

# References

Achilles, C. M., Finn, J. D., & Bain, H. P. (1997). Using class size to reduce the equity gap. *Education Leadership*, 55(4), 40–43.

Anderson, J., Hollinger, D., & Conaty, J. (1993). Re-examining the relationship between school poverty and student achievement. *ERS Spectrum*, 11(2), 21–31.

Angrist, J. D., Imbens, G. W., & Rubin, D. B. (1996). Identification of causal effects using instrumental variables (with commentary). *Journal of the American Statistical Association*, 91, 444–472.

Angrist, J. D., & Krueger, A. B. (1991). Does compulsory school attendance affect schooling and earnings? *Quarterly Journal of Economics*, 106(4), 979–1014.

Angrist, J. D., & Krueger, A. B. (1995). Split-sample instrumental variables estimates of the returns to schooling. *Journal of Business and Economic Statistics*, 13(2), 225–235.

Angrist, J. D., & Krueger, A. B. (2001). Instrumental variables and the search for identification: From supply and demand to natural experiments. *Journal of Economic Perspectives*, 15(4), 69–85.

- Angrist, J.D., & Lavy, V. (1999). Using Maimonides' rule to estimate the effects of class size on academic achievement. *Quarterly Journal of Economics*, 114(2), 533–576.
- Baker, D., & Leary, R. (1995). Letting girls speak out about science. *Journal of Research in Science Teaching*, 32, 3–7.
- Barnow, B., Cain, G., & Goldberger, A. (1980). Issues in the analysis of selectivity bias. In E. Stromsdorfer & G. Farkas (Eds.), *Evaluation studies* (Vol. 5, pp. 43–59). Beverly Hills, CA: Sage.
- Benbow, C., & Minor, L. (1986). Mathematically talented males and females and achievement in the high school sciences. *American Educational Research Journal*, 23, 425–439.
- Beutel, A. M., & Marini, M. M. (1995). Gender and values. *American Sociological Review*, 60(3), 436–448.
- Bidwell, C. (1965). The school as a formal organization. In J. G. March (Ed.), *Handbook of organizations* (pp. 972–1022). Chicago: Rand McNally.
- Bidwell, C. (2000). School as context and construction: A social psychological approach to the study of schooling. In M. T. Hallinan (Ed.), *Handbook of the sociology of education* (pp. 13–37). New York: Kluwer Academic/Plenum.
- Bidwell, C., Frank, K. A., & Quiroz, P. (1997). Teacher types, workplace controls, and the organization of schools. *Sociology of Education*, 70(4), 285–307.
- Bifulco, R., & Ladd, H. F. (2006). The impacts of charter schools on student achievement: Evidence from North Carolina. *Education Finance and Policy*, 1(1), 50–99.
- Bloch, M. (2004). A discourse that disciplines, governs, and regulates: The National Research Council's report on scientific research in education. *Qualitative Inquiry*, 10(1), 96–110.
- Bloom, H. S. (1984). Accounting for no-shows in experimental evaluation designs. *Evaluation Review*, 8, 225–246.

Bound, J., Jaeger, D. A., & Baker, R. M. (1995). Problems with instrumental variables estimation when the correlation between the instruments and the exogenous explanatory variable is weak. *Journal of the American Statistical Association*, *90*, 443–450.

Bryk, A. S., Lee, V. E., & Holland, P. B. (1993). *Catholic schools and the common good*. Cambridge, MA: Harvard University Press.

Bryk, A. S., & Raudenbush, S. W. (2002). *Hierarchical linear models* (2nd ed.). Thousand Oaks, CA: Sage Publications.

Burkham, D., Lee, V., & Smerdon, B. (1997). Gender and science learning early in high school: Subject matter and laboratory experiences. *American Educational Research Journal*, *34*, 297–332.

Burstein, L. (Ed.). (1993). *The IEA study of mathematics III: Student growth and classroom processes* (Vol. 3). Oxford, UK: Pergamon Press.

Byrnes, D. A. (1989). Attitudes of students, parents, and educators toward repeating a grade. In L. A. Shepard & M. L. Smith (Eds.), *Flunking grades: Research and policies on retention* (pp. 108–131). Philadelphia: Falmer Press.

Campbell, D. T. (1969). Reforms as experiments. *American Psychologist*, *25*, 409–429.

Campbell, D. T., & Stanley, J. C. (1963). Experimental and quasi-experimental designs for research on teaching. In N. L. Gage (Ed.), *Handbook of research on teaching*. Chicago: Rand McNally.

Cochran, W. G., & Cox, G. M. (1950). *Experimental design* (2nd ed.). New York: Wiley.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.

Coleman, J. (1961). *The adolescent society*. New York: Free Press.

Coleman, J., & Hoffer, T. (1987). *Public and private high schools: The impact of communities*. New York: Basic Books.

- Cook, T.D. (2002). Randomized experiments in educational policy research: A critical examination of the reasons the educational evaluation community has offered for not doing them. *Educational Evaluation and Policy Analysis*, 24(3), 175–199.
- Cook, T.D. (2007). Randomized experiments in education: Assessing the objections to doing them. *Economics of Innovation and New Technology*, 16(2), 31–49.
- Cook, T.D. (in press). “Waiting for life to arrive”: A history of the regression-discontinuity design in psychology, statistics and economics. *Journal of Econometrics*.
- Correll, S. (2001). Gender and the career choice process: The role of biased self-assessments. *American Journal of Sociology*, 106, 1697–1730.
- Cox, D. R. (1958a). *Planning of experiments*. New York: Wiley.
- Cox, D. R. (1958b). The interpretation of the effects of non-additivity in the Latin square. *Biometrika*, 45, 69–73.
- Crosnoe, R., Cavanagh, S., & Elder Jr., G.H. (2003). Adolescent friendships as academic resources: The intersection of friendship, race, and school disadvantage. *Sociological Perspectives*, 46, 331–352.
- Cuddeback, G., Wilson, E., Orme, J. G., & Combs-Orme, T. (2004). Detecting and correcting sample selection bias. *Journal of Social Service Research*, 30(3), 19–33.
- Cuijpers, P. (2003). Examining the effects of prevention programs on the incidence of new cases of mental disorders: The lack of statistical power. *American Journal of Psychiatry*, 160(8), 1385–1391.
- Currie, J. (2003, June). *When do we really know what we think we know? Determining causality*. Invited paper presented at Work, Family, Health and Well-Being conference, NICHD Administration for Children and Families, Washington, DC.
- Currie, J., & Thomas, D. (1995). Does Head Start make a difference? *American Economic Review*, 85, 341–364.

Currie, J., & Thomas, D. (1999). Does Head Start help Hispanic children? *Journal of Public Economics*, 74, 235–262.

Darling-Hammond, L. (1995). Inequality and access to knowledge. In J. A. Banks (Ed.), *The handbook of research on multicultural education* (pp. 465–483). New York: Macmillan.

Dignam, J. (2003, November). *From efficacy to effectiveness: Translating randomized controlled trial findings into treatment standards*. Paper presented at the invitational conference Conceptualizing Scale-Up: Multidisciplinary Perspectives, Data Research and Development Center, Washington, DC.

Dreeben, R., & Gamoran, A. (1986). Race, instruction, and learning. *American Sociological Review*, 51(5), 660–669.

Dryler, H. (1998). Parental role models, gender, and educational choice. *British Journal of Sociology*, 49, 375–398.

Eccles, J., Jacobs, J., & Harold, R. (1990). Gender role stereotypes, expectancy effects, and parents' socialization of gender differences. *Journal of Social Issues*, 46, 183–202.

Education Sciences Reform Act of 2002, Pub. L. No. 107-279. Retrieved September 3, 2004, from <http://www.ed.gov/legislation/EdSciencesRef/>

Eisenhart, M., & Towne, L. (2003). Contestation and change in national policy on scientifically based research. *Educational Researcher*, 32(7), 3138.

Epstein, J. L. (1983). The influence of friends on achievement and affective outcomes. In J. L. Epstein & N. Karweit (Eds.), *Friends in school: Patterns of selection and influence in secondary schools* (pp. 177–200). New York: Academic Press.

Erikson, E. H. (1968). *Identity, youth, and crisis*. New York: Norton.

Felmlee, D. H. (1999). Social norms in same- and cross-gender friendships. *Social Psychology Quarterly*, 62(1), 53–67.

- Feuer, M., Towne, L., & Shavelson R. J. (2002). Scientific research and education. *Educational Researcher*, 31(8), 4–14.
- Finn, J.D., & Achilles, C. (1990). Answers and questions about class size: A statewide experiment. *American Educational Research Journal*, 27(3), 557–577.
- Finn, J.D., & Achilles, C. (1999). Tennessee's class size study: Findings, implications, and misconceptions. *Educational Evaluation and Policy Analysis*, 21(2), 97–109.
- Firestone, W. A. (1985). The study of loose coupling: Problems, progress, and prospects. In A.C. Kerckhoff (Ed.), *Research in sociology of education and socialization* (Vol. 5, pp. 3–30). Greenwich, CT: JAI Press.
- Fish, S. (2003). Truth but no consequences: Why philosophy doesn't matter. *Critical Inquiry*, 29(3), 389–417.
- Fisher, R. A. (1935). *The design of experiments*. Edinburgh, UK: Oliver & Boyd.
- Foster, E. M., & Fang, G. Y. (2004). Alternative methods for handling attrition: An illustration using data from the Fast Track evaluation. *Evaluation Review*, 28(5), 434–464.
- Frangakis, C.E., & Rubin, D.B. (1999). Addressing complications of intention-to-treat analysis in the combined presence of all-or-none treatment-noncomplicance and subsequent missing outcomes. *Biometrika*, 86, 365–379.
- Gamoran, A. (1989). Measuring curriculum differentiation. *American Journal of Education*, 97, 129–143.
- Gamoran, A., & Berends M. (1988). The effects of stratification in secondary schools: Synthesis of survey and ethnographic research. *Review of Educational Research*, 57, 415–435.
- Gee, J.P. (2005). It's theories all the way down: A response to *Scientific Research in Education*. *Teachers College Record*, 107(1), 10–18.

Giangreco, M.F., & Taylor, S.J. (2003). "Scientifically based research" and qualitative inquiry. *Research and Practice for Persons with Severe Disabilities*, 28(3), 133–137.

Giordano, P. (2003). Relationships in adolescence. *Annual Review of Sociology*, 29, 257–281.

Granic, I., & Dishion, T. (2003). Deviant talk in adolescent friendships: A step toward measuring a pathogenic attractor process. *Social Development*, 12(3), 314–334.

Greenwald, R., Hedges, L. V., & Laine, R. D. (1996). The effect of school resources on student achievement. *Review of Educational Research*, 66, 361–396.

Hahn, J., Todd, P., & Van der Klaauw. (1999). *Evaluating the effect of an antidiscrimination law using a regression-discontinuity design* (Brief No. 7131). Cambridge, MA: National Bureau of Economic Research.

Hahn, J., Todd, P., & Van der Klaauw. (2001). Identification and estimation of treatment effects with a regression-discontinuity design. *Econometrica*, 69(1), 201–209.

Hallinan, S., & Williams, R. (1990). Students' characteristics and the peer-influence process. *Sociology of Education*, 63, 122–132.

Halpern, S.D., Karlawish, J.H., & Berlin, J.A. (2002). The continuing unethical conduct of underpowered clinical trials. *Journal of the American Medical Association*, 288, 358–362.

Hanushek, E. A. (1986). The economics of schooling: Production and efficiency in public schools. *Journal of Economic Literature*, 24, 1141–1177.

Heckman J.J. (1976). The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models. *Annals of Economic and Social Measurement*, 5, 475–492.

Heckman, J.J. (1979). Sample selection bias as a specification error. *Econometrica*, 47(1), 153–161.

Hedberg, E. C., Santana, R., & Hedges, L. V. (2004, April). *The variance structure of academic achievement in America*. Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA.

Hedges, L. V. (2004). Examining the effects of school organization on student achievement. In M. Ross (Ed.), *Instructional performance consequences of high poverty schooling*. Washington, DC: U.S. Government Printing Office.

Hedges, L. V. (2006). *Fixed effects versus mixed effects models*. Manuscript in preparation.

Higginbotham, H. N., West, S. G., & Forsyth, D. R. (1988). *Psychotherapy and behavior change: Social, cultural, and methodological perspectives*. New York: Pergamon.

Hirano, K., Imbens, G., Rider, G., & Rubin, D. B. (2001). Combining panel data sets with attrition and refreshment samples. *Econometrica*, 69(6), 1645–1659.

Holland, P. W. (1986). Statistics and causal inference. *Journal of the American Statistics Association*, 81, 945–970.

Holland, P. W. (1988). Causal inference, path analysis, and recursive structural equations models. In C. Clogg (Ed.), *Sociological methodology* (pp. 449–484). Washington, DC: American Sociological Association.

Holland, P. W., & Rubin, D. B. (1983). On Lord's Paradox. In H. Wainer & S. Messick, (Eds.), *Principals of modern psychological measurement* (pp. 3–25). Hillsdale, NJ: Lawrence Erlbaum.

Holmes, C. T. (1989). Grade-level retention effects: A meta-analysis of research studies. In L. A. Shepard & M. L. Smith (Eds.), *Flunking grades: Research and policies on retention* (pp. 16–33). Philadelphia: Falmer.

Hong, G., & Raudenbush, S. W. (2005). Effects of kindergarten retention policy on children's cognitive growth in reading and mathematics. *Educational Evaluation and Policy Analysis*, 27(3), 205–224.

Horner, R. H., Carr, E. G., Halle, J., McGee, G., Odom, S., & Wolery, M. (2005). The use of single-subject research to identify evidence-based practice in special education. *Exceptional Children*, 71(2), 165–179.

Imbens, G., & Rubin, D. B. (1997). Bayesian inference for causal effects in randomized experiments with noncompliance. *Annals of Statistics*, 25, 305–327.

Ingersoll, R. (1993). Loosely coupled organizations revisited. *Research in the Sociology of Organizations*, 11, 81–112.

Jackson, G. B. (1975). The research evidence on the effect of grade retention. *Review of Educational Research*, 45(3), 613–635.

Jacob, B., & Lefgren, L. (2004). Remedial education and student achievement: A regression discontinuity analysis. *Review of Economics and Statistics*, 86(1), 226–244.

Jimerson, S. R. (2001). Meta-analysis of grade retention research: Implications for practice in the 21st century. *School Psychology Review*, 30, 313–330.

Jurs, S. G., & Glass, G. V. (1971). Experimental mortality. *Journal of Experimental Education*, 40, 62–66.

Kaestle, C. F. (1993). The awful reputation of educational research. *Educational Researcher*, 22, 23–31.

Kamil, M. L. (2004). The current state of quantitative research. *Reading Research Quarterly*, 39(1), 100–107.

Karweit, N. L. (1992). Retention policy. In M. Alkin (Ed.), *Encyclopedia of educational research* (pp. 114–118). New York: Macmillan.

Kellam, S. G., & Langevin, D. J. (2003). A framework for understanding “evidence” in prevention research and programs. *Prevention Science*, 4(3), 137–153.

Kellam, S. G., Ling, X., Merisca, R., Brown, C. H., & Ialongo, N. (1998). The effect and level of aggression in the first grade classroom

on the course and malleability of aggressive behavior into middle school. *Development and Psychopathology*, 10(2), 165–185.

Kellam, S. G., & Van Horn, Y. V. (1997). Life course development, community epidemiology, and preventive trials: A scientific structure for prevention research. *American Journal of Community Psychology*, 25(2), 177–187.

Kemphorne, O. (1952). *Design and analysis of experiments*. New York: Wiley.

Kilgore, S. B. (1991). The organizational context of tracking in schools. *American Journal of Sociology*, 56(2), 189–203.

Kilgore, S. B., & Pendleton, W. W. (1993). The organizational context of learning: A framework for understanding the acquisition of knowledge. *Sociology of Education*, 66(1), 63–87.

King, G., Honaker, J., Joseph, A., & Scheve, K. (2001). Analyzing incomplete political science data: An alternative algorithm for multiple imputation. *American Political Science Review*, 95, 49–69.

Kish L. (1987). *Statistical design for research*. New York: John Wiley and Sons.

Krei, M. S. (1998). Intensifying the barriers: The problem of inequitable teacher allocation in low-income urban schools. *Urban Education*, 33, 71–94.

Krueger, A. B. (1999). Experimental estimates of education production functions. *Quarterly Journal of Economics*, 114(2), 497–532.

Lagemann, E. C. (1999). An auspicious moment for education research? In E. C. Lagemann & L. S. Shulman (Eds.), *Issues in education research: Problems and possibilities* (pp. 3–16). San Francisco: Jossey-Bass.

Lagemann, E. C. (2000). *An elusive science: The troubling history of education research*. Chicago: University of Chicago Press.

Lagemann, E. C. (2005). Does history matter in education research? A brief for the humanities in the age of science. *Harvard Educational Review*, 75(1), 3–19.

- Langford, H., Loeb, S., & Wyckoff, J. (2002). Teacher sorting and the plight of urban schools: A descriptive analysis. *Educational Evaluation and Policy Analysis, 24*, 37–62.
- Lather, P. (2004). Scientific research in education: A critical perspective. *British Educational Research Journal, 30*(6), 759–772.
- Lee, J. D. (2002). More than ability: Gender and personal relationships influence science and technology development. *Sociology of Education, 75*, 349–373.
- Lee, V., & Bryk, A. (1986). The effects of single-sex secondary schools on students' achievement and attitudes. *Journal of Educational Psychology, 78*, 381–396.
- Lee, V. E., Smith, J. B., & Croninger, R. G. (1997). How high school organization influences the equitable distribution of learning in math and science. *Sociology of Education, 70*(2), 128–150.
- Leinhardt, G. (1980). Transition rooms: Promoting maturation or reducing education? *Journal of Education Psychology, 72*, 55–61.
- Levin, J. R. (2003). Random thoughts on the (in)credibility of educational-psychological intervention research. *Educational Psychologist, 39*(3), 173–184.
- Levin, J. R., & O'Donnell, A. M. (1999). What to do about educational research's credibility gaps? *Issues in Education, 5*(2), 177–229.
- Lin, D. Y., Psaty, B. M., & Kronmal, R. A. (1998). Assessing the sensitivity of regression results to unmeasured confounders in observational studies. *Biometrics, 54*, 948–963.
- Little, R. J., & Rubin, D. B. (2002). *Statistical analyses with missing data*. New York: John Wiley.
- Little, R. J., & Yau, L. H. Y. (1998). Statistical techniques for analyzing data from preventive trials: Treatment of no-shows using Rubin's causal model. *Psychological Methods, 3*, 147–159.
- Locke, J. (1975). *An essay concerning human understanding*. Oxford, UK: Clarendon Press. (Original work published in 1690)

- Matsueda, R., & Anderson, K. (1998). The dynamics of delinquent peers and delinquent behavior. *Criminology*, 36, 269–308.
- Mayer, R.E. (2003). Learning environments: The case for evidence-based practice and issue-driven research. *Educational Psychology Review*, 15(4), 359–366.
- McCarthy, B., Felmlee, D., & Haga, J. (2004). Girl friends are better: Gender, friends, and crime among school and street youth. *Criminology*, 42(4), 805–835.
- McDonald, S.-K., Keesler, V., Kauffman, N., & Schneider, B. (2006). Scaling-up exemplary interventions. *Educational Researcher*, 35(3), 15–24.
- McKnight, C. C., Crosswhite, F. J., Dossey, J. A., Kifer, E., Swafford, J. O., Travers, K. J., et al. (1987). *The underachieving curriculum: Assessing U.S. school mathematics from an international perspective*. Champaign, IL: Stipes Publishing Company.
- Morgan, S. L. (2001). Counterfactuals, causal effect heterogeneity, and the Catholic school effect on learning. *Sociology of Education*, 74, 341–374.
- Morrison, F. J., Griffith, E. M., & Alberts, D. M. (1997). Nature-nurture in the classroom: Entrance age, school readiness, and learning in children. *Developmental Psychology*, 33(2), 254–262.
- Moss, P. (2005). Toward “epistemic reflexivity” in educational research: A response to *Scientific Research in Education*. *Teachers College Record*, 107(1), 19–29.
- Nagaoka, J., & Roderick, M. (2004). *Ending social promotion: The effects of retention*. Chicago: Consortium on Chicago School Research.
- National Research Council. (2002). *Scientific research in education*. Washington, DC: National Academy Press.
- National Research Council. (2004a). *Advancing scientific research in education*. Washington, DC: National Academy Press.

National Research Council. (2004b). *Implementing randomized field trials in education: Report of a workshop*. Washington, DC: National Academies Press.

National Science Foundation. (2004). *Final report on the National Science Foundation mathematics education portfolio review*. Washington, DC: National Science Foundation.

Neyman, J. (1923). On the application of probability theory to agricultural experiments: Essay on principles (D. M. Dabrowska & T. P. Speed, Trans.). *Statistical Sciences*, 5, 472–480.

Neyman, J. (1935). Statistical problems in agricultural experimentation. *Journal of the Royal Statistical Society*, 2, 107–180.

No Child Left Behind Act of 2001, Pub. L. No. 107-11, 115 Stat. 1425 (2002).

Norman, G.R., & Streiner, D.L. (2004). *PDQ statistics* (3rd ed.). Hamilton, Ontario, Canada: B. C. Decker.

Nye, B., Konstantopoulos, S., & Hedges, L. V. (2000). Effects of small classes on academic achievement: The results of the Tennessee class size experiment. *American Educational Research Journal*, 37, 123–151.

Nye, B., Konstantopoulos, S., & Hedges, L. V. (2004). How large are teacher effects? *Educational Evaluation and Policy Analysis*, 26, 237–257.

Plummer, D. L., & Graziano, W. G. (1987). Impact of grade retention on the social development of elementary school children. *Developmental Psychology*, 23(2), 267–275.

Popkewitz, T. S. (2004). Is the National Research Council Committee's report on *Scientific Research in Education* scientific? On trusting the manifesto. *Qualitative Inquiry*, 10(1), 62–78.

Raudenbush, S. W. (1997). Statistical analysis and optimal design for cluster randomized trials. *Psychological Methods*, 2(2), 173–185.

- Raudenbush, S.W. (2005). Learning from attempts to improve schooling: The contribution of methodological diversity. *Educational Researcher*, 34(5), 25–31.
- Raudenbush, S.W., Fotiu, R.P., & Cheong, Y.F. (1998). Inequality of access to educational resources: A national report card for eighth-grade math. *Educational Evaluation and Policy Analysis*, 20(4), 253–267.
- Raudenbush, S.W., & Liu, X. (2000). Statistical power and optimal design for multisite randomized trials. *Psychological Methods*, 5(2), 199–213.
- Reynolds, A.J. (1992). Grade retention and school adjustment: An explanatory analysis. *Educational Evaluation and Policy Analysis*, 14(2), 101–121.
- Riegle-Crumb, C., Farkas, G., & Muller, C. (2006). The role of gender and friendship in advanced course-taking. *Sociology of Education*, 79(3), 206–228.
- Rivkin, S.G., Hanushek, E.A., & Kain, J.F. (2005). Teachers, schools, and academic achievement. *Econometrica*, 72(3), 417–458.
- Rosenbaum, P.R. (1986). Dropping out of high school in the United States: An observational study. *Journal of Educational Statistics*, 11(3), 207–224.
- Rosenbaum, P.R. (2002). *Observational studies* (2nd ed.). New York: Springer.
- Rosenbaum, P.R., & Rubin, D.B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1), 41–55.
- Rossi, J.S. (1990). Statistical power of psychological research: What have we gained in 20 years? *Journal of Consulting and Clinical Psychology*, 58, 646–656.
- Rover, D.T. (2005). Centered on education [Review of the book *Scientific research in education*]. *Journal of Engineering Education*, 94(1), 195–197.

Rubin, D.B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*, 66, 688–701.

Rubin, D.B. (1977). Assignment of treatment group on the basis of a covariate. *Journal of Educational Statistics*, 2, 1–26.

Rubin, D.B. (1978). Bayesian inference for causal effects: The role of randomization. *Annals of Statistics*, 6, 34–58.

Rubin, D.B. (1980). Discussion of “Randomization analysis of experimental data in the Fisher randomization test” by Basu. *Journal of the American Statistical Association*, 75, 591–593.

Rubin, D.B. (1986). Which ifs have causal answers? Discussion of “Statistics and causal inference” by Holland. *Journal of the American Statistical Association*, 81, 961–962.

Rubin, D.B. (1997). Estimating causal effects from large data sets using propensity scores. *Annals of Internal Medicine*, 127(8), 757–763.

Rubin, D.B. (1998). More powerful randomization-based  $p$ -values in double-blind trials with noncompliance (with discussion). *Statistics in Medicine*, 17, 371–389.

Rubin, D.B. (2006). *Matched sampling for causal effects*. New York: Cambridge University Press.

Schmidt, W.H., McKnight, C., Cogan, L.S., Jakwerth, P.M., & Houang, R.T. (1999). *Facing the consequences: Using TIMSS for a closer look at U.S. mathematics and science education*. Dordrecht/Boston/London: Kluwer.

Schmidt, W.H., McKnight, C.C., Houang, R.T., Wang, H., Wiley, D.E., Cogan, L.S., et al. (2001). *Why schools matter: A cross-national comparison of curriculum and learning*. New York: John Wiley & Sons.

Schneider, B., Kertcher, Z., & Offer, S. (2006). Global trends towards education and science: Tension and resistance. In J. Ozga, T. Seddon, & T. Popkewitz (Eds.), *Education research and*

*policy: Steering the knowledge-based economy* (pp. 200–215). New York: Routledge.

Schneider, B., & McDonald, S.-K. (Eds.). (2007a). *Scale-up in education: Vol. 1. Ideas in principle*. Lanham, MD: Rowman & Littlefield.

Schneider, B., & McDonald, S.-K. (Eds.). (2007b). *Scale-up in education: Vol. 2. Issues in practice*. Lanham, MD: Rowman & Littlefield.

Schneider, B., McDonald, S.-K., Brown, K. B., Schalliol, D., Makela, S., Yamaguchi, K., et al. (2006). *Evaluating the efficacy of the Center for Education at the National Academies: Report to the Center for Education*. Chicago: University of Chicago.

Schneider, B., & Stevenson, D. (1999). *The ambitious generation: America's teenagers, motivated but directionless*. New Haven, CT: Yale University Press.

Scientific research in education. (2002). Special issue of *Educational Researcher*, 31(8).

Scientific research in education. (2005). Special issue of *Teachers College Record*, 107(1).

Secretary of Education. (2005, January 25). Scientifically based evaluation methods (RIN 1890–ZA00). *Federal Register*, 70(15), 3586.

Seymour, E., & Hewitt, N. (1997). *Talking about leaving: Why undergraduates leave the sciences*. Boulder, CO: Westview Press.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton Mifflin.

Shapka, J. D., & Keating, D. P. (2003). Effects of a girls-only curriculum during adolescence: Performance, persistence, and engagement in mathematics and science. *American Educational Research Journal*, 40, 929–960.

Shavelson, R. J., & Berliner, D. C. (1988). Erosion of the education research infrastructure. *Educational Researcher*, 17(1), 9–12.

Shepard, L. A. (1989). A review of research on kindergarten retention. In L. A. Shepard & M. L. Smith (Eds.), *Flunking grades: Research and policies on retention* (pp. 64–78). London: Falmer Press.

Shu, X., & Marini, M. M. (1998). Gender-related change in occupational aspirations. *Sociology of Education*, 71(1), 43–67.

Slavin, R. E., & Madden, N. A. (2001). *One million children: Success for All*. Thousand Oaks, CA: Corwin.

Slavin, R. E., Madden, N. A., Dolan, L. J., Wasik, B. A., Ross, S. M., & Smith, L. M. (1994). “Whenever and wherever we choose”: The replication of Success for All. *Phi Delta Kappan*, 75, 639–647.

Slavin, R. E., Madden, N. A., Karweit, N. L., Dolan, L., & Wasik, B. A. (1992). *Success for All: A relentless approach to prevention and early intervention in elementary schools*. Arlington, VA: Educational Research Service.

Smith, M. L., & Shepard, L. A. (1988). Kindergarten readiness and retention: A qualitative study of teachers’ beliefs and practices. *American Educational Research Journal*, 25(3), 307–333.

Sorenson, A. B. (1970). Organizational differentiation of students and educational opportunity. *Sociology of Education*, 43(4), 355–376.

Sorenson, A. B. (1987). The organization and differentiation of students in schools as an opportunity structure. In M. T. Hallinan (Ed.), *The social organization of schools*. New York: Plenum Press.

South, S., & Haynie, D. (2004). Friendship networks of mobile adolescents. *Social Forces*, 83, 315–350.

Spooner, F., & Browder, D. M. (2003). Scientific-based research in education and students with low incidence disabilities. *Research in Practice for Persons With Severe Disabilities*, 28(3), 177–125.

- Sroufe, G. E. (1997). Improving the “awful reputation” of education research. *Educational Researcher*, 26(7), 26–28.
- Staiger, D., & Stock, J. (1997). Instrumental variables regression with weak instruments. *Econometrica*, 65, 557–587.
- Stake, J., & Nicken, S. (2005). Adolescent girls’ and boys’ science peer relationships and perceptions of the possible self as scientist. *Sex Roles*, 52, 1–11.
- Stevenson, D. L., Schiller, K. S., & Schneider, B. (1994). Sequences of opportunities of learning. *Sociology of Education*, 67(3), 184–198.
- Stoiber, K. C. (2002). Revisiting efforts on constructing a knowledge base of evidence-based intervention within school psychology. *School Psychology Quarterly*, 17(4), 533–546.
- Stolzenberg, R. M., & Relles, D. A. (1997). Tools for intuition about sample selection bias and its correction. *American Sociological Review*, 62(3), 494–507.
- Tanner, C. K., & Gallis, S. A. (1997). Student retention: Why is there a gap between the majority of research findings and school practice? *Psychology in the Schools*, 34(2), 107–114.
- Thompson, B., Diamond, K. E., McWilliam, R., Snyder, P., & Snyder, S. W. (2005). Evaluating the quality of evidence from correlational research for evidence-based practice. *Exceptional Children*, 71(2), 181–194.
- Van der Klaauw, W. (2002). Estimating the effect of financial aid offers on college enrollment: A regression-discontinuity approach. *International Economic Review*, 43(4), 1249–1287.
- Warr, M. (1993). Parents, peers, and delinquency. *Social Forces*, 72(1), 247–264.
- Weerland, F., & Smeenk, W. (2005). Peer similarity in delinquency for different types of friends: A comparison using two measurement methods. *Criminology*, 43(2), 499–523.

Weiss, C. H. (1999). Research-policy linkages: How much influence does social science research have? In UNESCO, *World Social Science Report 1999* (pp. 194–205). Paris: UNESCO/Elsevier.

West, S. G., Biesanz, J. C., & Pitts, S. C. (2000). Causal inference and generalization in field settings. In H. T. Reis & C. M. Judd (Eds.), *Handbook of research methods in social and personality psychology* (pp. 40–84). Cambridge, UK: Cambridge University Press.

Whiteley, B. J., Seelig, S. E., Weinshenker, M., & Schneider, B. (2002). *The AERA research grants program: Key findings of selected studies* (A report to the AERA Grants Board). Chicago: University of Chicago and NORC, Academic Research Centers.

Willinksky, J. (2005). Scientific research in a democratic culture: Or, What's a social science for? *Teachers College Record*, 107(1), 38–51.

Winship, C., & Mare, R. D. (1992). Models for sample selection bias. *Annual Review of Sociology*, 18, 327–350.

Winship, C., & Morgan, S. L. (1999). The estimation of causal effects from observational data. *Annual Review of Sociology*, 25, 659–706.

Xie, Y., & Shauman, K. (2003). *Women in science: Career processes and outcomes*. Cambridge, MA: Harvard University Press.