Educational efforts center around helping students succeed in school. Defining success, however, is a difficult and elusive task. Indicators of success and how to measure them are at the center of much controversy. A review of the research on bilingual education in the United States over the past 30 years reveals different trends in the underlying notion of success. The debate on the effectiveness of bilingual education has dominated the research agenda in the United States since the 1970s. Most studies focus on the ability of bilingual programs to enhance students’ performance in English academic areas, especially language and mathematics. By the late 1980s bilingual education researchers shifted their focus to identify characteristics of effective bilingual programs, as was occurring in the effective schools movement. Presently, a trend toward in-depth case studies provides educators with profiles of bilingual programs in context.

This paper proposes a comprehensive framework for defining success in bilingual education. It also shows how research over the years has focused on certain components of the framework and how some recent trends are consistent with this broader definition.
Defining Success

Success is measured by outcomes, depends on quality of inputs, and is challenged by contextual factors (see Figure 1). A successful bilingual program develops students’ language and literacy proficiency, leads them in successful academic achievement, and nurtures sociocultural integration. The level of language and literacy development for each language will depend on the goals of the program. Academic achievement has been defined and measured in different ways. Most schools simply monitor students’ performance in standardized tests; others prefer to define it based on performance of real life tasks, reflecting instructional goals of the schools (Sizer, 1992). Sociocultural integration is the ability to function in the larger society as well as in the heritage community (Brisk, 1998). Programs rarely measure the latter outcome, nor is it a variable in their definition of success. Research on cultural affiliation has demonstrated the effects of a balanced attitude toward both cultures (Gardner and Lambert, 1972; Taylor, 1987). Given the wider role that schools are taking in the community, outcomes can also include general impact of the program and school on the families.

Fairness and accuracy of assessment procedures play an important role in defining success. In the case of bilingual programs, language and culture greatly influence assessment. Students’ successful performance can be measured directly by assessing language, literacy, and content knowledge or indirectly by such things as promotion, completion of high school, school attendance, successful participation in contests, parent satisfaction, and by long-term effects of education on helping students become productive members of society. Success of students’ performance depends on the quality of the bilingual program, including curriculum, instruction, resources, assessment practices, and preparation of personnel. Different contextual factors, outside the control of the program, determine the challenge in educating specific students. These factors are related to the individual students, their peers, their families, the school where the program is housed, and

FIGURE 1: Framework for Defining Success

<table>
<thead>
<tr>
<th>DEPENDS ON</th>
<th>MEASURED BY</th>
<th>INFLUENCED BY</th>
</tr>
</thead>
<tbody>
<tr>
<td>Program Characteristics</td>
<td>Students’ Performance</td>
<td>Contextual Factors</td>
</tr>
<tr>
<td>• teacher preparation</td>
<td>• language and literacy development</td>
<td>• students’ characteristics</td>
</tr>
<tr>
<td>• curriculum</td>
<td>• academic achievement</td>
<td>• peer pressures</td>
</tr>
<tr>
<td>• materials</td>
<td>• sociocultural integration</td>
<td>• family’s role</td>
</tr>
<tr>
<td>• instruction</td>
<td></td>
<td>• school characteristics</td>
</tr>
<tr>
<td>• assessment</td>
<td></td>
<td>• situational factors</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(historical perspective)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>INPUT</th>
<th>OUTCOME</th>
<th>SUPPORT</th>
<th>CHALLENGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>QUALITY</td>
<td>SUCCESS</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
the socio-historical context, including linguistic, cultural, political, economic, and social factors (Brisk, 1998). The amount of effort a bilingual program makes to successfully educate its students will increase relative to the diminished support of these external factors.

BILINGUAL EDUCATION DEBATE

The debate on bilingual education produced scores of studies, beginning in the 1970s, that mainly measured students’ performance on standardized tests, particularly in English. Most of the studies researched elementary Spanish-English bilingual education. The studies either compared bilingual education with English-only education, or evaluated particular types of bilingual education. These studies included little or no description of program characteristics.

Three studies (AIR, Baker and de Kanter, Rossell and Baker) of existing evaluations or individual studies concluded that bilingual education did not have a significant impact. The controversial AIR report (Danoff et al., 1978) reviewed evaluations of Title VII-funded Spanish-English bilingual programs. The purpose of this review was to establish the effects that federal funding of bilingual programs had on student achievement, particularly English reading, language, and math. Information was also gathered on Spanish language, reading, math and attitudes. The AIR study carried out some surveys on students’ and teacher’s backgrounds as well as instruction and curricular characteristics. The study concluded that students in bilingual programs did not achieve better than students not attending bilingual programs. The lack of strength in the control groups lead to questions about the validity of the results. Baker’s and de Kanter’s (1981, 1983) reviews focused on the ability of transitional bilingual programs to promote English language development. They concluded that the evidence from existing studies was weak. Rossell and Baker (1996) concluded that transitional bilingual education is no better than structure immersion programs in developing English language skills.¹

Two major meta-studies, Willig and Greene, were done to rebut the findings of opponents of bilingual education.² Willig (1985) analyzed 23 studies included in the Baker and de Kanter (1981) review of the literature. Willig’s report compared the results from various types of programs such as bilingual, ESL, two-way with submersion models. Student outcomes tested in those studies included English language, English language and math, as well as self-concept. Other studies reviewed evaluated outcomes with respect to English and Spanish language, as well as reading and writing, English math, attitudes/self-concept, and cognitive functioning. One of the studies included English aural comprehension, reading, math and French listening, language, math, and social studies.

Greene (1997) offered a meta-analysis of eleven studies included in Rossell and Baker (1996) comparing the

¹ The problem with this study lies in the authors’ misidentification of Canadian Immersion programs as structured immersion, rendering the conclusions invalid. (See critique in Brisk, 1998 and Cummins, 1999).
² For the sake of conciseness, I have chosen to report on published research that covered a number of individual studies. My report does not represent an exhaustive list.
effectiveness of various types of bilingual programs with English-only programs. The student outcomes reported were tests in English, tests in Spanish, and tests in English reading. Only four of the eleven studies measured native language. No program descriptions were included.

Other research supporting bilingual education reviews existing studies or evaluations. Zappert and Cruz (1977) surveyed twelve studies to look for evidence of the effectiveness of bilingual programs when compared to monolingual programs. They only reported on students’ outcomes on measures similar to those included in Willig. There was no mention of whether the studies described the programs themselves. Zappert and Cruz found that performance of students in bilingual education programs, except for one in the Philippines, was equal or superior to those of students not attending bilingual programs.

Troike (1978) analyzed twelve programs, mostly funded by Title VII, serving Spanish, Chinese, French, and Navajo populations. Outcomes evaluated include English language, reading readiness, reading, writing, general achievement, math, and social studies. Only programs designed for Spanish bilinguals measured achievement in the native language. Some programs measured attendance. This analysis supported the effectiveness of bilingual programs.

Some studies focused on defending specific types of programs, such as maintenance and two-way bilingual education. Medina and Escamilla (1992a) measured the effect maintenance bilingual programs had on English language proficiency of Spanish-speakers, while Medina (1991) also included Spanish language and math achievement. Medina and Escamilla (1992b) studied oral achievement in English and Spanish of students in maintenance programs. They compared the English results to those of Vietnamese speakers in transitional programs. Studies of two-way bilingual education reported student outcomes with respect performance in language, math, and other content areas in both languages (Lindholm, 1991), as well as attitudes toward bilingualism (Cazabon, Nicolaidis, and Lambert, 1998).

All of these studies and meta-studies compared students in different educational models, giving little detail to the school context. Several of the meta-studies and critiques commented on the fact that student backgrounds and program characteristics can influence student outcomes, but they did not include specifics (Gray, 1977; Troike, 1978; Willig, 1985). Bilingual maintenance and two-way program studies included some explanation of program characteristics, such as language distribution, in the curriculum.

Only two of these major studies analyzed program characteristics and contextual factors. Ramirez (1992) compared the effectiveness of long-term transitional bilingual programs, short-term transitional bilingual programs, and English-only programs. This study included, in addition to student outcomes in English reading and math, some evidence of instructional strategies, staff qualifications, and parental involvement. Thomas and Collier (1997) reported on English reading and math achievement with respect to different program models. They synthesized the characteristics common to good programs, such as quality of instruction and curriculum. They also described some necessary school contextual factors gleaned from focus interviews with staff in various sites.

3 In most cases, evaluations included only some of the measures listed.
TABLE 1: Studies on Effectiveness of Bilingual Education

<table>
<thead>
<tr>
<th>Students’ Outcomes</th>
<th>Program Characteristics</th>
<th>Contextual Factors</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>WHAT</td>
<td>HOW</td>
<td>WHAT</td>
<td>HOW</td>
</tr>
<tr>
<td>• Predominantly English reading, language arts, and math.</td>
<td>• Standardized tests</td>
<td>• Model type (TBE, Structured Immersion, etc.)</td>
<td>• Parent involvement; some school characteristics</td>
</tr>
<tr>
<td>• Some content area and native language achievement¹</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• English reading and math²</td>
<td>• Standardized tests</td>
<td>• Models + some program characteristics</td>
<td></td>
</tr>
<tr>
<td>• English oral and reading.</td>
<td>• Standardized tests</td>
<td>• Language use in the curriculum</td>
<td></td>
</tr>
<tr>
<td>• Native language and math³</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Language and content area achievement in English and Spanish and attitude toward bilingualism⁴</td>
<td>• Standardized tests + attitude survey</td>
<td>• Language use in the curriculum</td>
<td>• Effect on students</td>
</tr>
</tbody>
</table>

⁴ Cazabon et al. (1998), Lindholm (1991)

**Effective Bilingual Education**

In embracing the movement toward accountability, some educators interested in bilingual education have adopted the effective schools model of defining success (for a synthesis of these studies see August and Hakuta, 1997; Brisk, 1998). Although initially students’ outcomes were the basis for defining success, studies focused on what characteristics have made these schools successful. In the bilingual education studies influenced by this trend, students’ outcomes were usually vaguely reported. Identification of the programs as effective was done mostly through experts’ opinions, and by corroboration with extensive qualitative research done on school sites. Program quality and school factors were described in detail through extensive and thorough site research (Berman et al., 1995; Carter and Chatfield, 1986; Freeman, 1998; Lucas, Henze, and Donato, 1990; Mace-Matluck et al., 1989; Mace-Matluck et al., 1998). Other studies focused only on
classroom practices (Fisher et al., 1983; Garcia, 1990; Moll, 1988; Gersten, 1996; Pease-Alvarez et al., 1991; Tikunoff, 1983; Wong Fillmore et al., 1985).

A number of studies were designed to test the effectiveness of particular teaching strategies with bilingual students. The languages of instruction depended on the particular setting. Studies of classroom characteristics and specific instructional practices described classroom context in detail, but provided little information on the bilingual programs as a whole. These studies assessed student outcomes in one or both languages only with respect to skills that the particular instructional innovation addressed. A substantial number focused on different approaches to developing literacy (Calderon et al., 1996; Dianda and Flaherty, 1994; Escamilla, 1994; Goldenberg and Gallimore, 1991; Hernandez, 1991; Muñiz-Swicegood, 1994; Saunders et al., 1998; Slavin and Madden, 1994; Tharp, 1982).

A more limited number of studies highlighted strategies for teaching content areas. Cohen (1984) studied the effects of Finding Out/Descubrimiento, a bilingual approach to teaching math and science; Rosebery et al. (1992) focused on teaching of science; and Henderson and Landesman (1992) looked at the effects of thematic instruction on various content areas.

### TABLE 2: Studies on Effectiveness of Bilingual Education

<table>
<thead>
<tr>
<th>STUDENTS’ OUTCOMES</th>
<th>PROGRAM CHARACTERISTICS</th>
<th>CONTEXTUAL FACTORS</th>
<th>PURPOSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>WHAT</td>
<td>HOW</td>
<td>WHAT</td>
<td>HOW</td>
</tr>
<tr>
<td>• Prestige</td>
<td>• Nomination from experts or unspecified</td>
<td>• Curriculum, classroom characteristics&lt;sup&gt;1&lt;/sup&gt;</td>
<td>• School characteristics; parent involvement</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Classroom characteristics&lt;sup&gt;2&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>• Reading or content area</td>
<td>• Standardized and authentic assessments</td>
<td>• Implementation of instructional intervention&lt;sup&gt;3&lt;/sup&gt;</td>
<td>• Observations</td>
</tr>
</tbody>
</table>

<sup>1</sup> Berman et al. (1995); Carter and Chatfield (1986); Freeman (1998); Lucas et al. (1990); Mace-Matluck et al. (1989); Mace-Matluck et al. (1998); Tikunoff et al. (1991).

<sup>2</sup> Fisher et al. (1983); Garcia (1990); Moll (1988); Gersten (1996); Pease-Alvarez et al. (1991); Tikunoff (1983); Wong Fillmore et al. (1985).

<sup>3</sup> Calderon et al. (1996); Cohen (1984); Dianda and Flaherty (1994); Escamilla (1994); Goldenberg and Gallimore (1991); Henderson and Landesman (1992); Hernandez, 1991; Muñiz-Swicegood (1994); Rosebery et al. (1992); Saunders et al. (1998); Slavin and Madden (1994); Tharp (1982).
A recent research trend considers student outcomes and educational program characteristics. Schools can only be successful if their students are progressing. To understand the nature of this success, researchers documented in detail the characteristics of the educational program offered to students as well as the social and historical context of these programs. The purpose of these studies varies. Dentler and Hafner (1997) focused on districts that successfully adjusted to the influx of sizable numbers of immigrant students. They provided student outcome data with respect to reading and math as measured in English. Curriculum, instructional practices, and quality of personnel are described in detail. Dentler and Hafner investigated the history and characteristics of the district, its educational institutions, and other services offered to immigrant families through extensive field research. Christian et al (1997) studied three two-way programs “to contribute to a greater understanding of how two-way immersion works” (p. 4). They measured student outcomes with respect to language and academic achievement in English and Spanish. They also described in detail the districts where the schools are located, the development of the programs, characteristics of the programs, and instructional practices. They provided qualitative data on students’ language development and included a number of outcome measures both in English and Spanish for language and content areas. The research was carried out over a two-year period. Researchers visited schools frequently, interviewed school staff, observed class rooms, and obtained results of the various tests administered in both languages. By contrasting the three programs, the study provided comparisons of methods for approaching two-way education.

### TABLE 3: Case Studies

<table>
<thead>
<tr>
<th>Students’ Outcomes</th>
<th>Program Characteristics</th>
<th>Contextual Factors</th>
<th>PURPOSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>WHAT</td>
<td>HOW</td>
<td>WHAT</td>
<td>HOW</td>
</tr>
<tr>
<td>• English reading and math(^1)</td>
<td>• Standardized tests</td>
<td>• Teacher quality curriculum, instructional practices</td>
<td>• School and district characteristics, parent involvement</td>
</tr>
<tr>
<td>• Language/literacy in English and heritage language</td>
<td>• SOLOM</td>
<td>• SOPR</td>
<td>• LASS</td>
</tr>
<tr>
<td>• Academic achievement(^2)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

\(^1\) Dentler and Hafner (1997)


---

4 They focused specifically on Spanish/English programs because these are the most popular at present.
CONCLUSION

Research that aims at judging the effectiveness of bilingual education centers on the language of instruction and models. This research continues to support the notion that in the model lies the solution, disregarding the influence of program characteristics and contextual factors on students’ performance. In contrast to the focus on language and models emerged research that looked at schools, classrooms, and specific instructional practices in detail. These studies provide rich information to guide schools to adopt successful practices. Good practices are likely to promote learning, but they do not necessarily guarantee student performance.

The United States is a large and extremely varied country. The goal of bilingual research should be to understand how different communities employ different paths in the implementation of successful bilingual education. A given program is successful by the attainment of its students, is challenged by the amount of adverse factors it must conquer, and is exemplary by the quality of its practices. Case-study research reflected in the latest trends has the potential of covering all areas of the proposed definition of success—student outcomes, educational practices, and social context. The accumulation of such case studies presents a rich and robust picture of bilingual education that can benefit practice and influence policy (Cummins, 1999).
REFERENCES


The Northeast and Islands Regional Educational Laboratory

a program of The Education Alliance at Brown University

The LAB, a program of The Education Alliance at Brown University, is one of ten educational laboratories funded by the U.S. Department of Education’s Office of Educational Research and Improvement. Our goals are to improve teaching and learning, advance school improvement, build capacity for reform, and develop strategic alliances with key members of the region’s education and policy making community.

The LAB develops educational products and services for school administrators, policymakers, teachers, and parents in New England, New York, Puerto Rico, and the Virgin Islands. Central to our efforts is a commitment to equity and excellence.

Information about LAB programs and services is available by contacting:

LAB at Brown University
The Education Alliance
222 Richmond Street, Suite 300
Providence, RI 02903-4226

Phone: 800/521-9550
E-mail: info@lab.brown.edu
Web: www.lab.brown.edu
Fax: 401/421-7650

This publication is based on work supported by the Office of Educational Research and Improvement (OERI), U.S. Department of Education, under Contract Number RJ96006401. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of OERI, the U.S. Department of Education, or any other agency of the U.S. Government.

Copyright © 2000  Brown University. All rights reserved.
The Northeast and Islands Regional Educational Laboratory

a program of The Education Alliance at Brown University

Adeline Becker
Executive Director, The Education Alliance

Phil Zarleno
Executive Director, The LAB at Brown University

Vincent Ferrandino
Chair, LAB Board of Governors

Marjorie Medd
Vice Chair, LAB Board of Governors

BOARD MEMBERS

J. Duke Albanese Richard Mills
Barbara Bailey Thong Phamduy
Pamela Berry Daria Plummer
Paul Crowley Olga Lucia Sallaway
David Driscoll Theodore Sergi
Katharine Eneguess David Sherman
Victor Fajardo Ruby Simmonds
Charlotte K. Frank Jeanette Smith
Harold Levy Jill Tarule
Edward McElroy Elizabeth Twomey
Peter McWalters David Wolk