An Urban Myth

The “Poor Quality” of Bilingual Education Research

You’ve heard the charge leveled many times before: The “problem” with bilingual education is that there are no high-quality studies to support it. Friends and foes alike have decried the state of research on bilingual programs. The “poor quality” argument has appeared repeatedly in recent years, especially during the contentious fights over English-only initiatives in California, Arizona, and Massachusetts.

Yet, on close inspection, this accusation has little evidence to support it. Indeed, one could make a strong case for the opposite conclusion, that the field of bilingual education has more high quality research than most other areas of education. Certainly it has no less.

The Ten Percent Confusion

Let’s begin with the weakest of all the arguments behind the critique of bilingual education research. It is often claimed that a “disappointing percentage of studies… [are found] to be methodologically adequate” (August & Hakuta, 1997, p. 146). At first glance, this would appear to be true. The author of one prominent review of the research found that only about 10 percent of all studies done in bilingual education during the 1970s and 1980s were of high quality (Lam, 1992). Later reviews concluded that the percentage was a bit higher, though still quite low.

But suppose I were to tell you that “only” 10 percent of a thousand people were coming to dinner at your house this evening. It’s a small percentage but, most would agree, a rather large number for a dinner party. Now consider that the 10 percent figure is typically based on reviews of hundreds of studies and in some cases more than 1,400. Even two of the severest critics of bilingual education, Christine

Jeff McQuillan
Center for Educational Development
Russell and Keith Baker, found more than 70 good-quality studies to include in their review. The percentage statistic is meaningless unless we know how many studies a review started with. What counts in the end is the absolute number of studies, not the “percentage acceptable.”

When we compare these numbers with those in early reading instruction, for example, bilingual education research fares quite well. The most prominent review in early literacy (Stahl & Miller, 1989) found that 52 studies—of an unspecified number examined—were of high enough quality to be included. Yet there were few complaints about the “poor quality” of early reading instruction when this review was published. Indeed, most instructional interventions and techniques have far fewer studies to support them than bilingual education—contrary to the implication of those who espouse the “only 10 percent” analysis.

**What Counts as a Study?**

Even the 10 percent figure given above is very misleading when we consider just what gets counted as a “study” in research reviews of bilingual education. Because many bilingual programs were funded through federal Title VII funds (from 1968 to 2002), they were required to submit a written evaluation report. Yet, as anyone who has worked on a federal grant evaluation knows, these efforts to collect and analyze data are rarely done with the aim of being published in a scientific journal or being presented at a conference.

Evaluations are not “studies” at all in that sense. Rather, these are reports by school personnel or outside evaluators who often have inadequate training for such an effort. Yet they are often classified as “studies” of the effectiveness of bilingual education.

In addition, evaluation reports undergo no peer review by other researchers to ensure quality. They are simply filed with the appropriate state or federal office. It is therefore not surprising that, subjected to the rigors of university-level examination, they fall short of the mark. In Rossell and Baker’s 1996 review of 300 bilingual education “studies,” 89 percent were unpublished evaluation reports of precisely this nature. It’s no wonder that these researchers found only 24 percent of the studies methodologically adequate. But it’s also no reflection on the true quality of bilingual education studies, properly understood.

When we compare bilingual education to other federally funded programs that have similar evaluation report requirements, we find that it is no less “adequate” than other areas of education. For example, the 1997 General Accounting Office review of 200 Head Start evaluations determined that approximately 10 percent (22) met their methodological criteria—higher percentage of “acceptable” studies than those in bilingual education? To answer this question, I looked at two major journals, the *Review of Educational Research* and *Psychological Bulletin*, and selected 26 empirical reviews at random. My goal was to compile an average of how many studies in other fields of education are typically rejected for low methodological quality.

It was impossible, however, to compute an average percentage of “low quality” studies. Only one of the 26 articles actually reported the number of studies that appeared to be rejected for inadequate methodology (as opposed to other possible exclusion criteria, such as not examining the constructs or population of interest).

That single study (Greenwald, Hedges, & Laine, 1996, on school finance) found that 18 percent of the 175 articles and books initially reviewed met all of their criteria for inclusion, a figure not much different from the average percentage of empiricism in bilingual education reviews through 1996 (15 percent, with a range of 5 percent to 44 percent).

Moreover, in the average number of studies included, other research reviews were very similar to reviews of bilingual education. For education reviews, an average of 55 studies were used to make a determination as to the success or failure of a given approach or research hypothesis. For psychology reviews, an average of 90 studies were used.

The truth is that bilingual education research is no worse—and, in many cases, considerably better—than research in other branches of education.

*continued on page 30*
An Urban Myth

These comparisons are quite favorable to bilingual education research—remember that Rossell and Baker found 72 high quality studies—especially when one considers that research design difficulties, such as the establishment of a comparable control group, are much less severe in other areas of education and psychology than they are in bilingual program evaluation.

Always Room for Improvement

Clearly, we need more high-quality studies of bilingual education, as August and Hakuta (1997) and others have argued. But this conclusion is quite different from judging the current research in bilingual education to be inferior in quality to other educational research. There is no logical or empirical basis for such harsh assessments.

The “percentage acceptable” method used by other reviewers has little acceptance in either education or psychology as a measure of quality. In any case, it is an unstable product of shifting acceptability criteria, with little regard for the absolute number of studies available. This is, in other words, the sort of crude and context-free “single statistic” that research methodologists have warned against as a poor way to make a reasoned argument. The truth is that bilingual education research is no worse—and, in many cases, considerably better—than research in other branches of education.

For a more detailed presentation of these ideas, see McQuillan, J. (2005). The “poor quality” of bilingual education research: An examination of the logic and metric of judgment. International Journal of Foreign Language Teaching, 1(2), 1-5.

References


