

Rossell, Christine H. und Julia Kuder (2005): Meta-Murky: A
Rebuttal to Recent Meta-Analyses of Bilingual Education.

In: Arbeitsstelle Interkulturelle Konflikte und gesellschaftliche Integration (AKI) (ed.).
The Effectiveness of Bilingual School Programs for Immigrant Children.
Berlin: Wissenschaftszentrum Berlin für Sozialforschung,
Discussion Paper SP IV 2005-601: 43-76.

Online-Publication of Appendices 1a, 1b, 2, 3, 4, 5, 6, 7

Appendix 1a

Effects* of TBE on Second Language Reading, Language, and Math
Compared to Other Instructional Techniques
as Found in Methodologically Acceptable Studies

	READING (or ORAL **)	LANGUAGE	MATH
<u>TBE v. Submersion</u>			
TBE Better	AIR (Corpus Christi), 1975b; Bacon, Kidd & Seaberg, 1982; Burkheimer et al., 1989; Campeau et al., 1975; Carsrud & Curtis, 1980; Covey, 1973; Kaufman, 1968; Legaretta, 1979; McConnell, 1980; Morgan, 1971; Olesini, 1971; Plante, 1976; Zirkel, 1972	Burkheimer et al., 1989	Cohen, 1975a; Bacon, Kidd & Seaberg, 1982; Burkheimer et al., 1989
	(N=13)	(N=1)	(N=3)
No Difference	AIR (Corpus Christi), 1975b; Alvarez, 1975; Ariza, 1988; Barclay, 1969**; Campeau, et al., 1975; Carsrud & Curtis, 1980; Ciriza, 1990a**; Cohen, 1975a; Cottrell, 1971; Huzar, 1973; Kaufman, 1968; Lampman, 1973; Legaretta, 1979; Maldonado, 1977; Matthews, 1979; McSpadden, 1979; 1980; Morgan, 1971; Plante, 1976; Powers, 1978; Prewitt-Diaz, 1979; Rothfarb, Ariza, & Urrutia, 1989; Stebbens et al., 1977; Skoczylas, 1972; Vasquez, 1990; de Weffer, 1972; Zirkel, 1972	Ariza, 1988; Ed. Op. Concepts, 1991b; Maldonado, 1977; Rothfarb, Ariza & Urrutia, 1989	Alvarez, 1975; Ariza, 1988; Bates, 1970; Carsrud & Curtis, 1980; Cohen, 1975a; Covey, 1973; Danoff et al., 1977; 1978 Ed. Op. Concepts, 1991b; Layden, 1972; Maldonado, 1977; McSpadden, 1979; 1980; Moore & Parr, 1978; Powers, 1978; Rothfarb, Ariza, & Urrutia, 1989; Stebbins, et al., 1977; Vasquez, 1990; de Weffer, 1972
	(N=27)	(N=4)	(N=19)

TBE Worse	Bates, 1970; Burkheimer et al., 1989; Cohen, Fathman, & Merino, 1976; Curiel, 1979; Curiel, Stenning, & Cooper, 1980; Danoff et al., 1977; 1978; Ed. Op. Concepts, 1991a; 1991b; El Paso, 1987; 1990; 1992; Layden, 1972; McSpadden, 1980; Melendez, 1980; Moore and Parr, 1978; Stern, 1975; Teschner, 1990; Valladolid, 1991; Webb, Clerc & Gavito, 1987	Burkheimer et al., 1989; Curiel, 1979; Curiel, Stenning, & Cooper, 1980; Ed. Op. Concepts, 1991a; El Paso, 1987; 1990; 1992; Teschner, 1990; Valladolid, 1991	Burkheimer et al., 1989; Cohen, Fathman, & McSpadden, 1980; Ed. Op. Concepts, 1991a; El Paso, 1987; 1990; 1992; Maldonado, 1977; Merino, 1976; Skoczylas, 1972; Stern, 1975; Teschner, 1990; Valladolid, 1991
	(N=20)	(N=9)	(N=12)

TBE v. ESL

TBE Better			Ames & Bicks, 1978
	(N=0)	(N=0)	(N=1)
No Difference	Ames & Bicks, 1978; Balasubramonium et al., 1973; Lum, 1971; Rossell, 1990; Yap, Enoki & Ishitani, 1988	Rossell, 1990; Yap, Enoki & Ishitani, 1988	Rossell, 1990; Yap, Enoki, & Ishitani, 1988
	(N=5)	(N=2)	(N=2)
TBE Worse	Lum, 1971; Rossell, 1990	Rossell, 1990	Rossell, 1990
	(N=2)	(N=1)	(N=1)

TBE v. Structured Immersion

No Difference	Ramirez et al., 1991; Ramos et al., 1967	Ramiriz et al., 1991	Barik, Swain, & Nwanunobi, 1977; Barik & Swain, 1975; Lambert & Tucker, 1972; Ramiriz et al., 1991; Ramos et al., 1967
	(N=2)	(N=1)	(N=5)
TBE Worse	Barik, Swain, & Nwanunobi, 1977; Barik & Swain, 1978; Bruck, Lambert, & Tucker, 1977; Day & Shapson, 1988;		Genessee & Lambert, 1983; Genessee et al., 1989; Gersten, 1985

Genessee & Lambert, 1983;
Genessee, Lambert & Tucker,
1977; Genessee et al., 1989;
Gersten, 1985; Malherbe,
1946; Pena-Hughes &
Solis, 1980

(N=10)

(N=0)

(N=3)

**Immersion v.
ESL**

Immersion Better Barik & Swain, 1975;
Becker and Gersten, 1982
Lambert & Tucker, 1972

(N=3)

(N=0)

(N=0)

TBE v. Maint. BE

TBE Better Medina & Escamilla, 1992**

(N=1)

(N=0)

(N=0)

*Studies are listed in more than one category if there were different effects for different grades or cohorts.

**Oral English achievement gains for preschool programs.

Appendix 1b

Revised Effects* of TBE on Second Language Reading, Language, and Math Compared to Other Instructional Techniques as Found in Methodologically Acceptable Studies

	READING (or ORAL**)	LANGUAGE	MATH
<p><u>TBE v. Submersion</u> TBE Better</p>	<p>AIR (Campeau, Corpus Christi), 1975b; Bacon, Kidd & Seaberg, 1982; Burkheimer et al., 1989; Campeau et al., 1975; Carsrud & Curtis, 1980; Covey, 1973; Kaufman, 1968; Legaretta, 1979; McConnell, 1980; Morgan, 1971; Olesini, 1971[†]; Plante, 1976; Zirkel, 1972</p> <p style="text-align: center;">(N=7)</p>	<p>Burkheimer et al., 1989</p> <p style="text-align: center;">(N=1)</p>	<p>Cohen, 1975a; Bacon, Kidd & Seaberg, 1982; Burkheimer et al., 1989</p> <p style="text-align: center;">(N=2)</p>
<p>No Difference</p>	<p>AIR (Campeau, Corpus Christi), 1975b; Alvarez, 1975; Ariza, 1988; Barclay, 1969**; Campeau, et al., 1975; Carsrud & Curtis, 1980; Ciriza, 1990a**; Cohen, 1975a; Cottrell, 1971; Huzar, 1973; Kaufman, 1968; Lampman, 1973; Legaretta, 1979; Maldonado, 1977; Matthews, 1979; McSpadden, 1979; 1980; Morgan, 1971; Plante, 1976; Powers, 1978; Prewitt-Diaz, 1979; Rothfarb, Ariza, & Urrutia, 1989; Stebbens et al., 1977; Skoczylas, 1972; Vasquez, 1990; de Weffer, 1972; Zirkel, 1972</p> <p style="text-align: center;">(N=18)</p>	<p>Ariza, 1988; Ed. Op. Concepts, 1991b; Maldonado, 1977; Rothfarb, Ariza & Urrutia, 1989</p> <p style="text-align: center;">(N=2)</p>	<p>Alvarez, 1975; Ariza, 1988; Bates, 1970; Carsrud & Curtis, 1980; Cohen, 1975a; D25; Danoff et al., 1977; 1978; Ed. Op. Concepts, 1991b; Layden, 1972; Maldonado, 1977; McSpadden, 1979; 1980; Moore & Parr, 1978; Powers, 1978; Rothfarb, Ariza, & Urrutia, 1989; Stebbins, et al., 1977; Vasquez, 1990; de Weffer, 1972</p> <p style="text-align: center;">(N=11)</p>

[†] Olesini, 1971 was accidentally inserted into Table 1 and Appendix 1 in Rossell and Baker. It was considered methodologically unacceptable in Baker and deKanter 1983b.

TBE Worse	Bates, 1970; Burkheimer et al., 1989; Cohen, Fathman, & Merino, 1976; Curiel, 1979; Curiel, Stenning, & Cooper, 1980; Danoff et al., 1977; 1978; Ed. Op. Concepts, 1991a; 1991b; El Paso, 1987; 1990; 1992; Layden, 1972; McSpadden, 1980; Melendez, 1980; Moore & Parr, 1978; Stern, 1975; Teschner, 1990; Valladolid, 1991; Webb, Clerc & Gavito, 1987	Burkheimer et al., 1989; Curiel, 1979; Curiel, Stenning, & Cooper, 1980; Ed. Op. Concepts, 1991a; El Paso, 1987; 1990; 1992; Teschner, 1990; Valladolid, 1991 <u>Moore & Parr, 1978</u>	Burkheimer et al., 1989; Cohen, Fathman, & Merino, 1976; McSpadden, 1980; Ed. Op. Concepts, 1991a; El Paso, 1987; 1990; 1992; Maldonado, 1977; Merino, 1976; Skoczytas, 1972; Stern, 1975; Teschner, 1990; Valladolid, 1991
	(N=10)	(N=4)	(N=7)

TBE v. ESL

TBE Better			Ames & Bicks, 1978
	(N=0)	(N=0)	(N=1)
No Difference	Ames & Bicks, 1978; Balasubramonium et al., 1973; Lum, 1974; Rossell, 1990; Yap, Enoki & Ishitani, 1988	Rossell, 1990; Yap, Enoki & Ishitani, 1988	Rossell, 1990; Yap, Enoki, & Ishitani, 1988
	(N=2)	(N=1)	(N=1)
TBE Worse	Lum, 1974; Rossell, 1990 <u>Valladolid, 1991</u>	Rossell, 1990 <u>Valladolid, 1991</u>	Rossell, 1990 <u>Valladolid, 1991</u>
	(N=2)	(N=2)	(N=2)

TBE v. Structured Immersion

No Difference	Ramirez et al., 1991; Ramos et al., 1967	Ramiriz et al., 1991	Barik, Swain, & Nwanunobi, 1977; Barik & Swain, 1975; Lambert & Tucker, 1972; Ramiriz et al., 1991; Ramos et al., 1967
	(N=2)	(N=1)	(N=5)
TBE Worse	Barik, Swain, & Nwanunobi, 1977; Barik & Swain, 1978; Bruck, Lambert, & Tucker, 1977; Day & Shapson, 1988;		Genessee & Lambert, 1983; Genessee et al., 1989; Gersten, 1985

Genessee & Lambert, 1983;
 Genessee, Lambert & Tucker,
 1977; Genessee et al., 1989;
~~Gersten, 1985~~; Malherbe,
 1946; Pena-Hughes &
 Solis, 1980

El Paso 1987; 1990,1992
 (N=12)

El Paso 1987; 1990,1992
 (N=3)

El Paso 1987; 1990,1992
 (N=5)

**Immersion v.
 ESL**

Immersion Better Barik & Swain, 1975;
 Becker and Gersten, 1982
 Lambert & Tucker, 1972

Gersten, 1985
 (N=4)

(N=0)

Gersten, 1985
 (N=1)

**TBE v. Maint.
 BE**

TBE Better ~~Medina & Escamilla, 1992**~~

(N=0)

(N=0)

(N=0)

*Studies are listed in more than one category if there were different effects for different grades or cohorts.

**Oral English achievement gains for preschool programs.

Appendix 2

METHODOLOGICALLY ACCEPTABLE STUDIES FROM ROSSELL & BAKER WITH NEW REJECTION CRITERIA:

- 1) U.S. SECONDARY TBE PROGRAM,
 - 2) NON-HISPANIC U.S. TBE PROGRAM,
 - 3) PROGRAM LESS THAN AN ACADEMIC YEAR,
 - 4) REDUNDANT STUDY (errors, not new criteria)
- (Original N = 70; New N=50)

- Alvarez, Juan M. (1975). "Comparison of Academic Aspirations and Achievement in Bilingual versus Monolingual Classrooms." Ph.D. dissertation, University of Texas.**
- American Institutes for Research. (Campeau, et al.,1975b). "Bilingual Education Program (Aprendemos En Dos Idiomas), Corpus Christi, Texas." Identification and Description of Exemplary Bilingual Education Programs. Palo Alto, California.** [CITED AS **CAMPEAU, ET AL IN SLAVIN AND CHEUNG (2004).**]
- Ames, J.S., and Bicks, Pat. 1978. An Evaluation of Title VII Bilingual/ Bicultural Program, 1977-78 School Year, Final Report, Community School District 22. Brooklyn, New York: School District of New York.*,**
- Ariza, Maria 1988. "Evaluating Limited English Proficient Students' Achievement: Does Curriculum Content in the Home Language Make a Difference?" Paper presented at the annual meeting of the American Educational Research Association, New Orleans. (4)
- Bacon, Herbert L. and Gerald D. Kidd, et al. 1982. "The Effectiveness of Bilingual Instruction with Cherokee Indian Students." Journal of American Indian Education :34-43. (2)
- Balasubramonian, K., H.N. Seelye, & C.R.E. de Weffer. 1973. "Do Bilingual Education Programs Inhibit English Language Achievement: A Report on an Illinois Experiment." Paper presented at the Seventh Annual Convention of Teachers of English to Speakers of Other Languages, San Juan, Puerto Rico.*,** (3)
- Barelay, Lisa. 1969. "The Comparative Efficacies of Spanish, English and Bilingual Cognitive Verbal Instruction with Mexican American Head Start Children." Ph.D. dissertation, Stanford University. (3)
- Barik, Henri, and Swain, Merrill. 1975. "Three Year Evaluation of a Large Scale Early Grade French Immersion Program: The Ottawa Study." Language Learning 25(1):1-30.*,**
- Barik, Henri C., and Merrill Swain. 1978. Evaluation of a Bilingual Education Program in Canada: The Elgin Study Through Grade Six. Switzerland: Commission Interuniversitaire Suisse de Linguistique Appliquee.*
- Barik, Henry C., Merrill Swain, and E.A. Nwanunobi. 1977. "English-French Bilingual Education: The Elgin Study Through Grade Five." Canadian Modern Language Review 33:459-475.*,**
- Bates, Enid May Buswell. 1970. "The Effects of One Experimental Bilingual Program on Verbal Ability and Vocabulary of First Grade Pupils." Ph.D. dissertation, Texas Tech University.** (3)
- Becker, Wesley C., and Russell Gersten. 1982. "A Follow-up of Follow Through: The Later Effects of the Direct Instruction Model on Children in Fifth and Sixth Grades." American Educational Research Journal 19:75-92.*

- Bruck, Margaret, Wallace E. Lambert, and G. Richard Tucker. 1977. "Cognitive Consequences of Bilingual Schooling: The St. Lambert Project Through Grade Six." Linguistics 24:13-33, January.*
- Burkheimer, Graham J., Conger, A.J., Dunteman, G.H., Elliott, B.G., Mowbray, K.A. 1989. Effectiveness of Services for Language-Minority Limited-English-Proficient Students. Report to the U.S. Department of Education.
- Campeau, Peggie L., A. Oscar H. Roberts, John E. Bowers, Melanie Austin, and Sarah J. Roberts. 1975. The Identification and Description of Exemplary Bilingual Education Programs. Palo Alto, CA: American Institutes for Research.* (4)
- Carsrud, Karen, and Curtis, John. 1980. ESEA Title VII Bilingual Program: Final Report. Austin, Texas: Austin Independent School District.*,**
- Ciriza, Frank. 1990a. Evaluation Report of the Preschool Project for Spanish-Speaking Children, 1989-90. San Diego City Schools, Planning, Research and Evaluation Division.
- Cohen, Andrew D. 1975a. A Sociolinguistic Approach to Bilingual Education. Rowley, MA: Newbury House Press, Publishers, Inc. *,**
- Cohen, Andrew D., Ann K. Fathman, and Barbara Merino. 1976. The Redwood City Bilingual Education Project, 1971-74: Spanish and English Proficiency, Mathematics and Language Use Over Time. Toronto: Ontario Institute for Studies in Education. *
- Cottrell, Milford C. 1971. "Bilingual Education in San Juan County, Utah: A Cross Cultural Emphasis". Paper presented at the annual meeting of the American Educational Research Association, New York.*,**
- Covey, D.D. 1973. "An Analytical Study of Secondary Freshman Bilingual Education and its Effect on Academic Achievement and Attitudes of Mexican American Students". Ph.D. dissertation, Arizona State University.*,** (1)
- Curiel, Herman, Stenning, Walter & Cooper Stenning, Peggy. 1980. "Achieved Reading Level, Self-Esteem, and Grades as Related to Length of Exposure to Bilingual Education." Hispanic Journal of Behavioral Sciences 2(4):389-400. (4)
- Curiel, Herman. 1979. "A Comparative Study Investigating Achieved Reading Level, Self-Esteem, and Achieved Grade Point Average Given Varying Participation." Ph.D. dissertation, Texas A. & M. University.
- Danoff, Malcolm N.; Arias, Beatriz M.; Coles; Gary J.; McLaughlin, Donald H.; and Reynolds, Dorothy J. 1977. Evaluation of the Impact of ESEA Title VII Spanish/English Bilingual Education Programs, Volume I and II. Palo Alto, Calif.: American Institutes for Research.*,**
- _____ 1978. Evaluation of the Impact of ESEA Title VII Spanish/English Bilingual Education Program, Vol III and IV. Palo Alto: American Institutes for Research.*,**
- Day, Elaine M. and Shapson, Stan M. 1988. "Provincial Assessment of Early and Late French Immersion Programs in British Columbia, Canada." Paper presented at the annual meeting of the American Educational Research Association, New Orleans, April.
- de Weffer, Rafaela de Carmen Elizondo. 1972. "Effects of First Language Instruction in Academic and Psychological Development of Bilingual Children." Ph.D. dissertation, Illinois Institute of Technology.** (3)

de la Garza, Jesus Valenzuela and Medina, Marcello. 1985. "Academic Achievement as Influenced by Bilingual Instruction for Spanish Dominant Mexican American Children." Hispanic Journal of Behavioral Sciences 7(3):247-259. **[ERROR WAS SUPPOSED TO BE IN REJECTION BIBLIOGRAPHY.]**

~~Educational Operations Concepts, Inc. 1991a. An Evaluation of the Title VII ESEA Bilingual Education Program for Hmong and Cambodian Students in Junior and Senior High School. St. Paul, MN: (2)~~

~~Educational Operations Concepts, Inc. 1991b. An Evaluation of the Title VII ESEA Bilingual Education Program for Hmong and Cambodian Students in Kindergarten and First Grade. St. Paul, MN: (2)~~

El Paso Independent School District. 1987. Interim Report of the Five-Year Bilingual Education Pilot 1986-87 School Year. El Paso, TX: Office for Research and Evaluation.

El Paso Independent School District. 1990. Bilingual Education Evaluation: the Sixth Year in a Longitudinal Study. El Paso, TX: Office for Research and Evaluation, September.

El Paso Independent School District. 1992. Bilingual Education Evaluation. El Paso, TX: Office for Research and Evaluation, November.

Genesee, Fred and W.E. Lambert. 1983. "Trilingual Education for Majority-Language Children." Child Development 54:105-114.

Genesee, Fred; Holobow, Naomi E., Lambert, Wallace E., and Chartrand, Louise. 1989. "Three Elementary School Alternatives for Learning through a Second Language." The Modern Language Journal 73:250-263.

Genesee, Fred., Wallace E. Lambert, and G.E. Tucker. 1977. An Experiment in Trilingual Education. Montreal: McGill University.*

Gersten, Russell. 1985. "Structured Immersion for Language Minority Students Results of a Longitudinal Evaluation." Educational Evaluation and Policy Analysis 7:187-196.*

Huzar, Helen. 1973. "The Effects of an English-Spanish Primary Grade Reading Program on Second and Third Grade Students." M.Ed. thesis, Rutgers University.*,**

~~Kaufman, Maurice. 1968. "Will Instruction in Reading Spanish Affect Ability in Reading English?" Journal of Reading 11:521-27.*,** (1)~~

Lambert, W.E., and Tucker, G.R. 1972. Bilingual Education of Children: The St. Lambert Experience. Rowley, Mass.: Newbury House Press.*,**

Lampman, Henry P. 1973. "Southeastern New Mexico Bilingual Program. Final Report." Artesia, N.M.: Artesia Public Schools.**

~~Layden, Russell Glenn. 1972. "The Relationship between the Language of Instruction and the Development of Self Concept, Classroom Climate and Achievement of Spanish Speaking Puerto Rican Children." Ph.D. dissertation, University of Maryland.** (3)~~

Legarreta, Dorothy. 1979. "The Effects of Program models on Language Acquisition by Spanish Speaking Children." TESOL Quarterly 13(4):521-34.*,**

- Lum, John Bernard. 1971. "An Effectiveness Study of English as a Second Language (ESL) and Chinese Bilingual Methods." Ph.D. dissertation, University of California at Berkeley.**,** (2)
- Maldonado, Jesus Ruben. 1977. "The Effect of the ESEA Title VII Program on the Cognitive Development of Mexican American students." Ph.D. dissertation, University of Houston.
- Malherbe, E.C. 1946. The Bilingual School. London: Longmans Green.**
- ~~Matthews, T. 1979. An Investigation of the Effects of Background Characteristics and Special Language Services on the Reading Achievement and English Fluency of Bilingual Students. Seattle, Wash: Seattle Public School, Department of Planning Research and Evaluation.**,** (2)~~
- McConnell, Beverly Brown. 1980a. "Effectiveness of Individualized Bilingual Instruction for Migrant Students." Ph.D. dissertation, Washington State University.**,**
- _____ 1980b. Individualized Bilingual Instruction. Final Evaluation, 1978-79 Program. Pullman, Wash.** [NEVER COUNTED AS ADDITIONAL STUDY/]
- McSpadden, J.R. 1979. Acadiana Bilingual Bicultural Education Program: Interim Evaluation Report, 1978-79. Lafayette Parish, LA.**,** (2)
- _____ 1980. Acadiana Bilingual Bicultural Education Program. Interim Evaluation Report 1979-80. Lafayette Parish, La.**,** (2)
- Medina, Marcello and Escamilla, Kathy. 1992. "Evaluation of Transitional and Maintenance Bilingual Programs." Urban Education 27(3):263-290. (2)
- Melendez, William Anselmo. 1980. "The Effect of the Language of Instruction on the Reading Achievement of Limited English Speakers in Secondary Schools." Ph.D. dissertation, Loyola University of Chicago.**
- Moore, Fernie.B., and Gerald D. Parr. 1978. "Models of Bilingual Education: Comparisons of Effectiveness." The Elementary School Journal 79:93-97. **,**
- Morgan, Judith Claire. 1971. "The Effects of Bilingual Instruction on the English Language Arts Achievement of First Grade Children." Ph.D. dissertation, Northwestern State University of Louisiana.** (2)
- Pena-Hughes, Eva, and Juan Solis. 1980. ABCs. McAllen, Texas: McAllen Independent School District.**,**
- Plante, Alexander, J. 1976. A Study of the Effectiveness of the Connecticut "Pairing" Model of Bilingual-Bicultural Education. Hamden, Conn.: Connecticut Staff Development Cooperative.**,**
- Powers, Stephen. 1978. "The Influence of Bilingual Instruction on Academic Achievement and Self-Esteem of Selected Mexican-American Junior High School Students." Ph.D. dissertation, University of Arizona.**
- Prewitt Diaz, Joseph O. 1979. "An Analysis of the Effects of a Bilingual Curriculum on Monolingual Spanish Ninth Graders as Compared with Monolingual English and bilingual Ninth Graders with Regard to Language Development, Attitude toward School and Self-Concept." Ph.D. dissertation, University of Connecticut.**

- Ramirez, J. David, Pasta, David J., Yuen, Sandra D., Billings, David K., Ramey, Dena R. 1991. Final Report: Longitudinal Study of Structured Immersion Strategy, Early-Exit and Late-Exit Transitional Bilingual Education Programs for Language-Minority Children. San Mateo, CA: Aguirre International, report to the U.S. Department of Education, Washington, D.C.
- Ramos, M.; Aguilar, J.V., and Sibayan, B.F. 1967. The Determination and Implementation of Language Policy. Philippine Center for Language Study Monograph Series 2. Quezon City, The Philippines: Alemor/Phoenix.*,**
- Rossell, Christine H. 1990. "The Effectiveness of Educational Alternatives for Limited-English-Proficient Children." In Learning in Two Languages. Ed. Gary Imhoff. New Brunswick, N.J.: Transaction Publishers.
- Rothfarb, Sylvia H., Ariza, Maria J. and Urrutia, Rafael. 1987. Evaluation of the Bilingual Curriculum Content (BCC) Project: A three-Year Study Final Report. Dade County: Office of Educational Accountability.
- Skoczylas, Rudolph V. 1972. "An Evaluation of Some Cognitive and Affective Aspects of a Spanish-English Bilingual Education Program." Ph.D. dissertation, University of New Mexico.*,**
- Stebbins, Linda B.; St. Pierre, Robert G.; Proper, Elizabeth C.; Anderson, Richard B.; and Carva, Thomas R. 1977. Education as Experimentation: A Planned Variation Model Volume IV-A An Evaluation of Follow Through. Cambridge, Mass.: ABT Associates.*,**
- Stern, Carolyn. 1975. Final Report of the Compton Unified School District's Title VII Bilingual-Bicultural Project: September 1969 through June 1975. Compton City, Calif.: Compton City Schools.*,**
- Teschner, Richard V. 1990. "Adequate Motivation and Bilingual Education." Southwest Journal of Instruction 9:1-42.
- Valladolid, Lupe A. 1991. "The Effect of Bilingual Education on Students' Academic Achievement as They Progress Through a Bilingual Program." Ph.D. dissertation, San Diego, CA: United States International University.
- Vasquez, Miriam. 1990. "A Longitudinal Study of Cohort Academic Success and Bilingual Education." Ph.D. dissertation, University of Rochester.
- ~~Yap, Kim O. and Enoki, Donald Y. and Ishitani, Patricia. 1988. "SLEP Student Achievement: Some Pertinent Variables and Policy Implications." A paper presented at the annual meeting of the American Educational Research Association, New Orleans, April 5-9. (2)~~
- Webb, John A. Clerc, R. J., and Gavito, Alfredo. 1987. Houston Independent School District: Comparison of Bilingual and Immersion Programs Using Structural Modeling. Houston Independent School District.
- Zirkel, Perry A. 1972. "An Evaluation of the Effectiveness of Selected Experimental Bilingual Education Programs in Connecticut". Ph.D. dissertation, University of Connecticut.*,**

* From Rossell and Ross, 1986.

**From Baker and de Kanter, 1983b.¹

¹ Olesini, 1971 was accidentally inserted into Table 1 and Appendix 1. It had been considered methodologically acceptable in Baker and deKanter 1991, but not in Baker and deKanter, 1983b. The merger of the earlier Baker and deKanter review with the Rossell and Ross review was supposed to be based on Baker and deKanter, 1993b, not Baker and deKanter, 1991.

Appendix 3

Greene List of Unacceptable Studies with Rossell Comments

Studies Excluded Because They Are Redundant

Ariza, M. (1988). *Evaluating limited English proficient students' achievement: Does curriculum content in the home language make a difference?* Paper presented at the April meetings of the American Educational Research Association, New Orleans. Redundant with Rothfarb et al., 1987.

- I agree.
- R & B Finding: No difference between transitional bilingual education (TBE) and mainstream.

Barik, H., and Swain, M. (1978). *Evaluation of a bilingual education program in Canada: The Elgin Study through grade six*. Switzerland: Commission Interuniversitaire Suisse de Linguistique Appliquee. Redundant with Barik et al. 1977.

- I disagree. This evaluation by Barik and Swain analyzed the 1975-76 school year and the evaluation by Barik, Swain, and Nwanumobi analyzed the 1974-75 school year.
- R & B finding: TBE worse than structured immersion.

Cohen, A. D., Fathman, A. K., & Merino, B. (1976). *The Redwood City bilingual education report, 1971-1974: Spanish and English proficiency, mathematics, and language-use over time*. Toronto: Ontario Institute for Studies in Education. Redundant with Cohen 1975.

- I disagree. Cohen, 1975 evaluated bilingual education for grades 1-3 from 1969-1972 whereas Cohen, Fathman, and Merino evaluated it for grades 3-5 from 1972-1975.
- R & B finding: TBE worse than mainstream in Cohen, Fathman, and Merino; No difference in Cohen.

Curiel, H., Stenning, W., & Cooper-Stenning, P. (1980). Achieved reading level, self-esteem, and grades as related to length of exposure to bilingual education. *Hispanic Journal of Behavioral Sciences*, 2, 389-400. Redundant with Curiel, 1979.

- I agree.
- R & B finding: TBE worse than mainstream.

Danoff, M. N., Coles, G. J., McLaughlin, D. H., & Reynolds, D. J. (1977b). *Evaluation of the impact of ESEA Title VII Spanish/English bilingual education programs, Vol. I: Study design and interim findings*. Palo Alto: American Institutes for Research. Redundant with Danoff et al. 1977a.

- Greene is wrong. We did not count 1977a and 1977b as two separate studies (see Appendix 1a).
- R & B finding: TBE worse than mainstream.

(1978a). *Evaluation of the impact of ESEA Title VII Spanish/English bilingual education programs, Vol. III: Year two impact designs*. Palo Alto: American Institutes for Research.

- I disagree. The 1977 study is of 37 school districts during the 1975-76 school year; the 1978 study is an analysis of data collected after the 1977 study in a smaller sample of schools.
- R & B finding: TBE worse than mainstream.

(1978b). *Evaluation of the impact of ESEA Title VII Spanish/English bilingual education programs, Vol. IV: Overview of the study and findings*. Palo Alto: American Institutes for Research.

- Greene is wrong. We did not count this as a separate study (see Appendix 1a).

Educational Operations Concepts, Inc. (1991b). *An evaluation of the Title VII ESEA bilingual education program for Hmong and Cambodian students in kindergarten and first grade*. St. Paul. Redundant with Educational Operations Concepts, Inc. 1991a.

- I disagree. Educational Operations Concepts 1991a is of junior and senior high school students and 1991b is of kindergarten and first grade students.
- R & B finding: No difference between TBE and mainstream for K-1 students in language and math; TBE worse than mainstream for K-1 students in reading and junior/senior high students in reading, language, and math.

El Paso Independent School District. (1990). *Bilingual education evaluation: The sixth year in a longitudinal study*. El Paso: Office for Research and Evaluation. Redundant with El Paso 1987.

- I disagree. El Paso 1987 analyzed grades 1-3 in 1986-87; El Paso 1990 analyzed grades PK-6 in 1989-90.
- R & B finding: TBE worse than mainstream. R & B made an error in constructing the table, it should be TBE worse than structured immersion.

El Paso Independent School District. (1992). *Bilingual education evaluation*. El Paso: Office for Research and Evaluation. Redundant with El Paso 1987.

- I disagree. El Paso 1987 analyzed grades 1-3 in 1986-87; El Paso 1990 analyzed grades PK-6 in 1989-90; and El Paso 1992 analyzed grades 3-11 in the 1990-91 and 1991-92 school years.
- R & B finding: TBE worse than mainstream. R & B made an error in constructing the table, it should be TBE worse than structured immersion.

Genesee, F., Lambert, W. E., & Tucker, G. E. (1977). *An experiment in trilingual education*. Montreal: McGill University. Redundant with Genesee et al 1983.

- I disagree. Genesee, Lambert, and Tucker 1977 analyzed students in grades 3-5 in two immersion schools in 1976-77; Genesee and Lambert, 1983 analyzed only fifth graders who had been in the program for five years. Although there is no information on when the data was collected in the latter study, it is obviously a different sample from the earlier study.
- R & B Finding: TBE worse than structured immersion.

McConnell, B. B. (1980b). *Individualized bilingual instruction, final evaluation, 1978-1979 program*. Pullman. Redundant with McConnell 1980a.

- Greene is wrong. We did not count the various versions of McConnell's study of the Pullman bilingual education program as separate studies (see Appendix 1a).
- R & B Finding: TBE better than submersion (a mainstream classroom) in reading.

(1980c). *Individualized bilingual instruction for migrants*. Paper presented at the October meeting of the International Congress for Individualized Instruction, Windsor.

- Greene is wrong. We did not count the various versions of McConnell's study of the Pullman bilingual education program as separate studies (see Appendix 1a).
- R & B Finding: TBE better than submersion (mainstream classroom) in reading.

McSpadden, J. R. (1980). *Arcadia bilingual bicultural education program: Interim evaluation report, 1979-80*. Lafayette Parish. Redundant with McSpadden 1979.

- I disagree. Neither of these studies is available any longer (thrown out when Keith Baker retired from the Department of Education in 1997) so Greene must have just guessed they were redundant.
- If you look at Appendix 1a, you can see that McSpadden 1979 shows no difference between TBE and a mainstream classroom and McSpadden 1980 shows TBE to be worse so clearly they have different samples.

Teschner, R. V. (1990). Adequate motivation and bilingual education. *Southwest Journal of Instruction*, 9, 1-42. Redundant with El Paso, 1990.

- I disagree. Teschner analyzes third graders from spring 1987 through spring 1989; El Paso 1990 analyzed grades PK-6 in 1989-90.
- R & B finding: TBE worse in reading, language, and math.

Studies Excluded Because They Are Unavailable

American Institutes for Research. (1975b). *Bilingual education program (Aprendemos En Dos Idiomas)*. Corpus Christi. Palo Alto: Identification and Description of Exemplary Bilingual Education Programs.

- Greene is wrong. The study is available and both Slavin and Cheung (2004) and I have copies. It is cited in Slavin and Cheung as Campeau, et al. (1975).
- R & B finding: TBE better in reading in some grades and no different in others.

Lambert, W. E., & Tucker, G. R. (1972). *Bilingual education of children: The St. Lambert experience*. Rowley, MA: Newbury House.

- Greene is wrong. It is available at many libraries. I have a copy of it that I would have provided to him had he asked.
- R & B finding: No difference between TBE and structured immersion in math; structured immersion better than ESL in reading.

McSpadden, J. R. (1979). *Arcadia bilingual bicultural education program: Interim evaluation report, 1978-79*. Lafayette Parish.

- Greene is correct, but the 1980 report has also disappeared.
- I would now exclude this study as it is U.S., but not of Spanish speakers.
- R & B finding: TBE worse (1980); TBE no different (1979)

Morgan, J. C. (1971). *The effects of bilingual instruction of the English language arts achievement of first grade children*. Doctoral dissertation, Northwestern State University of Louisiana.

- Greene is wrong. This is available through Dissertation Abstracts at the University of Michigan. I have a copy of it that I would have provided to him had he asked.
- I would, however, now exclude this study as it is U.S. but not Spanish speakers.
- R & B finding: TBE better than mainstream.

Ramos, M., Aguilar, J. V., & Sibayan, B. F. (1967). *The determination and implementation of language policy* (Monograph Series 2). Quezon City: Philippine Center for Language Study.

- Greene is wrong. It is available at many libraries. I have a copy of it that I would have provided to him had he asked.
- R & B finding: No difference between TBE and structured immersion

Studies Excluded Because They Are Not Evaluations Of Bilingual Programs

Becker, W. C. & Gersten, R. (1982). A follow-up of follow through: The latter effects of the Direct Instruction Model on children in fifth and sixth grades. *American Educational Research Journal*, 19, 75-92.

- I agree with the point, but R & B did not need to exclude---our goal was broader.
- R & B finding: structured immersion better than ESL

Campeau, P. L., Roberts, A., Oscar H., Bowers, J. E., Austin, M., & Roberts, S. J. (1975). *The identification and description of exemplary bilingual education programs*. Palo Alto: American Institutes for Research.

- I disagree. This is an evaluation of bilingual education programs as the title indicates. However, I have now decided to exclude on the grounds that there is insufficient information to justify inclusion.
- R & B finding: TBE better than mainstream.

Webb, J. A., Clerc, R. J., & Gavito, A. (1987). *Houston Independent School District: Comparison of bilingual and immersion programs using structural modeling*. Houston Independent School District.

- I disagree. This is an evaluation of bilingual education programs as the title indicates and the outcome is achievement as well as other outcomes.
- R & B finding: TBE worse than mainstream.

Studies Excluded Because There Is Not An Appropriate Control Group

Barik, H., Swain, M. & Nwanunobi, E. A. (1977). English-French bilingual education: The Elgin Study through grade five. *Canadian Modern Language Review*, 33, 459-475.

- I disagree. There were two comparisons in this study. The treatment group is native English speakers in Partial French Immersion (i.e. bilingual education). In the first comparison for English language outcomes, the control group was native English speakers in English education (immersion). The PFI (i.e. TBE) students did worse in English reading and math than the students educated completely in English. In the second comparison for French language outcomes, the control group was Native English speakers enrolled in TFI (total French Immersion). The PFI (i.e. TBE) students did worse in French reading, but there was no difference in math. We only used the findings for French outcomes, but we could have justified using the first findings as well.
- R & B finding: TBE worse than structured immersion in reading, no different in math.

Bruck, M., Lambert, W. E., & Tucker, G. R. (1977). Cognitive consequences of bilingual schooling: The St. Lambert project through grade six. *Linguistics*, 24, 13-33.

- I disagree. There were two control groups: native English speakers receiving some French instruction and native French speakers receiving all French instruction. The groups were

carefully compared for equivalence using socioeconomic status, IQ, language achievement, and home background factors based on home interviews.

- R & B finding: TBE worse than structured immersion in reading.

Burkheimer, G. J., Conger, A.J., Dunteman, G.H., Elliott, B.G., & Mowbray, K.A. (1989). *Effectiveness of services for language-minority limited- English-proficient students*. Report to the U.S. Department of Education.

- I disagree. Burkheimer et al. is a very sophisticated multiple regression analysis controlling for many instructional variables including the amount of instruction in Spanish. The only students studied were limited English proficient Spanish speakers.
- R & B finding: TBE better than mainstream in reading, language, and math.

Day, E. M., & Shapson, S. M. (1988). *Provincial assessment of early and late French immersion programs in British Columbia, Canada*. Paper presented at the April meetings of the American Educational Research Associates, New Orleans. No background controls or individual level data reported.

- I disagree. Randomly selected native English speakers in early and late French immersion were compared to each other and to both native English speakers in English education and Francophone students matched on the basis of socioeconomic status and academic ability. Analysis of variance was used. Early immersion students did better than late immersion (i.e. TBE) students.
- R & B finding: TBE worse than structured immersion.

El Paso Independent School District. (1987). *Interim report of the five-year bilingual education pilot 1986-1987 school year*. El Paso: Office for Research and Evaluation. No background or pretest controls.

- I disagree. The students were all Spanish speaking ELLs matched on important demographic variables. However, I now believe this study compared TBE to structured immersion. Therefore, the new finding is that TBE is worse than structured immersion.
- R & B finding: TBE worse than mainstream classroom.

Genesee, F., & Lambert, W. E. (1983). Trilingual education for majority-language children. *Child Development, 54*, 105-114. No background controls.

- I disagree; pre-test controls were used and similar groups were compared.
- R & B finding: TBE worse than structured immersion in reading and math.

Genesee, F., Holobow, N. E., Lambert, W. E., & Chartrand, L. (1989). Three elementary school alternatives for learning through a second language. *The Modern Language Journal, 73*, 250-263. No background controls.

- I disagree; pre-test controls were used and similar groups were compared.
- R & B finding: TBE worse than structured immersion in reading and math.

Gersten, R. (1985). Structured immersion for language-minority students: Results of a longitudinal evaluation. *Educational Evaluation and Policy Analysis, 7*, 187-196. No background controls.

- I disagree. The pre-test is sufficient. However, I now believe that because the students were Asian and of different languages, the program the district called bilingual education was in fact ESL pullout. Only the primary (first & second grade) results are used since the

intermediate elementary group does not have an appropriate control group. The comparison should be structured immersion v. ESL (incorrectly called bilingual).

- R & B finding: Structured immersion superior to TBE (should have been structured immersion superior to ESL).

Malherbe, E. C. (1946). *The bilingual school*. London: Longmans Green. No background or pretest controls.

- I disagree. This was a random sample of Afrikaans and English speaking students taught in various language environments, Afrikaans, English, bilingual. Despite the fact that there was no significant difference in intelligence between the students in the different school language environments, the more language exposure a student received in school the better the student did in that language compared to similar home language background students, even if the language of the school was not their native tongue.
- R & B finding: TBE worse than structured immersion.

McConnell, B. B. (1980a). *Effectiveness of individualized bilingual instruction for migrant students*. Doctoral dissertation, Washington State University.

- I disagree. The control group at each grade is the pretest scores of all the students who are in the treatment as they enter at each grade level. So the control group for a student who entered in kindergarten and exits at 4th grade is the pretest scores for students who entered in 4th grade.
- R & B finding: TBE is better than submersion in reading and math.

Medina, M., & Escamilla, K. (1992). Evaluation of transitional and maintenance bilingual programs. *Urban Education*, 27, 263-290.

- I agree. The TBE group was Vietnamese students which I now know means they were was not in a real bilingual education program. In addition, the maintenance bilingual education (MBE) was Hispanic and the comparisons and statistical analyses did not take into account the differences in ethnicity.
- R & B finding: TBE superior to MBE.

Melendez, W. A. (1980). *The effect of the language of instruction on the reading achievement of limited English speakers in secondary schools*. Doctoral dissertation, Loyola University of Chicago. No background controls.

- I disagree. A pretest is a sufficient background control.
- R & B finding: TBE worse than submersion in reading.

Stern, C. (1975). *Final report to the Compton Unified School District's Title VII Bilingual/Bicultural Project: September 1969 through June 1975*. Compton: Compton City Schools.

- I disagree. A pretest is a sufficient control.
- R & B finding: TBE worse than submersion in reading and math.

Vasquez, M. (1990). *A longitudinal study of cohort academic success and bilingual education*. Doctoral dissertation, University of Rochester. No background controls.

- I disagree. This is a two-way Spanish immersion program which R & B classified as TBE for the Spanish speakers. Parents chose the program for their children. Multiple regression analysis was used to test the effect of staying longer in the program. Control variables were previous years achievement, years in program, kindergarten English proficiency, and school attended.

- R & B finding: No difference in reading and math between TBE and submersion.

Studies Excluded Because The Effects Are Measured after An Unreasonably Short Period

Barclay, L. (1969). *The comparative efficacies of Spanish, English, and Bilingual Cognitive Verbal Instruction with Mexican American Head Start children*. Doctoral dissertation, Stanford University. Positive Average Effect.

- I agree.

Layden, R. G. (1972). *The relationship between the language of instruction and the development of self-concept, classroom climate, and achievement of Spanish speaking Puerto Rican children*. Doctoral dissertation, University of Maryland. Negative Average Effect.

- I agree.

Studies Excluded Because They Inadequately Control Differences Between Bilingual And English-Only Students

Alvarez, J. (1975). *Comparison of academic aspirations and achievement in bilingual versus monolingual classrooms*. Doctoral dissertation, University of Texas at Austin. Negative Average Effect.

- I disagree and so do Slavin and Cheung (2004).
- R & B finding: No difference.

Ames, J., & Bicks, P. (1978). *An evaluation of Title VII Bilingual/Bicultural Program, 1977-1978 school year, final report*. Community School District 22. Brooklyn. School District of New York.

- I disagree. The students in the ESL pullout were students who were eligible for bilingual education, but for one reason or another were not enrolled. In other words, they were initially equal. In addition, the analysis controlled for pretest scores of the treatment and control groups.
- R & B finding: Positive effect on math; no effect on reading.

Balasubramonian, K., Seelye, H., & de Weffer, R.C.E.(1973). *Do bilingual education programs inhibit English language achievement: A report on an Illinois experiment*. Paper presented at the 7th Annual Convention of Teachers of English to Speakers of Other Languages, San Juan. Positive Average Effect.

- I disagree, but will exclude on grounds that it is of too short a duration.
- R & B finding: No effect on reading.

Barik, H., & Swain, M. (1975). Three year evaluation of a large-scale early grade French immersion program: The Ottawa-Study. *Language Learning*, 25, 1-30. Negative Average Effect.

- I disagree. The analysis controlled for pretest scores of the treatment and control groups.
- R & B finding: Structured immersion better than TBE in reading.

Bates, E. M. B. (1970). The effects of one experimental bilingual program on verbal ability and vocabulary of first grade pupils. Doctoral dissertation, Texas Tech University. Negative Average Effect.

- I disagree, but will exclude on grounds that the program is of too short a duration.
- R & B finding: TBE no different in math, but inferior in reading.

Carsrud, K, & Curtis, J. (1980). *ESEA Title VII Bilingual Program: Final report*. Austin: Austin Independent School District. No statistical tests reported. Positive Average Effect.

- I disagree. Analysis of covariance used to statistically control for differences between treatment and control group. Like groups compared.
- R & B finding: no difference in reading or math.

Ciriza, F. (1990a). *Evaluation report of the Preschool Project for Spanish-speaking children, 1989-1990*. San Diego: Planning, Research and Evaluation Division. San Diego City Schools. Positive Average Effect.

- I disagree. Since the treatment and the control group had identical LAS scores in English before the treatment and virtually identical scores in Spanish, we considered analysis of variance to be adequate.
- R & B finding: no difference in reading.

Cohen, A. D. (1975). *A sociolinguistic approach to bilingual education*. Rowley, MA: Newbury House Press. Negative Average Effect.

- I disagree. So do Slavin and Cheung (2004).
- R & B finding: TBE better in math and no different in reading.

Cottrell, M. C. (1971). *Bilingual education in San Juan Co., Utah: A cross-cultural emphasis*. Paper presented at the April meetings of the American Educational Research Association, New York City. Negative Average Effect.

- I disagree, but would exclude it now because it is in the U.S., but not Spanish speakers.
- R & B Finding: No difference.

Curiel, H. (1979). *A comparative study investigating achieved reading level, self-esteem, and achieved grade point average given varying participation*. Doctoral dissertation, Texas A&M. Negative Average Effect.

- I disagree. The students were Hispanic 6th graders and the same students in 7th grade. The control group was randomly selected from among Hispanic students who had not been in bilingual education in elementary school. Analyses of family background indicated that the two groups of students were virtually identical in socioeconomic status and home language. Analysis of variance was used to analyze the differences in academic achievement between the two groups.
- R & B finding: TBE worse than submersion in reading and language.

de la Garza, J. V., & Marcella, M. (1985). Academic achievement as influenced by bilingual instruction for Spanish-dominant Mexican American children. *Hispanic Journal of Behavioral Sciences*, 7, 247-259. Positive Average Effect.

- I agree, but this study was not analyzed as methodologically acceptable (see Appendix 1aa). This was a bibliographical error in Rossell and Baker. The study should have been in the bibliographic listing in Appendix B, Methodologically Unacceptable Studies.

de Weffer, R. C. Elizondo (1972). *Effects of first language instruction in academic and psychological development of bilingual children*. Doctoral dissertation, Illinois Institute of Technology. Positive Average Effect.

- I disagree, but would exclude on grounds of too short a time period.
- R & B finding: No difference in reading (oral) and math.

Educational Operations Concepts, Inc. (1991a). St. Paul: An evaluation of the Title VII ESEA Bilingual Education Program for Hmong and Cambodian students in junior and senior high school. Positive Average Effect.

- I disagree, but would exclude on the grounds this is a U.S. program and Asian students not in a real bilingual education program.
- R & B finding: TBE worse in reading, language, and math than submersion.

Lampman, H. P. (1973). Southeastern New Mexico bilingual program: Final report. Artesia: Artesia Public Schools. Positive Average Effect.

- I disagree. The students are very carefully matched on age, non-verbal IQ, verbal IQ, family poverty status, family income, family structure, number of children, parents' occupation, parents' education, and home language. We probably will not be able to construct an effect size for this study, however, as there is no standard deviation or other information that could be used to estimate a standard deviation.
- R & B finding: TBE no different in reading than submersion.

Legarreta, D. (1979). The effects of program models on language acquisition by Spanish-speaking children. *TESOL Quarterly*, 13, 521-534. Positive Average Effect.

- I disagree. The author uses a sophisticated multivariate analysis with pre-test information to compare different kinds of bilingual programs and all-English programs.
- R & B finding: TBE superior to submersion in oral comprehension and communication (classified as reading) in one comparison; TBE no different in same areas in another comparison.

Lum, J. B. (1971). An effectiveness study of English as a second language (ESL) and Chinese bilingual methods. Doctoral dissertation, University of California, Berkeley. Negative Average Effect.

- I disagree, but would exclude it now because it is U.S., but not Spanish speakers.
- R & B finding: TBE worse in some grades, but no different in others in reading.

Maldonado, J. R. (1974). The effect of the ESEA Title VII Program on the cognitive development of Mexican American students. Doctoral dissertation, University of Houston. Negative Average Effect.

- I disagree and so do Slavin and Cheung (2004).
- R & B finding: No difference in reading, language, and math.

Matthews, T. (1979). An investigation of the effects of background characteristics and special language services on the reading achievement and English fluency of bilingual students. Seattle: Seattle Public Schools: Department of Planning, Research and Evaluation. Negative Average Effect.

- I would exclude on the grounds this is a U.S. program and Asian students not in a real bilingual education program. The author implied he conducted a multiple regression analysis controlling for many variables, but there is no quantitative data.
- R & B finding: TBE no different in reading than submersion.

Moore, F. B. & Parr, G. D. (1978). Models of bilingual education: Comparisons of effectiveness. *The Elementary School Journal*, 79, 93-97. Negative Average Effect.

- I disagree. Analysis of covariance used to control for pretreatment differences.
- R & B finding: TBE worse in reading and no different in math from submersion; we neglected to add TBE worse in language to our summary table (see Appendix 1aa).

Peña-Hughes, E., & Solis, J. (1980). *ABC's*. McAllen: McAllen Independent School, District. Positive Average Effect.

- I disagree. Although we no longer have most of the text of the study, Baker and deKanter have a two page description of the study. It is a randomized experiment which also used analysis of covariance. All students were Hispanic limited English proficient.
- R & B finding: Structured immersion superior to TBE in reading.

Prewitt Diaz, J. O. (1979). An analysis of the effects of a bicultural curriculum on monolingual Spanish ninth graders as compared with monolingual English and bilingual ninth graders with regard to language development, attitude toward school, and self-concept. Doctoral dissertation, University of Connecticut. Positive Average Effect.

- I would now exclude because it is a high school program and high school programs are not real bilingual education programs.
- R & B finding: No difference between TBE and mainstream classroom in reading.

Stebbins, L. B., St. Pierre, R. G., Proper, E. C., Anderson, R. B., & Carva, T. (1977). *Education as experimentation: A Planned Variation Model, Vol. IV-A. An evaluation of follow through*. Cambridge: ABT Associates. Positive Average Effect.

- I disagree. This is a very sophisticated study with pretest and many other control variables.
- R & B finding: No difference between TBE and submersion in reading and math.

Valladolid, L. A. (1991). The effects of bilingual education of students' academic achievement as they progress through a bilingual program. Doctoral dissertation, United States International University. No background or pretest controls. Negative Average Effect.

- I disagree. Hispanic students were matched on important variables: they were required to have entered school limited-English speaking, been in the bilingual or mainstream program for three years, and to be in grade five. A split-plot factorial ANOVA was used to remove initial differences between students.
- R & B finding: TBE worse than submersion in reading, language, and math. Upon rereading, we should have classified this as TBE worse than ESL in reading, language, and math. I have moved this study in the revised tables.

Yap, K. O., Enoki, D. Y., & Ishitani, P. (1988). *SLEP student achievement: Some pertinent variables and policy implications*. Paper presented at the April meetings of the American Educational Research Association, New Orleans. No background or pretest controls. Negative Average Effect.

- I agree. This particular paper has insufficient information on the programs, the analyses, and the students. (I believe that when we were reviewing this study we also had a longer school district report with more data that is now lost). If the lack of information didn't disqualify it, the fact that Asian students are in the bilingual program should.
- R & B finding: TBE no different in reading, language, and math than ESL.

Zirkel, P. A. (1972). An evaluation of the effectiveness of selected experimental bilingual education programs in Connecticut. Doctoral dissertation, University of Connecticut. Positive Average Effect.

- I disagree. Hispanic students in bilingual and mainstream classrooms were matched on grade level, number, school attended, age, socioeconomic status, sex, and language dominance and many other home characteristics. The researcher also personally observed all programs and changed their label to fit the facts of what was observed. Analysis of covariance was used to control for pretreatment characteristics. R & B only analyzed the true bilingual-mainstream comparison, not the quasi-bilingual comparisons.
- R & B finding: TBE superior to submersion in reading for later grades; no difference in grade 1.

Appendix 4

Slavin & Cheung (SC) Studies with Rossell (CR) Comments

Cited by*	CR Comments	Authors	Slavin & Cheung Remarks	CR on Problems w/ ELLGroup	CR on Problems with Treatment	Rossell Remarks on Other Problems (See also Appendix 3 Comments)
Methodologically Adequate--Elementary Reading						
RB		Alvarez (1975)			English and Spanish Reading	INCONSISTENT: 2nd grade students; pretest given in 1st grade AFTER treatment began, students probably fluent in English
RB	Greene	Bacon et al (1982)		U.S.-Not Spanish		INCONSISTENT: 8th grade students; pretest given 8 years AFTER treatment began
RB	RB accepted only Corpus Christi study	Campeau et al (1975)	5 separate studies met criteria		English and Spanish Reading	INCONSISTENT: Santa Fe: tests apparently given at the end of first grade to children already fluent in English; SC ERROR: R & B only considered the Corpus Christi evaluation scientific (listed under AIR author)
RB & W		Cohen (1975)			English and Spanish Reading	INCONSISTENT: Not clear when pretest given
		Doebler & Mardis (1980)		U.S.-Not Spanish		PROBLEM: 2nd grade students already fluent in English

RB & W	Greene	Huzar (1973)			English and Spanish Reading	
		J. A. Maldonado (1994)				Problem: Special Education students.
RB		J.R. Maldonado (1977)			English and Spanish Reading	
RB		Morgan (1971)		U.S.: -Not Spanish	English and French Reading	
RB	Greene	Plante (1976)			English and Spanish Reading	
RB	Greene	Ramirez et al (1991)				
		Saldade et al (1985)				INCONSISTENT: Pretest given after treatment began
Methodologically Adequate--Secondary Reading						
RB & W	Greene	Covey (1973)		Not Elementary		Reject because secondary--not real bilingual.
RB & W	Greene	Kaufman (1968)		Not Elementary		Reject because secondary--not real bilingual.
Canadian Studies of French Immersion A62						
RB		Barik & Swain (1975)		Disagree--see discussion in text and in Appendix 3.		
RB		Barik et al (1977)				
RB		Bruck et al (1977)				

RB		Day & Shapson (1988)		Disagree--see discussion in text and in Appendix 3.
RB		Genesee & Lambert (1983)		
RB		Genesee et al (1989)		
RB & W	W is error	Lambert & Tucker (1972)		
Students Were Not Learning the Societal Language				
RB		Ramos et al (1967)	Learning English in the Phillipines	Disagree--see discussion in text and in Appendix 3.
No Reading Outcomes (Oral Language Only)				
RB & W		Lum (1971)		U.S.: -Not Spanish Rossell and Baker had a different goal, one that included oral outcomes. Reject Lum because of ELL group.
RB	Author error >	Bates & May (1970)	6 months; no pretest data provided	
RB		Elizondo de Weffer (1972)	4 months; no reading outcomes; also preference for English language usage C>E	
RB & W		Legarreta (1979)		
RB	Greene	Rothfarb et al (1987)		
Pretests Were Given After Treatments Were Under Way				Both Greene (1997) and Rossell & Baker (1996) disagree that this is a problem that disqualifies a study.
RB & W	Greene	Danoff, Arias & Coles (1977a)		Disagree this is a problem. See rebuttal in text.
RB		Melendez (1980)		Disagree this is a problem. See rebuttal in text and Appendix 3.

RB	RB is error, should be W	Olesini (1971)		SC Error-Rossell and Baker did not consider this scientific.		
		Rosier & Holm (1980)		Rossell & Baker did not consider this scientific.		
RB	Greene	Rossell (1990)		Disagree--see discussion in text. Greene disagrees also.		
RB & W	Greene	Skoczylas (1972)	Large pretest differences; No separate analysis for Spanish dominant students; more English dominant children in the control group	Disagree--see discussion in text. Greene disagrees also.		
RB & W		Stern (1975)		Only one month delay.		
		Thomas & Collier (2002)	Separate studies in Maine & Houston			
RB		Valladolid (1991)		Disagree--see Appendix 3.		
RB		Yap, Enoki, & Ishitani (1988)		U.S.: -Not Spanish		Reject because of ELL group.
Redundant						
RB		Ariza (1988)	Redundant with Rothfarb (1987)	See rebuttal in Appendix 3. Only Ariza and Curiel are actually redundant by our standards.		
RB		Barik & Swain (1978)	Redundant with Barik et al (1977)			
RB		Cohen et al (1976)	Redundant with Cohen (1975)			
RB		Curiel et al (1980)	Redundant with Curiel (1979)			
RB & W		Danoff et al (1977b & 1978)	Redundant with Danoff (1977a)			
RB		El Paso ISD (1987 & 1990)	Redundant with El Paso ISD (1992)			
RB		Genesee, Lambert and Tucker (1977)	Redundant with Genesee et al (1983)			
RB		McConnell (1980a)	Redundant with McConnell (1980b)			

No Evidence of Initial Equality						
RB		Ames & Bicks (1978)	Large pretest difference; mixed grades and mixed languages			Disagree. R & B accepted studies that were methodologically sound, even if they did not report all data since we did not need all data for the vote count method. These studies matched students or otherwise controlled for pretreatment differences. See comments in text and Appendix 3.
RB		Barclay (1969)	Large pretest differences; 7 months			Reject because of short duration.
RB & W		Carsrud & Curtis (1979 & 1980)	Mixed Spanish and English dominant children in the analysis			Disagree (see above and Appendix 3)
RB		Cottrell (1971)	Poorly matched on SES. ANCOVA was used but no pretest data provided			Disagree (see above and Appendix 3)
RB		Curriel (1979)	No measure of early academic ability			Disagree (see above and Appendix 3)
RB		El Paso ISD (1992)	No measure of early academic ability			Disagree (see above and Appendix 3)
RB		Layden (1972)	Large pretest difference in both Spanish and English; 10 weeks			Reject because of short duration.
RB		Malherbe (1946)	Lacked information about initial comparability			Disagree (see above and Appendix 3)
RB		Matthews (1979)	Lacked information about initial comparability	U.S.-Not Spanish		Reject because of ELL group.
RB	Greene	Powers (1978)	No measure of early academic ability			Disagree (see above and Appendix 3)
RB & W		Stebbins et al (1977)	No measure of early academic ability			Disagree (see above and Appendix 3)
RB		Vasquez (1990)	No measure of early academic ability			Disagree (see above and Appendix 3)
RB&W		Zirkel (1972)	Large pretest differences in Hartford and Bridgeport. No bilingual instruction in New Britain, New London.			Disagree (see above and Appendix 3)

No Appropriate Comparison Group						
RB		Becker et al (1982)	Not an evaluation of bilingual programs			Disagree. See comments in text and Appendix 3
RB		Burkheimer et al (1989)	Compared actual performance to expected performance, no real control group			
		Carlisle & Beeman (2000)	Both groups were bilingual (80-20 vs. 20-80)			
RB	RB is error	de la Garza & Marcella (1985)	Compared Spanish dominant to English dominant; no pretest data			SC Error: Not in our review.
RB		Gersten (1985)	Study of Direct Instruction; No bilingual comparison group			
RB		Lampman (1973)	Mixed Spanish and English dominant children in the pretest analysis; only separate analysis for mean gains			See comments in text and Appendix 3
RB		McConnell (1980b)	Compared to a baseline group; No measure of initial comparability			
RB		Medina & Escamilla (1992)	Compared Vietnamese TBE to Hispanic Maintenance Bilingual; no reading outcomes			Agree. See comments in text and Appendix 3
RB		Moore & Parr (1978)	Mixed Spanish and English dominant children; also late pretests for grade 1 and 2			
RB		Prewitt-Diaz (1979)	17 weeks; initial group difference (control group had been in the US for 3 yrs; exp group just arrived from Puerto Rico); large pretest difference			Rejected because of short duration.
		Thomas & Collier (1997)	No control groups			
		Thomas & Collier (2002)	Separate studies in Oregon and Florida lacked control groups			
Brief Studies						
RB		Balasubramonian et al (1973)	4 months			Agree

Unavailable						
RB		Ciriza (1990)				SC error: is available
RB		Educational Operations Concepts (1991a & b)		U.S.: -Not Spanish		SC error: is available. Reject because of ELL group.
RB & W		McSpadden (1979, 1980)		U.S.: -Not Spanish		Reject because of ELL group.
RB & W		Pena-Hughes & Solis (1980)	Compared paired bilingual and transitional bilingual programs			Data tables are available; text is not.
RB		Teschner (1990)				SC error: is available
Slavin and Cheung Overlooked						
Slavin and Cheung overlooked Webb, Clerc, and Gavito which is in Rossell and Baker.						SC Error: Overlooked study

* RB=Rossell & Baker, 1996

W=Willig, 1985

Appendix 5

A Cohen's d Comparison of Slavin and Cheung's Effect Sizes to Rossell and Kuder's Effect Sizes for Reading

Study	Intervention description and Design	Duration	S&C N	R&K N exp	R&K N control	Grade	Sample Characteristics	Evidence of Initial Equality	Posttest	S&C Effect Size	S&C Mean ES	S&C 95% C. I.		R&K ES	R&K Mean ES	R&K 95% C.I.		R&K Stat Sig	
												lower limit	upper limit			lower limit	upper limit		
S&C Cohen's d: using R&K mean ES where S&C give none											0.34	0.26	0.43						Y
S&C Cohen's d: excluding the studies where S&C give no ES (shaded cells)											0.57	0.45	0.73						Y
R&K Spanish Elementary														0.14	0.03	0.26		Y	
Studies of Paired Bilingual Education																			
Plante (1976)	-Paired bilingual -Random assignment	2 yrs	55	31	22	1-2, 2-3	Spanish-dominant Puerto Rican students in New Haven, CT	Well matched on Spanish oral vocabulary but C>E in English pretest	English Inter-American Series										
										2nd grade	+0.62	+0.43	-0.12	0.98	0.62	0.43	-0.12	0.98	N
										3rd grade	+0.24								
Huzar (1973)	-Paired bilingual -Random assignment	2 & 3 yrs	160	84	76	1-2, 1-3	Disadvantaged Puerto Rican students in Perth Amboy, NJ	Well matched on IQ, SES, and initial achievement	English Inter-American Series										
										2nd grade	+0.01	+0.35	0.04	0.66	0.01	0.35	0.04	0.66	Y
										3rd grade	+0.68								
Campeau et al. (1975)-Corpus Christi	-Paired bilingual -Matched control	2 yrs	171	125	46	K-1	Spanish dominant students in Corpus Christi, Texas	Matched on English and Spanish pretests	English Inter-American Series	+0.45	+0.45	0.11	0.79	0.45	0.45	0.11	0.79	Y	
Campeau et al (1975)--Houston	-Paired bilingual -Matched control	3 yrs	206	461	151	K-2	Spanish dominant students in Houston, TX	Matched on language, SES, and academic achievement	English Inter-American Series	+1.00	+1.00	0.68	1.32	not scientific and/or insufficient info.					

J. R. Maldonado (1977)	-Paired bilingual -Matched control	5 yrs	126	47	79	1-5	Spanish dominant students in six elementary school in Corpus Christi, TX	Matched on SES and number of years in schools	English (SRAAS)									
									2nd	E=C	0.12	-0.24	0.48	0.15	0.12	-0.24	0.48	N
									3rd	E=C				0.23				
									4th	E=C				0.08				
									5th	E=C				0.04				
Alvarez (1975)	-Paired bilingual -Matched control	2 yrs	147	90	57	2	Spanish dominant children in two schools in Austin Texas	Matched on SES and initial language proficiency	CA Achievement Tests									
									vocab	E=C	-0.05	-0.38	0.28	0.12	-0.05	-0.38	0.28	N
									comp	E=C				-				
Cohen (1975)	-Paired bilingual -Matched control	2-3 yrs	90	45	45	K-1, 1-2, 1-3	Spanish dominant students in Redwood city, CA	Matched on SES and initial language proficiency	English Inter-American Series									
									Cohort 1	E=C	-0.14	-0.55	0.27	-	-0.14	-0.55	0.27	N
									Cohort 2	E=C				0.08				
									Cohort 3	none				-				
Campeau et al (1975)--Kingsville, TX	-Paired bilingual -Matched control	1 yr	89	48	41	K	Spanish dominant students in Kingsville, TX	Matched on SES and ethnic mix	English Inter-American Series	E>C	0.42	0.00	0.84	not scientific and/or insufficient info.				
Campeau et al (1975)--Santa Fe	-Paired bilingual -Matched control	1 yr	77	53	24	1	Hispanic students in Sante Fe, New Mexico	Pretests, E>C	English MAT	+0.28	+0.28	-0.20	0.76	not scientific and/or insufficient info.				
Studies of One-Year Transitional Bilingual Education																		
J. A. Maldonado (1994)	-Bilingual-1-year transition -Random assignment	3 yrs	20	10	10	2-4, 3-5	Spanish dominant special education students in Houston TX	Well matched on disability, language proficiency, & family background	English CTBS	+2.21	+2.21	1.10	3.32	2.22	2.22	0.64	3.79	Y

Campeau et al (1975)-- Alice, TX	-Bilingual-1-year transition -Matched control	2 yrs	125	106	19	K-1	Spanish dominant students in Alice ISD, Texas	Similar on English pretests but E>C on Spanish pretest	English Inter-American Series	+1.06	+1.06	0.55	1.57	not scientific and/or insufficient info.				
Studies of Two-Year Transitional Bilingual Education																		
Ramirez et al (1991)	-Bilingual-1-year transition -Matched control	4 yrs	varies	197	191	K-3	Spanish dominant LEP students	Fairly well matched on SES and home backgrounds.	English CTBS									
									3rd grade	Early=Imm	0.02	-0.18	0.22	0.02	0.02	-0.18	0.22	N
Studies of Bilingual Education (Unspecified)																		
Saldate et al (1985)	-Unspecified -Matched control	3 yrs	38	19	19	1-3	Spanish dominant students in Douglas, AZ	Well matched on pretests	English tests									
									MAT (2nd grade)	-0.29	+0.59	-0.06	1.24	-	0.59	-0.06	1.24	N
									WRAT (3rd grade)	+1.47				0.29				
Studies Involving Languages Other Than Spanish																		
Morgan (1971)	-Paired bilingual -Matched control	1 yr	193	93	100	1	French dominant students in Lafayette Diocese Catholic Schools of Louisiana	Well matched on initial mental ability and MRT pretests	English Stanford									
									Word Reading	+0.38	+0.26	-0.02	0.54	0.38	0.27	-0.14	0.68	N
									Paragraph meaning	+0.28				0.28				
									Vocab.	+0.19				0.19				
Word Study Skills	+0.23	0.23																
Bacon et al (1982)	-Paired bilingual -Matched control	4 & 5 yrs	53	35	18	1-5	Cherokee Indian students in Oklahoma	Well matched on control variables such as IQ and first language except for GPA & father's education, C>E	English SRA Reading									
									Cohort 1 (5 yrs)	+0.73	+0.70	0.12	1.28	0.73	0.70	0.01	1.38	Y
									Cohort 2 (4 yrs)	+0.67				0.67				

Doebler & Mardis (1980)	-Paired bilingual -Matched control	1 yr	63	26	37	2	Choctaw students in MS	Well matched on their initial English proficiency	English MAT	+0.15	+0.15	-0.35	0.65	0.15	0.15	-0.62	0.92	N	
Secondary Studies																			
Covey (1973)	-Paired bilingual -Random assignment	1 yr	200	86	87	9	Spanish dominant students	Well matched on pretests	English Stanford Diagnostic Reading	+0.82	+0.82	0.51	1.13	0.82	0.82	0.38	1.26	Y	
Kaufman (1968)	-Paired bilingual -Random assignment	1 & 2 yrs	139	51	44	7	Spanish dominant students in New York City	Initial CIA vocab and comprehension scores, language and non-language IQ, age, and Hoffman bilingual schedule scores were used as covariates	2-yr school										
									Word Meaning	E=C	0.23	-0.10	0.57	0.30	0.23	-0.32	0.78	N	
									Paragraph Meaning	E=C				0.11					
									1 yr school										
									Word Meaning	E>C				0.04					
Paragraph Meaning	E>C	0.48																	

* Shaded cells in N column denote large discrepancies in Ns.

** Shaded cells in Effect Size column denotes Rossell & Kuder effect size inserted.

Appendix 6

A Hedge's g Comparison of Slavin and Cheung's Effect Sizes to Rossell and Kuder's Effect Sizes for Reading

Study	Intervention description and Design	Duration	S&C N*	R&K N exp	R&K N control	Grade	Sample Characteristics	Evidence of Initial Equality	Posttest	S&C Effect Size	S&C Mean ES	S&C 95% C. I.		R&K ES	R&K Mean ES	R&K 95% C.I.		R&K Stat Sig.		
												lower limit	upper limit			lower limit	upper limit			
S&C Hedge's g (including arbitrary zeros)												0.25	0.17	0.34					Y	
R&K Spanish Elementary															0.10	-0.01	0.22	N		
Studies of Paired Bilingual Education																				
Plante (1976)	-Paired bilingual -Random assignment	2 yrs	55	31	22	1-2, 2-3	Spanish-dominant Puerto Rican students in New Haven, CT	Well matched on Spanish oral vocabulary but C>E in English pretest	English Inter-American Series											
										2nd grade	0.78	0.5	-0.05	1.05	0.80	0.51	-0.04	1.06	N	
										3rd grade	0.26				0.27					
Huzar (1973)	-Paired bilingual -Random assignment	2 & 3 yrs	160	84	76	1-2, 1-3	Disadvantaged Puerto Rican students in Perth Amboy, NJ	Well matched on IQ, SES, and initial achievement	English Inter-American Series											
										2nd grade	+0.01	0.16	-0.15	0.47	0.01	0.16	-0.15	0.47	N	
										3rd grade	0.31				0.31					
Campeau et al. (1975)-Corpus Christi	-Paired bilingual -Matched control	2 yrs	171	125	46	K-1	Spanish dominant students in Corpus Christi, Texas	Matched on English and Spanish pretests	English Inter-American Series	+0.45	+0.45	0.11	0.79	0.45	0.45	0.11	0.79	Y		
Campeau et al (1975)--Houston	-Paired bilingual -Matched control	3 yrs	206	461	151	K-2	Spanish dominant students in Houston, TX	Matched on language, SES, and academic achievement	English Inter-American Series	0.54	0.54	0.23	0.85	not scientific and/or insufficient info.						

J. R. Maldonado (1977)	-Paired bilingual -Matched control	5 yrs	126	47	79	1-5	Spanish dominant students in six elementary school in Corpus Christi, TX	Matched on SES and number of years in schools	English (SRAAS)									
									2nd	0	0.00	-0.36	0.36	0.15	0.12	-0.24	0.49	N
									3rd	0				0.23				
									4th	0				0.08				
5th	0	0.04																
Alvarez (1975)	-Paired bilingual -Matched control	2 yrs	147	90	57	2	Spanish dominant children in two schools in Austin Texas	Matched on SES and initial language proficiency	CA Achievement Tests									
									vocab	0.12	-0.05	-0.38	0.28	0.12	-0.05	-0.38	0.28	N
									comp	-0.23				-				
Cohen (1975)	-Paired bilingual -Matched control	2-3 yrs	90	45	45	K-1, 1-2, 1-3	Spanish dominant students in Redwood city, CA	Matched on SES and initial language proficiency	English Inter-American Series									
									Cohort 1	0	0.00	-0.41	0.41	-	-0.14	-0.55	0.27	N
									Cohort 2	0				0.08				
									Cohort 3	none				-				
Campeau et al (1975)--Kingsville, TX	-Paired bilingual -Matched control	1 yr	89	48	41	K	Spanish dominant students in Kingsville, TX	Matched on SES and ethnic mix	English Inter-American Series	0.42	0.42	0.00	0.84	not scientific and/or insufficient info.				
Campeau et al (1975)--Santa Fe	-Paired bilingual -Matched control	1 yr	77	53	24	1	Hispanic students in Sante Fe, New Mexico	Pretests, E>C	English MAT	0.03	0.03	-0.45	0.51	not scientific and/or insufficient info.				
Studies of One-Year Transitional Bilingual Education																		
J. A. Maldonado (1994)	-Bilingual-1-year transition -Random assignment	3 yrs	20	10	10	2-4, 3-5	Spanish dominant special education students in Houston TX	Well matched on disability, language proficiency, & family background	English CTBS	1.66	1.66	0.64	2.68	1.73	1.73	0.28	3.18	Y

Campeau et al (1975)--Alice, TX	-Bilingual-1-year transition -Matched control	2 yrs	125	106	19	K-1	Spanish dominant students in Alice ISD, Texas	Similar on English pretests but E>C on Spanish pretest	English Inter-American Series	0.49	0.49	0.00	0.98	not scientific and/or insufficient info.				
Studies of Two-Year Transitional Bilingual Education																		
Ramirez et al (1991)	-Bilingual-1-year transition -Matched control	4 yrs	varies	197	191	K-3	Spanish dominant LEP students	Fairly well matched on SES and home backgrounds.	English CTBS									
									3rd grade	0	0.00	-0.20	0.20	0.02	0.02	-0.18	0.22	N
Studies of Bilingual Education (Unspecified)																		
Saldate et al (1985)	-Unspecified -Matched control	3 yrs	38	19	19	1-3	Spanish dominant students in Douglas, AZ	Well matched on pretests	English tests									
									MAT (2nd grade)	-0.28	0.14	-0.50	0.78	-	0.14	-0.49	0.78	N
									WRAT (3rd grade)	0.89				0.29				
Studies Involving Languages Other Than Spanish																		
Morgan (1971)	-Paired bilingual -Matched control	1 yr	193	93	100	1	French dominant students in Lafayette Diocese Catholic Schools of Louisiana	Well matched on initial mental ability and MRT pretests	English Stanford									
									Word Reading	+0.38	+0.26	-0.02	0.54	0.38	0.26	-0.15	0.67	N
									Paragraph meaning	0.26				0.28				
									Vocab.	+0.19				0.19				
Word Study Skills	+0.23	0.23																
Bacon et al (1982)	-Paired bilingual -Matched control	4 & 5 yrs	53	35	18	1-5	Cherokee Indian students in Oklahoma	Well matched on control variables such as IQ and first language except for GPA & father's education, C>E	English SRA Reading									
									Cohort 1 (5 yrs)	+0.73	0.69	0.11	1.27	0.73	0.70	0.01	1.38	Y
									Cohort 2 (4 yrs)	0.68				0.67				

Doebler & Mardis (1980)	-Paired bilingual -Matched control	1 yr	63	26	37	2	Choctaw students in MS	Well matched on their initial English proficiency	English MAT	+0.15	+0.15	-0.35	0.65	0.15	0.15	-0.62	0.92	N		
Secondary Studies																				
Covey (1973)	-Paired bilingual -Random assignment	1 yr	200	86	87	9	Spanish dominant students	Well matched on pretests	English Stanford Diagnostic Reading	0.72	0.72	0.41	1.03	0.73	0.73	0.29	1.16	Y		
Kaufman (1968)	-Paired bilingual -Random assignment	1 & 2 yrs	139	51	44	7	Spanish dominant students in New York City	Initial CIA vocab and comprehension scores, language and non-language IQ, age, and Hoffman bilingual schedule scores were used as covariates	2-yr school											
									Word Meaning	0.23				0.23						
									Paragraph Meaning											
									1 yr school		0.23	-0.10	0.56		0.23	-0.32	0.78	N		
									Word Meaning	0.23				0.23						
Paragraph Meaning																				

* Shaded cells in N column denote large discrepancies in Ns.

** Shaded cells in Effect Size column denotes arbitrary assignment of 0 to effect size.

Appendix 7

**Appendix B from
Lipsey, Mark W. and Wilson, David B. 2001.
Practical Meta-Analysis. Newbury Park, Calif.: Sage Publications.**

can be imputed using Formula 5.

$$a = N \left(p_{r1} p_{c1} + \sqrt{\frac{\chi^2 p_{r1} p_{c1} (1 - p_{r1}) (1 - p_{c1})}{N}} \right)$$

$$a = 60 \left(0.5(0.117) + \sqrt{\frac{1.456(0.5)(0.117)(1 - 0.5)(1 - 0.117)}{60}} \right) = 5$$

Imputation of Odds-Ratio from Continuous Data

It may occur that a subset of studies eligible for inclusion in a meta-analysis of odds-ratios use a continuous dependent measure to contrast the groups. For example, many studies of diagnostic tests report data in a dichotomous form as the means on the test for the group with and that without the condition being diagnosed. Hasselblad and Hedges (1995) have shown how the standardized mean difference effect size can be converted into an odds-ratio (and vice versa) for meta-analysis. Using Formula 6 from Table B12, a standardized mean difference effect size computed with any of the formulas from Table B10 can be converted into an odds-ratio (or logged odds-ratio) equivalent. Suppose, for example, that a standardized mean difference effect size computed on means and standard deviations using Formula 1 in Table B10 has a value of .32. Formula 6 in Table B12 then yields the odds-ratio as follows.

$$ES_{OR} = e^{\left(\frac{\pi ES_{sm}}{\sqrt{3}}\right)} = e^{\left(\frac{3.14(.32)}{\sqrt{3}}\right)} = e^{.58} = 1.79$$

Table B10

Useful formulas for calculating ES_{sm} from a range of statistical data

Formula	Data needed and definition of terms
Direct calculation formula for ES_{sm}	
(1) $ES_{sm} = \frac{\bar{X}_1 - \bar{X}_2}{s_{pooled}}$ $s_{pooled} = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$	Means (\bar{X}), standard deviations (s), and sample sizes (n) for each group.
Algebraically equivalent formulas for ES_{sm}	
(2) $ES_{sm} = t \sqrt{\frac{n_1 + n_2}{n_1 n_2}}$	Independent t -test (t) and sample sizes (n) for each group.
(3) $ES_{sm} = \frac{2t}{\sqrt{N}}$	Independent t -test (t) and total sample size (N). Assumes $n_1 = n_2$.

Table B10
continued

Formula	Data needed and definition of terms
Direct calculation formula for ES_{tm}	
(4) $ ES_{tm} = \sqrt{\frac{F(n_1 + n_2)}{n_1 n_2}}$	F-ratio (F) from a one-way ANOVA and sample sizes (n) for each group.
(5) $ ES_{tm} = 2\sqrt{\frac{F}{N}}$	F-ratio (F) from a one-way ANOVA and total sample size (N). Assumes $n_1 = n_2$.
Exact probabilities levels for a t -value	
(6) $t = IDF(p, df)$	Determine the t -value for a p -value and df from Table B13 or an inverse distribution function (IDF) in a spreadsheet or statistical software program. Use result and Formula 3 to obtain effect size. Note: $t^2 = F$.
Calculation of means and standard deviations from a grouped frequency distribution	
(7) $\bar{X} = \frac{\sum x_i f_i}{\sum f_i}$	Frequency counts (f) for each level (i) of a variable (x).
(8) $s = \sqrt{\frac{(\sum f_i)(\sum x_i^2 f_i) - (\sum x_i f_i)^2}{(\sum f_i)^2}}$	Frequency counts (f) for each level (i) of a variable (x).
Approximations based on continuous data	
(9) $ES_{tm} = \frac{2r}{\sqrt{1-r^2}}$	Correlation (r) between group membership and dependent variable (assumes equal n in groups).
(10) $ES_{tm} = \frac{r}{\sqrt{(1-r^2)(p(1-p))}}$	Correlation (r) between group membership and dependent variable, and the proportion (p) of the total sample in one of the two groups.
Estimates $\bar{X}_1 - \bar{X}_2$ of (numerator of ES_{tm})	
(11) $\bar{X}_1 - \bar{X}_2 \approx \Delta_1 - \Delta_2$	Mean gain score (Δ) for each group.
(12) $\bar{X}_1 - \bar{X}_2 \approx \bar{X}_{1, \text{adjusted}} - \bar{X}_{2, \text{adjusted}}$	Covariate or regression adjusted means ($\bar{X}_{\text{adjusted}}$) for each group.
(13) $\bar{X}_1 - \bar{X}_2 \approx B$	Unstandardized regression coefficient (B) for group membership.
Estimates of s_{pooled} (denominator of ES_{tm})	
(14) $s_{\text{pooled}} = \sqrt{\frac{s^2(N-1) - \frac{(\bar{X}_1^2 + \bar{X}_2^2 - 2\bar{X}_1\bar{X}_2)(n_1 n_2)}{n_1 + n_2}}{N-1}}$	Full-sample standard deviation (s), group means (\bar{X}), group sample sizes (n), and total sample size (N).

Table B10
continued

Formula	Data needed and definition of terms
(15) $s_{\text{pooled}} = \frac{\bar{X}_1 - \bar{X}_2}{t \sqrt{\frac{n_1 + n_2}{n_1 n_2}}}$	Means (\bar{X}) and sample sizes (n) for each group, and associated t -value (t).
(16) $s = se \sqrt{n - 1}$	Standard error of the mean (se) and sample size (n) for any group.
(17) $s_{\text{pooled}} = \sqrt{\frac{MS_b}{F_{\text{oneway}}}}$ $MS_b = \frac{\sum n_j \bar{X}_j^2 - \frac{(\sum n_j \bar{X}_j)^2}{\sum n_j}}{k - 1}$	F -ratio (F) from a one-way ANOVA with k groups and the mean (\bar{X}) and sample size (n) for each group (j).
(18) $s_{\text{pooled}} = \sqrt{\frac{SS_B + SS_{AB} + SS_w}{df_B + df_{AB} + df_w}}$	The sums-of-squares (SS) and degrees of freedom (df) from a factorial (two-way) ANOVA. Subscripts indicate factors (A and B) and within groups or residual term (w).
(19) $s_{\text{pooled}} = \sqrt{\frac{(MS_{\text{error}})(df_{\text{error}} - 1)}{(1 - r^2)(df_{\text{error}} - 2)}}$	The mean-square error (MS_{error}) and associated degrees of freedom (df), and correlation (r) between covariate and dependent variable from a one-way ANCOVA.
(20) $s_{\text{pooled}} = \frac{s_{\text{gain}}}{\sqrt{2(1 - r)}}$	Standard deviation of the gain scores (s_{gain}) and the correlation (r) between time-one and time-two scores.
Approximations based on dichotomous data	
(21) $ES_{tm} = \text{probit}(p_1) - \text{probit}(p_2)$	Probit transformation (Table B15) of the proportion (p) of successes for each group.
(22) $ES_{tm} = \text{arcsine}(p_1) - \text{arcsine}(p_2)$	Arcsine transformation (Table B14) of the proportion (p) of successes for each group.
(23) $ ES_{tm} = 2 \sqrt{\frac{\chi^2}{N - \chi^2}}$	Chi-square (χ^2) with $df = 1$ and total sample size (N).
(24) $ES_{tm} = \frac{2r}{\sqrt{1 - r^2}}$	Phi-coefficient (r).