Response to "What Do Klein et al. Tell Us About Test Scores in Texas?"¹

Stephen P. Klein, Laura S. Hamilton, Daniel F. McCaffrey, Brian M. Stecher
RAND


We have reviewed the article by Toenjes (2005). Below we summarize our responses. First, Toenjes incorrectly describes the focus of our study. As we note in our paper, the findings from the 20-school analysis merely triggered and supplemented our statewide analyses. Most of our article examined the stark discontinuity between TAAS and NAEP trends for the entire state of Texas (namely, the meteoric increase in scores and the narrowing of the gap between racial/ethnic groups on TAAS not being reflected on NAEP). The huge discrepancy between NAEP and state test scores in Texas is well-documented, it is still present, and it has been described in other studies (Haney, 2000; McCombs et al., 2005; Peterson & Hess, 2005). Toenjes’ comments about the representativeness of the 20 schools are irrelevant to our statewide analyses, and that analysis was the focus of our paper.

Second, Toenjes appears to have misunderstood the purpose of our 20-schools analysis. The intent was not to make conclusions about relationships between achievement and socioeconomic status statewide. Rather, the purpose was to examine whether, in this particular sample for which scores on two different tests were available, the TAAS functioned in the same way as an alternative test that many other states use, namely, the Stanford 9. The fact that the TAAS and Stanford 9

¹ Accepted under the editorship of Sherman Dorn. Send commentary to Casey Cobb (casey.cobb@uconn.edu).
showed radically different relationships with SES does not tell us anything about what those relationships would be in a larger sample, and we never claim that it does. In particular, we do not, as Toenjes claims, use the 20-schools study to conclude that TAAS lacks validity.

Instead, we used the findings with the 20 schools to show that in these particular schools, the scores on a commonly used test of mathematics achievement did not correspond to scores on the TAAS. The difference between these two tests was much larger than can plausibly be explained by differences in content coverage, and therefore raises questions about the meaning of TAAS scores for these schools. As we state clearly, “We are therefore reluctant to draw conclusions from our findings with these schools or to imply that these findings are likely to occur elsewhere in Texas. Nevertheless, they do suggest the desirability of periodic administration of external tests to validate TAAS results. This procedure, which is sometimes referred to as ‘audit testing,’ could have been incorporated into the study of the Metropolitan Achievement Test discussed previously.” This is a much more cautious conclusion than Toenjes claims we make.

Third, Toenjes misunderstands the relationship between our study and the Grissmer, Flanagan, Kawata, & Williamson (2000) study. Toenjes states that we concluded that the observed increases in Texas student academic performance reported by Grissmer et al. (2000) were highly suspect. Our paper contains no such conclusion. We focused on whether the TAAS scores were suspect, whereas Grissmer et al. examined gains in NAEP. As we clearly state in our paper, “these studies differed in the questions they investigated, the data they analyzed, and the methodologies they employed.” The studies do not in fact produce conflicting findings; they simply address different questions.

References


About the Authors

Stephen P. Klein  
Laura S. Hamilton  
Daniel F. McCaffrey  
Brian Stecher  
RAND

Email: klein@rand.org

Dr. Stephen P. Klein is a Senior Research Scientist at RAND, where for the past 30 years he has led studies on health, criminal justice, military manpower, and educational issues. His current projects include analyzing licensing examinations in teaching and other professions, delivering computer performance tests over the Web, and measuring the effects of instructional practices and curriculum on student performance.

Dr. Laura S. Hamilton is a Senior Behavioral Scientist at RAND where she conducts research on educational assessment and the effectiveness of educational reform programs. Her current projects include a study of systemic reforms in math and science, an investigation of district and school responses to standards-based accountability, and a study of teachers’ and principals’ use of information from a value-added assessment system.

Dr. Daniel F. McCaffrey is a Senior Statistician at RAND where he works on studies of health and educational issues. His recent research has focused the study of value-added model estimates of teacher effects by evaluating the literature and statistical methods for value-added modeling, by developing new models and algorithms for estimating effects, and by conducting empirical evaluations of the various estimates. He is also studying the effect of a value-added assessment program on student outcomes and educational practice.

Dr. Brian Stecher is a Senior Social Scientist in the Education program at RAND. Dr. Stecher’s research emphasis is applied educational measurement, including the implementation, quality, and impact of state assessment and accountability systems and the cost, quality, and feasibility of performance-based assessments in mathematics and science.
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 Antropología Social Occidente, Guadalajara, México

Antonio Teodoro
Universidade Lusófona Lisboa,

Adrián Acosta
Universidad de Guadalajara

Alejandra Birgin
Ministerio de Educación, Argentina

Ursula Casanova
Arizona State University, Tempe, Arizona

Mariano Fernández
Enguia 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 Martín, 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 Políticas Publicas-Universidad de Buenos Aires, Argentina

Jurjo Torres Santomé
Universidad de la Coruña, España