</td>
<td>- -</td>
<td></td>
</tr>
<tr>
<td>LEP</td>
<td>3 16</td>
<td>5 26</td>
<td>- -</td>
<td>2 11</td>
<td></td>
</tr>
</tbody>
</table>

The table continues with the following entries:

<table>
<thead>
<tr>
<th></th>
<th>Criteria</th>
<th>Stated</th>
<th>Criteria</th>
<th>Stated</th>
<th>Criteria</th>
<th>Stated</th>
<th>Unable to Classify or no Information</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Unstated</td>
<td></td>
<td>Unstated</td>
<td></td>
<td>Stated</td>
<td></td>
<td>F %</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>F %</td>
</tr>
</tbody>
</table>

Note: The table provides data on the classification of delayed English reading strategies for NEP and LEP, with categories for not delayed, delayed until English oral, delayed until English oral and primary reading, and unable to classify or no information.
with the majority, 47%, being explicitly stated. Criteria for exit into the mainstream were reported in 58% of the reports. Objective measures were cited by all, but the actual criteria for exit were explained in only 45%. Follow-up of students who had left the program was reported in 43% of the cases, but in only 37% were the procedures described.

Table 9
Project Characteristics: Literacy Goals

<table>
<thead>
<tr>
<th>Student Category</th>
<th>F</th>
<th>%</th>
<th>F</th>
<th>%</th>
<th>F</th>
<th>%</th>
<th>F</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>NEP</td>
<td>5</td>
<td>26</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>LEP</td>
<td>9</td>
<td>47</td>
<td>3</td>
<td>16</td>
<td>2</td>
<td>11</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

As shown in Table 11, only about 50% of the projects reported the proportion of their teachers and aides who were bilingual in English and the primary language of their students. From the reports, it appears that 47% of the aides are bilingual, whereas only 26% reported having a fully bilingual teaching staff. More than half (58%) of the projects did not report staff bilingualism.
Table 10
Project Characteristics: Entry, Reclassification, Exit, Follow-up

<table>
<thead>
<tr>
<th>Procedure</th>
<th>Subjective Judgement</th>
<th>Objective Tests Criteria Unclear</th>
<th>Objective Tests Criteria Stated</th>
<th>Not Mentioned or Unable to Classify</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>F</td>
<td>%</td>
<td>F</td>
<td>%</td>
</tr>
<tr>
<td>Entry</td>
<td>-</td>
<td>-</td>
<td>2</td>
<td>10</td>
</tr>
<tr>
<td>Reclassification</td>
<td>-</td>
<td>-</td>
<td>4</td>
<td>21</td>
</tr>
<tr>
<td>Exit</td>
<td>-</td>
<td>-</td>
<td>3</td>
<td>16</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Procedure</th>
<th>Mentioned</th>
<th>Procedures Described</th>
<th>Not Mentioned or Unable to Classify</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>F</td>
<td>%</td>
<td>F</td>
</tr>
<tr>
<td>Follow-up</td>
<td>1</td>
<td>6</td>
<td>7</td>
</tr>
</tbody>
</table>
Table 11

Staff Characteristics: Percent of Teaching Staff Bilingual in English/Primary Language

<table>
<thead>
<tr>
<th></th>
<th>Teachers</th>
<th></th>
<th>Aides</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Freq. %</td>
<td></td>
<td>Freq. %</td>
<td></td>
</tr>
<tr>
<td>0-20% Bilingual</td>
<td>1</td>
<td>6</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>21-40% Bilingual</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>41-60% Bilingual</td>
<td>2</td>
<td>10</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>61-80% Bilingual</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>81-100% Bilingual</td>
<td>5</td>
<td>26</td>
<td>9</td>
<td>47</td>
</tr>
<tr>
<td>Unable to Classify or no Information</td>
<td>11</td>
<td>58</td>
<td>9</td>
<td>47</td>
</tr>
</tbody>
</table>

In Table 12, the frequency and percent of occurrence for various types of staff training are reported. Intensive training in the primary language of the students was mentioned in two reports, although only 26% reported having bilingual teachers. Some form of inservice training was mentioned, described and/or evaluated in 90% of the reports. Fifty-seven percent of these evaluated the training. For 42% of the projects, there were reports of degree programs for teachers. Fewer projects, 37%, reported similar training for aides. Professional meetings were attended by teaching staff from 26% of the projects; nothing is known of the other 74%.
Table 12

Staff Characteristics: Type and Emphasis of Staff Training

<table>
<thead>
<tr>
<th>Type</th>
<th>Emphasis</th>
<th>Freq. %</th>
<th>Freq. %</th>
<th>Freq. %</th>
<th>Freq. %</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mentioned Only</td>
<td>Described</td>
<td>Described &amp; Evaluated</td>
<td>Not Mentioned</td>
<td></td>
</tr>
<tr>
<td>Individualized Training</td>
<td>2 10</td>
<td>3 17</td>
<td>2 10</td>
<td>12 63</td>
<td></td>
</tr>
<tr>
<td>Inservice/Local Activities</td>
<td>1 6</td>
<td>5 27</td>
<td>11 57</td>
<td>2 10</td>
<td></td>
</tr>
<tr>
<td>Degree/Teachers Courses/Teachers</td>
<td>2 10</td>
<td>4 22</td>
<td>2 10</td>
<td>11 57</td>
<td></td>
</tr>
<tr>
<td>Degree/Aides Courses/Aides</td>
<td>1 5</td>
<td>5 27</td>
<td>1 5</td>
<td>12 63</td>
<td></td>
</tr>
<tr>
<td>Language Training</td>
<td>-</td>
<td>-</td>
<td>2 10</td>
<td>-</td>
<td>17 90</td>
</tr>
<tr>
<td>Professional Meetings</td>
<td>-</td>
<td>-</td>
<td>3 16</td>
<td>2 10</td>
<td>14 74</td>
</tr>
</tbody>
</table>

From the preceding tables, depicting the information that could be gleaned from the 19 project evaluation reports subjected to analysis, it is apparent that much detail is missing about the nature of the target population, the staff who serve them, and the nature and manner of their bilingual instruction. Some of the provided
information was not clearly stated. In general, however, projects emphasized the staff development and instructional components of their effort. They tended to treat NEP and LEP students similarly, but with more use of the primary language with LEP students, and with emphasis on the development of proficiency in English with all.

Achievement Results

Description of the sample for which achievement results were analyzed is provided in Table 13. Except for kindergarten and 12th grades, students were fairly evenly distributed over all grade levels. They represented four major language categories with each group being substantially represented. The majority of the students were limited English proficient.

Characteristics of the tests used to measure student achievement will be found in Table 14. The California Achievement Test (CAT) was the most frequently used test, followed by the Science Research Associates (SRA) and Language Assessment Scales (LAS). All the testing was performed in English. Test editions and levels were generally stated; if not, projects were contacted.
Table 13
Description of the Sample

<table>
<thead>
<tr>
<th>Grade Level</th>
<th>Major Language Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>K</td>
<td>1% Spanish</td>
</tr>
<tr>
<td>1</td>
<td>9% Middle Eastern</td>
</tr>
<tr>
<td>2</td>
<td>15% Asian Pacific</td>
</tr>
<tr>
<td>3</td>
<td>10% Native American</td>
</tr>
<tr>
<td>4</td>
<td>9%</td>
</tr>
<tr>
<td>5</td>
<td>8%</td>
</tr>
<tr>
<td>6</td>
<td>9% English Proficiency</td>
</tr>
<tr>
<td>Jr. High</td>
<td>11% No Proficiency</td>
</tr>
<tr>
<td>Sr. High</td>
<td>28% Limited Proficiency</td>
</tr>
</tbody>
</table>

Results of the analyses of student achievement data are reported in Tables 15 through 19. Instead of trying to quantify how well the project was doing, the gap-reduction design implemented here examined how well the project students were doing relative to some nonproject comparison group. In the tables and figures that follow, it must be remembered that the pooled standard deviation of the comparison group's pre- and posttest is the metric in which growth estimates for the project and comparison groups are cast. The gap measured is the gap between the mean achievement level of the project group and the mean
achievement level of the non-project comparison group. Finally, the Relative Growth Index (RGI) represents the growth of the project group versus the comparison group. Again, an RGI of less than 100% indicates the project students falling further behind the non-participants during the evaluation period. RGI's equal to 100% indicate the project group growing at the same rate as the non-participants, and RGI's greater than 100% indicate that project participants are catching up with the non-participants. Similarly, a negative gap reduction signals the project group falling behind, while a positive gap reduction indicates catching up or keeping up with the comparison group.

Table 14
Test Characteristics

<table>
<thead>
<tr>
<th>Test Used</th>
<th>Test Edition Stated</th>
<th>Stated</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAT</td>
<td>33%</td>
<td>Yes</td>
</tr>
<tr>
<td>SRA</td>
<td>19%</td>
<td>Derived</td>
</tr>
<tr>
<td>LAS</td>
<td>14%</td>
<td>100%</td>
</tr>
<tr>
<td>MAT</td>
<td>10%</td>
<td></td>
</tr>
<tr>
<td>ITBS</td>
<td>10%</td>
<td></td>
</tr>
<tr>
<td>Gates-MacGinite</td>
<td>10%</td>
<td>Yes</td>
</tr>
<tr>
<td>STEP</td>
<td>4%</td>
<td>Derived</td>
</tr>
<tr>
<td></td>
<td>100%</td>
<td>100%</td>
</tr>
</tbody>
</table>

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Gains in core area achievement by the LEP students are presented by grade level in Table 15. Gains in language arts in these projects, in general, are higher than those for mathematics and reading. Project participants were catching up with non-participants. Gains in mathematics were higher than those for reading and appeared to be just keeping up, while the reading RGI's indicated a falling behind by the project group. The students in the elementary grades outgained secondary students in all three core areas, and were catching up with the comparison group. A decline was indicated beginning with junior high and continued through 12th grade, except for mathematics, which declined initially after elementary but regained after junior high.

The effects of the program model by core area are reported in Table 16. The great majority of models used in these projects were either English only (ESL) or maintenance. Only one of the core areas, reading, shows students falling behind. This occurred in the maintenance model. In mathematics and language the maintenance model showed higher gains than the ESL model, and they all showed students keeping up with comparison group students. The majority of these models are in urban settings.
<table>
<thead>
<tr>
<th>Core Area</th>
<th>Number of Students</th>
<th>Number of Projects</th>
<th>Gap Reduction</th>
<th>RGI</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Language Arts</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Elementary</td>
<td>597</td>
<td>7</td>
<td>.15</td>
<td>121.26</td>
</tr>
<tr>
<td>Jr. High</td>
<td>155</td>
<td>7</td>
<td>.23</td>
<td>134.49</td>
</tr>
<tr>
<td>Sr. High</td>
<td>485</td>
<td>6</td>
<td>.07</td>
<td>103.94</td>
</tr>
<tr>
<td><strong>Reading</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Elementary</td>
<td>1,509</td>
<td>12</td>
<td>.10</td>
<td>106.15</td>
</tr>
<tr>
<td>Jr. High</td>
<td>227</td>
<td>10</td>
<td>-.17</td>
<td>95.39</td>
</tr>
<tr>
<td>Sr. High</td>
<td>471</td>
<td>9</td>
<td>-.04</td>
<td>88.74</td>
</tr>
<tr>
<td><strong>Mathematics</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Elementary</td>
<td>1,412</td>
<td>10</td>
<td>.25</td>
<td>117.64</td>
</tr>
<tr>
<td>Jr. High</td>
<td>174</td>
<td>8</td>
<td>-.06</td>
<td>96.66</td>
</tr>
<tr>
<td>Sr. High</td>
<td>619</td>
<td>9</td>
<td>.09</td>
<td>115.62</td>
</tr>
</tbody>
</table>

Presented in Table 17 is the effect of community type on the relative growth index. Urban communities represent 74% of the projects, rural 16% and mixed 10%. In all the core areas the rural communities showed gains exceeding 100% RGI. They were achieving at the level of their comparison group counterparts. The urban students were catching up in language arts and mathematics, but were falling behind in reading. Similarly, the mixed
community, according to the table, was falling behind in reading, but catching up in mathematics.

Table 16
Effect of Program Model by Core Area

<table>
<thead>
<tr>
<th>Core Area</th>
<th>Number of Students</th>
<th>Number of Projects</th>
<th>Gap Reduction</th>
<th>RGI</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Language Arts</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ESL*</td>
<td>632</td>
<td>6</td>
<td>.09</td>
<td>102.94</td>
</tr>
<tr>
<td>Structured Immersion</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Early-Exit Trans-</td>
<td>47</td>
<td>1</td>
<td>.06</td>
<td>148.47</td>
</tr>
<tr>
<td>itional**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maintenance***</td>
<td>555</td>
<td>4</td>
<td>.09</td>
<td>115.79</td>
</tr>
<tr>
<td><strong>Reading</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ESL*</td>
<td>849</td>
<td>6</td>
<td>.03</td>
<td>107.94</td>
</tr>
<tr>
<td>Structured Immersion</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Early-Exit Trans-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>itional**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maintenance***</td>
<td>1,339</td>
<td>9</td>
<td>-.03</td>
<td>92.82</td>
</tr>
<tr>
<td><strong>Mathematics</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ESL*</td>
<td>849</td>
<td>6</td>
<td>.18</td>
<td>115.282</td>
</tr>
<tr>
<td>Structured Immersion</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Early-Exit Trans-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>itional**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maintenance***</td>
<td>1,337</td>
<td>8</td>
<td>.06</td>
<td>125.64</td>
</tr>
</tbody>
</table>

* All ESL projects are urban.  
** Early-Exit transitional is urban.  
*** Maintenance projects are 66% urban and 34% rural (Native American).
### Table 17

**Effect of Community Type on RGIs**

<table>
<thead>
<tr>
<th>Core Area</th>
<th>Number of Students</th>
<th>Gap Reduction</th>
<th>RGI</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Language Arts</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urban</td>
<td>1,188</td>
<td>.18</td>
<td>118.87</td>
</tr>
<tr>
<td>Rural</td>
<td>217</td>
<td>.09</td>
<td>129.62</td>
</tr>
<tr>
<td>Mixed</td>
<td>0</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><strong>Reading</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urban</td>
<td>1,888</td>
<td>-.05</td>
<td>93.48</td>
</tr>
<tr>
<td>Rural</td>
<td>259</td>
<td>.04</td>
<td>107.98</td>
</tr>
<tr>
<td>Mixed</td>
<td>50</td>
<td>-.15</td>
<td>69.93</td>
</tr>
<tr>
<td><strong>Mathematics</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urban</td>
<td>2,055</td>
<td>.03</td>
<td>109.83</td>
</tr>
<tr>
<td>Rural</td>
<td>90</td>
<td>.40</td>
<td>143.76</td>
</tr>
<tr>
<td>Mixed</td>
<td>50</td>
<td>.23</td>
<td>149.76</td>
</tr>
</tbody>
</table>

The effect of student linguistic characteristics on the relative growth index is revealed in Table 18. The majority of the students participating in the reported projects, as might be expected, were dominant in a language other than English. Although their RGI indicated that they were keeping up with the comparison group, they gained the least of the three linguistic categories reported. Those students classified as knowledgeable
equally limited in both English and native language showed significant gain and appeared to be easily catching up with the comparison group. Those students monolingual in their native language also showed gains indicating catching up with their English speaking counterparts.

Although students dominant in or monolingual in English are permitted to participate under Title VII rules, none were reported in the evaluations.

Table 18
Effect of Student Linguistic Characteristics on RGI

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Number of Students</th>
<th>Gap Reduction</th>
<th>RGI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monolingual in Native Language</td>
<td>876</td>
<td>.04</td>
<td>114.76</td>
</tr>
<tr>
<td>Dominant in Language Other than English</td>
<td>1,100</td>
<td>.15</td>
<td>101.87</td>
</tr>
<tr>
<td>Bilingual in both English &amp; Native Language or Equally Limited</td>
<td>523</td>
<td>.09</td>
<td>130.61</td>
</tr>
<tr>
<td>Dominant in English</td>
<td>0</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Monolingual in English</td>
<td>0</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Finally, a discussion of the effect of program type on the relative growth index per Table 19. The two basic program types, self-contained and pull-out, were represented in the evaluations, as well as in a combination of
the two. In a few evaluation reports it was not possible to identify the program type. Generally, the students appeared to be catching up with the nonproject comparison group. The projects using self-contained classrooms clearly succeeded in keeping up with and/or surpassing their counterparts. Projects using the pull-out method or a combination of the two appeared to be just keeping up. The results portrayed by this table are in keeping with all subsequent results.

Table 19
Effect of Program Type on RGI

<table>
<thead>
<tr>
<th>Program Type</th>
<th>Number of Students*</th>
<th>Gap Reduction</th>
<th>RGI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self-contained</td>
<td>1,110</td>
<td>.24</td>
<td>132.97</td>
</tr>
<tr>
<td>Pull-out (Individualized learning plans)</td>
<td>311</td>
<td>.06</td>
<td>102.56</td>
</tr>
<tr>
<td>Combination of the two</td>
<td>908</td>
<td>.01</td>
<td>103.63</td>
</tr>
</tbody>
</table>

* All students aren't represented because of the difficulty in some evaluations of identifying type of program.

Summary of Results

The preceding synthesis of results identified three major types of findings. The first major finding showed that the quality of the evaluation reports included contributed significantly to the results of the study. Those reports conformed to the exacting standards of the revised
coding form developed by Okada et al. (1982).

Project description, in terms of its correlation with learning, represents the second major group of findings. Again, the significance of the missing and/or incomplete data is most apparent in the description of populations and program procedures. Project variables such as these that play a role in the results of this study include student classification and reclassification. Too often program criteria was vague in this area.

The third major finding of this study involves student achievement data and the fact that, regardless of test, bilingual education programs, overall, produce positive effects in the academic realm. The evaluation reports generally revealed data in the fields of language arts, mathematics, and reading. Data was generally unavailable for science, social studies, self-concept and attitude toward school. Some projects did include descriptive check sheets in the affective area. In the chapter that follows the results reported are discussed according to project results and their future implications.
CHAPTER V

CONCLUSIONS AND RECOMMENDATIONS

The topic examined in this study is the differential effectiveness of bilingual education models in terms of types of students, programs and settings. This question was investigated using Title VII evaluation reports and the gap reduction research design as modified for bilingual education by Tallmadge (1985). Project means, using scale scores, were used as the unit of analysis, from which the gap-reduction and relative growth indexes were computed. The degree to which the project groups kept pace with the national comparison groups was reported for 13 grade levels on three subject area tests, reading, math and language. Conclusions based on the results of these analyses follow.

In spite of the sometimes inadequate quality of many of the evaluation reports included in this meta-analysis, the results have indicated that participation in bilingual education programs produces small to moderate, though statistically significant, differences favoring bilingual education. These differences are found for reading, language and mathematics. Additionally, results of this analysis suggest that regardless of the program setting in the Midwest, all three models, ESL, EET, and maintenance
have a similar positive impact. These findings suggest that the bilingual education program models studied, except for a few instances, produced results that show project participants keeping up with or catching up with non-participants.

The following discussion includes figures showing the EET model clearly superior to all others. This represents only one project and is discussed in the limitations section of this chapter.

Do Different Models Make a Difference With Different Types of Students?

As pointed out in Chapter II, earlier research emphasized one model or another, while more recent research, primarily, has shown that any model using the native language has generally outperformed the English only type. The findings of this study support recent research in this respect. When judging the differential effectiveness of program models for different types of student, the most effective model appears to be ESL for those students monolingual in their native language. But, for those students dominant in another language or bilingual, the relative growth index indicates the maintenance model as more successful. Figure 2 shows these results. Although, the results indicate that some models work better than others with particular students, these differences are not statistically significant and all models show a relative
growth index exceeding 100% for students who speak more than one language.

Relative Growth Index

<table>
<thead>
<tr>
<th>Models</th>
</tr>
</thead>
<tbody>
<tr>
<td>Early-Exit</td>
</tr>
<tr>
<td>Transitional</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Maintenance</th>
</tr>
</thead>
<tbody>
<tr>
<td>ESL</td>
</tr>
</tbody>
</table>

Types of Students

- Monolingual
- Dominant in Native Lang.
- Bilingual Lang. other than English

Figure 2. Program Models/Types of Student

An additional measure of student type is socio-economic status. Eighty-four percent of the project students in this study were classified as being from a low socio-economic group. This group, according to a United States General Accounting Office Report (1987), tends to show more efficient learning of academic skills in bilingual programs, while middle class children tend to do well using any language. This study indicates similar findings with this type of student, but also reveals that low socioeconomic monolingual students may do as well in an ESL program. Based on the preceding findings, the model...
used to instruct LEP, low socioeconomic group students, might have a bearing on their achievement level.

Do Different Models Make a Difference With Different Types of Programs?

The Title VII evaluation reports used in this study were generally much improved over what had been reported in past research. One area that continues to be somewhat vague is in the description of program types. Many reports generalized, giving the impression of self-containment in the secondary programs, with pull-out types in the elementary. In some instances this evidence was contradictory, so this researcher concluded that the project used a combination of program types. When referring to Figure 3, the mean result indicates equal success with each of two program types, pull-out and a combination of pull-out and self-contained. Although these types are clearly successful in helping students keep up with the mainstream, the self-contained classroom appears to be a factor in catching up with the mainstream in all models studied. This finding corroborates the recent research of Hakuta and Gould (1987).
Relative Growth Index

<table>
<thead>
<tr>
<th>Models</th>
<th>160</th>
<th>150</th>
<th>140</th>
<th>130</th>
<th>120</th>
<th>110</th>
<th>100</th>
<th>90</th>
<th>80</th>
<th>70</th>
<th>60</th>
</tr>
</thead>
<tbody>
<tr>
<td>Early-Exit</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transitional</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maintenance</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ESL</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Program Types

- Self-contained
- Pull-out
- Combination

Figure 3. Program Models/Types of Program

Do Different Models Make a Difference in Different Settings?

The majority, 74%, of the Title VII projects included in this study are in urban settings. The rural projects are primarily represented by native American students and comprise the minority. Within these two settings there is a clear preference for one model over another and a likewise obvious success rate. Figure 4 shows the results. The maintenance model was used in all of the rural projects and indicates substantial success in catching up with non-project student achievement. Urban projects overwhelmingly chose the ESL model for their bilingual programs and achieved results in keeping with the mainstream, but did not achieve the success rate of rural,
maintenance projects. Once again, as has been seen throughout the literature, use of the native language seems to enhance academic achievement of LEP students. The current SRA Technologies four-year longitudinal study (Crawford, 1986) found that students in bilingual programs performed better on tests in mathematics, reading, and language arts if the program included native-language components. These results were contrary to the expectations of the researchers. The findings of this study would be more conclusive had there been projects representing each model within each setting, but urban projects greatly outnumber rural projects in the Midwest.

![Relative Growth Index](chart.png)

**Program Settings**

<table>
<thead>
<tr>
<th>Relative Growth Index</th>
<th>Models</th>
</tr>
</thead>
<tbody>
<tr>
<td>160</td>
<td>Early-Exit</td>
</tr>
<tr>
<td>150</td>
<td>Transitional</td>
</tr>
<tr>
<td>140</td>
<td></td>
</tr>
<tr>
<td>130</td>
<td></td>
</tr>
<tr>
<td>120</td>
<td></td>
</tr>
<tr>
<td>110</td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>Maintenance</td>
</tr>
<tr>
<td>90</td>
<td>ESL</td>
</tr>
<tr>
<td>80</td>
<td></td>
</tr>
<tr>
<td>70</td>
<td></td>
</tr>
<tr>
<td>60</td>
<td></td>
</tr>
</tbody>
</table>

**Figure 4. Program Models/Settings**

This study examined the differential effectiveness of bilingual education models in relation to students,
program types, and settings. Data were obtained to assess the effectiveness of bilingual education through academic achievement. Study results indicated that, generally, regardless of student characteristics, program types and/or settings, bilingual education has a positive effect on program participants as measured by the gap-reduction design.

The sample of project means used in this study was derived from Title VII evaluation reports. The reports represent a wide range of educational conditions, student characteristics, teacher characteristics, etc. The basic question of interest in this study was in determining whether such differences were likely, overall, to have a differential effect on achievement. That is, is there an educational basis for selecting an optimum instructional model? This was not found to be the case. Based on the results of this study, which are summarized in Figure 5, there is no reason to assume that any one model, ESL, transitional or maintenance, is better on average than any other model. Analyses of the models in conjunction with other variables suggest that one may be favored over another depending on certain community factors.
Figure 5. A comprehensive View of Academic Achievement Project by Project
Limitations and Implications for Further Research

This study, as common with all meta-analysis, was dependent on other research, Title VII evaluations, for its data. These evaluations have often been criticized, in past meta-analytic research (Okada et al., 1982; Willig, 1985), for their lack of precision and tendency to be incomplete. Recommendations have generally revealed a need for improving evaluations and ultimately the efficacy of bilingual education. The evaluations used in this study were similarly insufficient, but to a much lesser extent. While Okada et al. (1982) was able to include only 12% of the evaluations reviewed, this study realized the inclusion of 40%. The quality of evaluations remains a serious obstacle to precise research, principally for what is not contained in the reports, but this study points to an improvement in this area.

This study was further limited in its findings because of the very nature of Title VII governmental rules and regulations. Although many cities, towns, etc., are carrying out programs of bilingual education, they may or may not qualify for Title VII funding because of an inadequate proposal, insufficient number of students, proximity to a larger program, etc. For example, the state of Wisconsin operates exemplary programs of bilingual education with state and local resources, not opting for
federal grants. Consequently, Wisconsin is not included in this study. Additionally, because the federal government, in the past, has not been specific as to program model used, all are not represented here. In fact, the early-exit transitional model, which research indicates is widely used, was present in only one project represented in this study. This particular project had the highest relative growth index on all variables, but posed a limitation because of the small number of projects. While data from this study support the question of bilingual effectiveness, further verification of the question and research on the topic is needed.

One variable of practical importance which should be investigated is the program type, self-contained and pull-out, in the elementary and secondary projects. This area is very confusing as many projects use a combination and make it impossible to discern what is being used and where the results of this study indicate that the self-contained classroom is highly successful, these results cannot be extrapolated because of the confusion resulting from the method of reporting.

An even more important area which needs investigation is the extent to which Title VII project students are comparable to the national norm group with whom they are compared. This has long presented a problem for bilingual education because law demands that all LEP students
receive services, making a similar control group an impossibility. Current Title VII regulations allow for appropriate non-project comparison groups under the assumption that Title VII project students are similar to students in the norm group who scored at the same percentile level. Because of student population differences between Title VII student subpopulations and the cross-section of students comprising the national norm, Title VII students may progress at either a faster or slower rate than expectations derived from the national norm. How significant this difference is needs to be investigated.
APPENDIX A

Conceptual Framework
## Bilingual Program Outcomes

### Instructional Components

1. Culture/History
2. ESL/English Language Arts
3. Native Language Arts
4. English Reading
5. Native Language Reading
6. English Language Math
7. Native Language Math
8. English Language Social Science/Science
9. Native Language Social Science/Science
10. English Language Misc. Courses
11. Native Language Misc. Courses

### Non-Instructional Components

1. Development/Acquisition of Curriculum Instructional Materials
2. Community Parent Participation
3. Staff Training
4. Improvement of Program Management
5. Improvement of Evaluations

### Affective Components

1. Self-Concept
2. Attitude Toward School

The models and outcomes indicated by an X were studied for particular types of student, in different settings, and in different types of program.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
APPENDIX B

Implementation of the Gap-Reduction Design and Relative Growth Index
Steps for the Implementation of the
Gap-Reduction Design and
Relative Growth Index

Step 1. Convert each project and comparison group student's raw pretest and posttest scores to scale scores using the correct conversion table for the form and level of the test you used.

Step 2. Compute the means of all scores and standard deviations of the comparison group.

<table>
<thead>
<tr>
<th></th>
<th>Project Group</th>
<th>Comparison Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>pretest mean</td>
<td>55.34</td>
<td>61.63</td>
</tr>
<tr>
<td>pretest std. dev.</td>
<td>n/a</td>
<td>10.48</td>
</tr>
<tr>
<td>posttest mean</td>
<td>65.88</td>
<td>70.63</td>
</tr>
<tr>
<td>posttest std. dev.</td>
<td>n/a</td>
<td>9.50</td>
</tr>
</tbody>
</table>

Step 3. Subtract the project group's mean pretest score from the comparison group's mean pretest score. Divide the difference by the comparison group's pretest standard deviation and label the result the pretest gap. (61.63 - 55.34) / 10.48 = .60 pretest gap

Step 4. Follow same procedure for posttest gap.
(70.63 - 65.88) / 9.50 = .50 posttest gap

Step 5. Subtract posttest gap from pretest gap and label the difference the gap reduction. .60 - .50 = .10 gap reduction

Step 6. Subtract the comparison group's mean pretest score from its mean posttest score and label the difference the comparison group's unstandardized growth estimate.
70.63 - 61.63 = 9.00 comparison group's unstandardized growth estimate
Step 7. Using the comparison group's pre- and posttest standard deviations, calculate the following value:

\[
\sqrt{\frac{(10.48)^2 + (9.50)^2}{2}} = \sqrt{100.04} = 10.00 \text{ comparison group's pooled standard deviation}
\]

Step 8. Divide the comparison group's unstandardized growth estimate by the comparison group's pooled standard deviation. Label this value the comparison group's standardized growth estimate.

9.00 / 10.00 = .90 comparison group's standardized growth estimate

Step 9. Add the gap reduction to the comparison group's standardized growth estimate. Label this sum the project group's standardized growth estimate.

.90 + .10 = 1.00 project group's standardized growth estimate.

Step 10. Divide the project group's standardized growth estimate by the comparison group's standardized growth estimate. Multiply the result by 100 to convert it to a percent and label it the Relative Growth Index (RGI).

(1.00 / .90) 100 = 111% Relative Growth Index
BIBLIOGRAPHY


Crawford, J. (1986, April 23). Immersion method is faring poorly in bilingual study. Education Week, 1, 10.


Johnson, D. W., Maruyama, G., Johnson, R. T., Nelson, D.,


Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.


Riverside Publishing Co. (1983). Iowa tests of basic


Saville-Troike, R. (1977, Sept./Oct.) The view from the center: Warning-ESLC (traditional) may be hazardous to children. The Linguistic Reporter.


Walton, S. (1984, February 8). Research and the quest
for 'effective' bilingual methods. Education Week, pp. 18-20.


Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.