Education Policy Analysis Archives/Archivos Analíticos de Políticas Educativas
ISSN: 1068-2341
epaa@alperin.ca
Arizona State University
Estados Unidos

Wenglinsky, Harold
On Ideology, Causal Inference and the Reification of Statistical Methods: Reflections on "Examining Instruction, Achievement and Equity with NAEP Mathematics Data"
Education Policy Analysis Archives/Archivos Analíticos de Políticas Educativas, vol. 14, 2006, pp. 1-8
Arizona State University
Arizona, Estados Unidos

Available in: http://www.redalyc.org/articulo.oa?id=275020543017

How to cite
Complete issue
More information about this article
Journal's homepage in redalyc.org

Scientific Information System
Network of Scientific Journals from Latin America, the Caribbean, Spain and Portugal
Non-profit academic project, developed under the open access initiative

EDUCATION POLICY ANALYSIS ARCHIVES

A peer-reviewed scholarly journal
Editor: Sherman Dorn
College of Education
University of South Florida

Volume 14 Number 17 July 1, 2006 ISSN 1068–2341

On Ideology, Causal Inference and the Reification of Statistical Methods: Reflections on "Examining Instruction, Achievement and Equity with NAEP Mathematics Data"

Harold Wenglinsky
Academy for Educational Development
Regional Educational Laboratory of the Southeast

Citation: Wenglinsky, H. (2006). On ideology, causal inference and the reification of statistical methods: Reflections on "Examining instruction, achievement and equity with NAEP mathematics data." *Education Policy Analysis Archives, 14*(17). Retrieved [date] from http://epaa.asu.edu/epaa/v14n17/.

Abstract
The purpose of this article is to comment on the prior article entitled "Examining Instruction, Achievement and Equity with NAEP mathematics data," by Sarah Theule Lubienski. That article claims that a prior article by the author suffered from three weaknesses: (1) An attempt to justify No Child Left Behind (NCLB); (2) drawing causal inferences from cross-sectional data; (3) and various statistical quibbles. The author responds to the first claim, by indicating that any mention of NCLB was intended purely to make the article relevant to a policy journal; to the second claim, by noting his own reservations about using cross-sectional data to draw causal inferences; and to the third claim by noting potential issues of quantitative methodology in the Lubienski article. He concludes that studies that use advanced statistical methods are often so opaque as to be difficult to compare, and suggests some advantages to the quantitative transparency that comes from the findings of randomly controlled field trials.
Keywords: equity, mathematics achievement; mathematics instruction; NAEP.

SOME RIGHTS RESERVED Readers are free to copy, display, and distribute this article, as long as the work is attributed to the author(s) and **Education Policy Analysis Archives,** it is distributed for non-commercial purposes only, and no alteration or transformation is made in the work. More details of this Creative Commons license are available at http://creativecommons.org/licenses/by-nc-nd/2.5/. All other uses must be approved by the author(s) or **EPAA**. **EPAA** is published jointly by the Colleges of Education at Arizona State University and the University of South Florida. Articles are indexed by H.W. Wilson & Co. Please contribute commentary at http://epaa.info/wordpress/ and send errata notes to Sherman Dorn (epaa-editor@shermandorn.com).

Resumen

El objetivo de este articulo es comentar el trabajo "Examinando instrucción, logros, y equidad usando resultados de matemáticas de NAEP" de Sarah Theule Lubienski previamente publicado en EPAA. Ese articulo sostenía que otro articulo publicado por mi sufría de tres debilidades: (1) Intentaba justificar la ley federal sin abandonar ningún niño (2) Realizaba inferencias causales a partir de datos obtenidos en estudios cross seccionales; (3) otras objeciones sobre las estadísticas usadas. Este autor responde a la primera objeción indicando que cualquier mención a la ley federal sin abandonar ningún niño tuvo como objetivo establecer la relevancia de este trabajo para una revista de análisis político; a la segunda objeción, respondo indicando mis propias reservas al uso de estudios cross seccionales para establecer relaciones causales; y a la tercer objeción, indicando algunos problemas potenciales en la metodología cuantitativa del trabajo de Lubienski. Este autor concluye que los estudios que usan métodos cuantitativos avanzados suelen ser opacos y difíciles de entender y comparar, y sugiere que la transparencia cuantitativa deviene obtener resultados en experimentos con poblaciones seleccionadas al azar y campos controlados.

Editor's Note: This article is a response to Sarah Lubienski's (2006) article that appears at http://epaa.asu.edu/epaa/v14n14/*, which discussed Wenglinsky's (2004) article available at* http://epaa.asu.edu/v12n64/*. It is the practice of Education Policy Analysis Archives to publish one round of responses to articles where it is merited. Additional discussion of this and other articles is welcome online at* http://epaa.info/wordpress.

Over the last decade, the author has published half a dozen studies of relationships among school and teacher characteristics and student achievement using data from the National Assessment for Educational Progress (NAEP). Otherwise known as "the Nation's Report Card," NAEP provides test data on nationally representative samples of fourth, eighth and twelfth graders in a variety of subjects over multiple years. There are many methodological challenges faced in the analysis of NAEP data, but none as nettlesome as its cross-sectional nature. Because the data are cross-sectional, the finding of relationships between school characteristics, such as class size, and student achievement cannot be used to draw causal inferences. This point has been made by the author in nearly all of his publications on the topic, and is reiterated by opponents of the author's conclusions. If the policy conclusions are of a constructivist nature, the author finds himself attacked by conservative-learning researchers on the grounds that he is making causal inferences. If the policy conclusions are of a didactic nature, the author finds himself attacked by liberal-leaning researchers, generally on the same grounds. The critics usually also sprinkle in some methodological quibbles, such as wondering what the results would have been if variable X had been measured slightly differently, but the core criticism is that the cross-sectionality of the data make causal inferences impossible.

A recent instance of this occurred with Sarah Theule Lubienski's (2006) response to the current author's article on the achievement gap, "Closing the Racial Achievement Gap: The Role of Reforming Instructional Practices" (Wenglinsky, 2004). The current author's article distinguished between two types of racial achievement gap, that "between schools," meaning between predominantly minority and predominantly white schools, and "within schools," meaning between

White and minority students in the same school. The author found that a variety of instructional practices, mostly of the constructivist variety, but some not, were negatively related to the within-school achievement gap, but unrelated to the between-school achievement gap. The author concluded that using the identified practices might be a viable strategy for reducing the achievement gap within schools, but not the one between schools, which would require some more macro-institutional change. Lubienski's (2006) response, "Examining Instruction, Achievement and Equity with NAEP Mathematics Data," also found some instructional practices to be associated with the racial achievement gap, but with the caveats that only constructivist techniques evinced a relationship and that these techniques did not go so far as eliminating achievement gaps (the gaps she examined were analogous to the within-school gaps the author examined).

Lubienski's article raised the three questions that the current author finds are commonly raised about NAEP secondary analyses. First, Lubienski suggested an ideological underpinning to her critique: The current author's study was supposed to suggest how the Bush Administration's No Child Left Behind (NCLB) could succeed, whereas her study focused on empirical support for the National Council of Teachers of Mathematics' "reform-minded" (read constructivist) instructional practices. Second, Lubienski raised the issue of causal inferences, claiming that the current author used causal language in his study. And, third, she proposed some statistical quibbles which amounted to the notion that she approached her analysis slightly differently than the current author did.

With regard to the first argument, the current author had no ideological agenda. The mention of NCLB did not constitute and endorsement of it, but simply an attempt to find some relevance to policymakers in an article submitted to a policy journal. It has typically been the author's experience that the findings of NAEP secondary analyses rarely fit neatly into one ideological framework or another. Thus the author's study of school finance found that the effectiveness of school dollars depended upon how they were spent (Wenglinsky, 1997). As another example, the author's foray into the debate about whether educational technology made a difference found that technology effects depended upon how the technology was used (Wenglinsky, 2005). Rarely are the findings from statistical analyses of large-scale data unequivocal, and "Closing the Racial Achievement Gap" was no different, finding that the effective practices were an ideological potpourri, leaning somewhat towards the constructivist side.

The second argument, about causal inferences, is to some extent a red herring. As Lubienski admits, the current author acknowledged repeatedly that causal inferences cannot be drawn from cross-sectional data. He specifically noted that while he would use the phrase "school effect," he meant it in the statistical sense (as in "effect size") and did not intend it to connote causality. Lubienski herself sometimes falls into causal language, such as when she refers to instructional practices as "predictors" of achievement, suggesting that they are temporally, and thus causally, prior to test scores (Lubienski, 2006, p. 7). And in another analysis of NAEP data, in which Lubienski seeks to measure the relationship between attending a private school and student achievement, she talks of private-school effects (C. Lubienski & S. T. Lubienski, 2006). Given this challenge in both Wenglinsky's and Lubienski's work, one may rightly ask whether it is worthwhile to engage in secondary analyses of NAEP at all. The answer is that, while correlational analyses are not good at establishing causal relationships they are good at identifying variables that should subsequently be subject to more rigorous analyses. This view is the rationale behind the Institute for Education Sciences framework for its discretionary grants program; it has a continuum of research goals under which researchers can apply, beginning with secondary analyses to identify variables of importance, proceeding to developmental studies that make the transition from variable identification to the creation of an intervention, finally to experimental studies that can establish causation for interventions. The reason that correlational analyses are an important first step is because they

suggest where to concentrate and where not to concentrate. Given that, in the Lubienski study, the use of calculators proved unrelated to the racial achievement gap, it is unlikely that providing calculators to students, alone, is a worthwhile intervention. Thus Lubienski is correct in suggesting that cross-sectional data do not support causal inferences, but such secondary analyses are a crucial piece of work prior to the development and testing of educational interventions.

There is perhaps a more important reason that secondary analyses must be viewed as preliminary to more rigorous research designs, and that is because secondary analyses nearly always fall victim to "statistical reification." This term refers to the fact that researchers typically treat the statistical method of the day as an absolute basis for truth, even though the rationale behind using the technique, to say nothing of the mathematics involved, is typically opaque. Why use hierarchical linear modeling rather than structural equation modeling? When is multicollinearity severe enough to discredit results, given that most of the research questions of interest involve creating a significant degree of multicollinearity? Statistical quibbles can be raised about any secondary analysis. Lubienski's is no exception. Although she does not present the correlations among her factors, she refers to their being highly correlated. Highly-correlated factors suggest a poorly-fitting factor model and therefore potentially invalidate it. In addition, the proper way to verify a factor structure is through confirmatory factor analysis using a separate replication sample, not through creating scales and running Cronbach Alphas on the same data. And, disturbingly, here Cronbach Alpha's are mostly below what many consider to be the cutoff of .7. One other issue is that her models disaggregate teacher effects to the student level, which means that the instructional variables are only partitioning student-level variance in achievement, not school-level variance in achievement, and thus may understate the size of instructional effects. Does all of the foregoing invalidate her conclusions? The reader will probably decide by comparing the findings to his or her own experiences with education reform.

Secondary analyses need to be conducted with limited goals, and some kind of more robust research design (quasi-experimental or experimental) used for the more ambitious goal of demonstrating the efficacy of an intervention. One reason for this is that experiments are designed to support causal inferences, by holding constant all variables besides exposure to the treatment. They therefore address selection bias in a way that statistical analyses cannot. But a more important reason is the transparent nature of the results of an experiment. The results are transparent because they generally involve performing a student's t-test on two raw scores, that of the treatment and that of the control. These kinds of comparisons are more likely to be persuasive to a policymaker or educator than the elaborate debates over the appropriate multivariate method, the appropriate fit statistic, or any number of other running debates among quantitative methodologists. This is not to say that experiments are the "gold standard," but simply that they are less subject to reification, and therefore more trustworthy from the standpoint of making policy decisions.

References

Lubienski, C., & Lubienski, S.T. (2006). *Charter, private, public schools and academic achievement: New evidence from NAEP mathematics data.* New York: National Center for the Study of Privatization in Education, Teachers College, Columbia University.

Lubienski, S.T. (2006). Examining instruction, achievement, and equity with NAEP mathematics data. *Educational Policy Analysis Archives, 14*(14). Retrieved June 30, 2006, from http://epaa.asu.edu/epaa/v14n14/.

Wenglinsky, H. (1997). *When money matters: How educational expenditures improve student achievement and how they don't.* Princeton, NJ: Educational Testing Service.

Wenglinsky, H. (2004). Closing the racial achievement gap: The role of reforming instructional practices. *Educational Policy Analysis Archives, 12*(64). Retrieved June 30, 2006, from http://epaa.asu.edu/epaa/v12n64/.

Wenglinsky, H. (2005). *Using technology wisely: The keys to success in schools.* New York, NY: Teachers College Press.

About the Author

Harold Wenglinsky
Academy for Educational Development
Regional Educational Laboratory of the Southeast

Email: hwenglinsky@aed.org

Harold Wenglinsky is a program officer at the Academy for Educational Development and the Co-investigator of a randomized controlled trial of a state-level intervention, the Alabama Math, Science and Technology Initiative. He still enjoys analyzing NAEP data.

EDUCATION POLICY ANALYSIS ARCHIVES **http://epaa.asu.edu**

Editor: Sherman Dorn, University of South Florida

Production Assistant: Chris Murrell, Arizona State University

General questions about appropriateness of topics or particular articles may be addressed to the Editor, Sherman Dorn, epaa-editor@shermandorn.com.

Editorial Board

Michael W. Apple
University of Wisconsin

Robert Bickel
Marshall University

Casey Cobb
University of Connecticut

Gunapala Edirisooriya
Youngstown State University

Gustavo E. Fischman
Arizona State Univeristy

Gene V Glass
Arizona State University

Aimee Howley
Ohio University

William Hunter
University of Ontario Institute of Technology

Benjamin Levin
University of Manitoba

Les McLean
University of Toronto

Michele Moses
Arizona State University

Michael Scriven
Western Michigan University

John Willinsky
University of British Columbia

David C. Berliner
Arizona State University

Gregory Camilli
Rutgers University

Linda Darling-Hammond
Stanford University

Mark E. Fetler
California Commission on Teacher Credentialing

Richard Garlikov
Birmingham, Alabama

Thomas F. Green
Syracuse University

Craig B. Howley
Ohio University

Daniel Kallós
Umeå University

Thomas Mauhs-Pugh
Green Mountain College

Heinrich Mintrop
University of California, Berkeley

Anthony G. Rud Jr.
Purdue University

Terrence G. Wiley
Arizona State University

EDUCATION POLICY ANALYSIS ARCHIVES

English-language Graduate-Student Editorial Board

Noga Admon
New York University

Jessica Allen
University of Colorado

Cheryl Aman
University of British Columbia

Anne Black
University of Connecticut

Marisa Cannata
Michigan State University

Chad d'Entremont
Teachers College Columbia University

Carol Da Silva
Harvard University

Tara Donahue
Michigan State University

Camille Farrington
University of Illinois Chicago

Chris Frey
Indiana University

Amy Garrett Dikkers
University of Minnesota

Misty Ginicola
Yale University

Jake Gross
Indiana University

Hee Kyung Hong
Loyola University Chicago

Jennifer Lloyd
University of British Columbia

Heather Lord
Yale University

Shereeza Mohammed
Florida Atlantic University

Ben Superfine
University of Michigan

John Weathers
University of Pennsylvania

Kyo Yamashiro
University of California Los Angeles

Archivos Analíticos de Políticas Educativas

Associate Editors
Gustavo E. Fischman & Pablo Gentili
Arizona State University & Universidade do Estado do Rio de Janeiro

Founding Associate Editor for Spanish Language (1998—2003)
Roberto Rodríguez Gómez

Editorial Board

Hugo Aboites
Universidad Autónoma Metropolitana-Xochimilco
Dalila Andrade de Oliveira
Universidade Federal de Minas Gerais, Belo Horizonte, Brasil
Alejandro Canales
Universidad Nacional Autónoma de México
Erwin Epstein
Loyola University, Chicago, Illinois
Rollin Kent
Universidad Autónoma de Puebla. Puebla, México
Daniel C. Levy
University at Albany, SUNY, Albany, New York
María Loreto Egaña
Programa Interdisciplinario de Investigación en Educación
Grover Pango
Foro Latinoamericano de Políticas Educativas, Perú
Angel Ignacio Pérez Gómez
Universidad de Málaga

Diana Rhoten
Social Science Research Council, New York, New York
Susan Street
Centro de Investigaciones y Estudios Superiores en Antropologia Social Occidente, Guadalajara, México
Antonio Teodoro
Universidade Lusófona Lisboa,

Adrián Acosta
Universidad de Guadalajara México
Alejandra Birgin
Ministerio de Educación, Argentina
Ursula Casanova
Arizona State University, Tempe, Arizona
Mariano Fernández Enguita Universidad de Salamanca. España
Walter Kohan
Universidade Estadual do Rio de Janeiro, Brasil
Nilma Limo Gomes
Universidade Federal de Minas Gerais, Belo Horizonte
Mariano Narodowski
Universidad Torcuato Di Tella, Argentina
Vanilda Paiva
Universidade Estadual Do Rio De Janeiro, Brasil
Mónica Pini
Universidad Nacional de San Martin, Argentina
José Gimeno Sacristán
Universidad de Valencia, España
Nelly P. Stromquist
University of Southern California, Los Angeles, California

Carlos A. Torres
UCLA

Claudio Almonacid Avila
Universidad Metropolitana de Ciencias de la Educación, Chile
Teresa Bracho
Centro de Investigación y Docencia Económica-CIDE
Sigfredo Chiroque
Instituto de Pedagogía Popular, Perú
Gaudêncio Frigotto
Universidade Estadual do Rio de Janeiro, Brasil
Roberto Leher
Universidade Estadual do Rio de Janeiro, Brasil
Pia Lindquist Wong
California State University, Sacramento, California
Iolanda de Oliveira
Universidade Federal Fluminense, Brasil
Miguel Pereira
Catedratico Universidad de Granada, España
Romualdo Portella do Oliveira
Universidade de São Paulo
Daniel Schugurensky
Ontario Institute for Studies in Education, Canada
Daniel Suarez
Laboratorio de Politicas Publicas-Universidad de Buenos Aires, Argentina

Jurjo Torres Santomé
Universidad de la Coruña, España