



Education Policy Analysis Archives/Archivos
Analíticos de Políticas Educativas

ISSN: 1068-2341

epaa@alperin.ca

Arizona State University
Estados Unidos

Pullin, Diana

Legal Issues in the Use of Student Test Scores and Valueadded Models (VAM) to Determine
Educational Quality

Education Policy Analysis Archives/Archivos Analíticos de Políticas Educativas, vol. 21, enero, 2013,
pp. 1-24

Arizona State University
Arizona, Estados Unidos

Available in: <http://www.redalyc.org/articulo.oa?id=275029728006>

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

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

SPECIAL ISSUE

Value-Added: What America's Policymakers Need to Know and Understand

education policy analysis archives

A peer-reviewed, independent,
open access, multilingual journal



Arizona State University

Volume 21 Number 6

January 29th, 2013

ISSN 1068-2341

Legal Issues in the Use of Student Test Scores and Value-added Models (VAM) to Determine Educational Quality

Diana Pullin

Boston College

United States of America

Citation: Pullin, D. (2013). Legal issues in the use of student test scores and value-added models (VAM) to determine educational quality. *Education Policy Analysis Archives*, 21(6). This article is part of EPAA/AAPE's Special Issue on *Value-Added: What America's Policymakers Need to Know and Understand*, Guest Edited by Dr. Audrey Amrein-Beardsley and Assistant Editors Dr. Clarin Collins, Dr. Sarah Polasky, and Ed Sloat. Retrieved [date], from <http://epaa.asu.edu/ojs/article/view/1160>.

Abstract: A growing number of states and local schools across the country have adopted educator evaluation and accountability programs based on the use of student test scores and value-added models (VAM). A wide array of potential legal issues could arise from the implementation of these programs. This article uses legal analysis and social science evidence to discuss potential legal challenges by educators to the use of VAM that should be considered by public policy makers. It also discusses potential ways VAM might be used as evidence in support of legal claims by students concerning access to educational opportunity.

Keywords: Law; value-added models (VAM); teacher quality; teacher evaluation; merit pay; testing; accountability

Cuestiones legales en el uso de resultados de exámenes de estudiantes y modelos de valor añadido (MVA) para determinar calidad educativa

Resumen: Un número creciente de estados y distritos escolares en todo el país han aprobado

Journal website: <http://epaa.asu.edu/ojs/>

Facebook: /EPAAA

Twitter: @epaa_aape

Manuscript received: 7/3/2012

Revisions received: 1/19/2013

Accepted: 1/19/2013

programas de evaluación docente y rendición de cuentas basados en el uso de resultados de los exámenes de los estudiantes y modelos de valor añadido (MVA). Una amplia gama de posibles problemas legales pueden derivarse de la aplicación de estos programas. Este artículo utiliza análisis jurídicos para analizar posibles desafíos legales para el uso de MVA que deben ser considerados por los encargados de las políticas públicas. Asimismo, se discute sobre posibles maneras en que MVA pueden ser utilizados como pruebas en apoyo de demandas legales de estudiantes relativas a las oportunidades educativas.

Palabras clave: Leyes; modelos de valor añadido (MVA); calidad docente; evaluación docente; pago por mérito; pruebas; la rendición de cuentas.

Questões legais no uso de notas dos exames dos alunos em modelos de valor agregado (MVA) para determinar a qualidade do ensino

Resumo: Um número crescente de estados e distritos escolares em todo o país aprovaram programas de avaliação de professores e de responsabilização com base na utilização de resultados de testes de estudantes e modelos de valor agregado (MVA). Uma vasta gama de problemas legais podem surgir a partir da implementação desses programas. Este artigo utiliza a análise jurídica para analisar os potenciais desafios legais para o uso de MVA a ser considerado pelos formuladores de políticas públicas. Também discute maneiras possíveis que MVA possam ser usados por de estudantes em apoio a ações judiciais relativas as oportunidades educacionais.

Palavras-chave: Leis; modelos de valor agregado (MVA); qualidade dos professores; avaliação docente; pagamento por mérito; avaliações; prestação de contas.

Introduction

The quality of the teaching force is the current focus of education reform. There has been rapid adoption across the country of new programs to evaluate teachers using student test scores as well as metrics based on those scores, usually value-added models (VAM).¹ And there is increasing consideration of the use of VAM to evaluate school leaders and institutions of higher education. Advocates for VAM assert that it can promote accountability and improve teaching and learning (Harris, 2011). This article argues that public policy makers should be aware of potential legal ramifications of the use of VAM when they are considering the adoption or implementation of VAM approaches. As this paper outlines, a wide array of potential legal issues could arise from the implementation of these programs. While there are few court cases to date directly addressing the use of VAM, there are statutes and regulations as well as past cases that shed light on how courts might assess the use of VAM. Many of the relevant legal issues include consideration of the social science evidence concerning VAM and its use. This article uses legal analysis and social science evidence to discuss potential ways educators might challenge the use of VAM, as well as potential ways VAM data might be used as evidence in support of legal claims by students or citizens alleging denials of educational opportunity.

A small number of states and local school systems adopted VAM approaches or use of student achievement test scores to evaluate educators on their own initiative (Amrein-Beardsley,

¹ Within the range of approaches to assessing student achievement growth over time, VAM is a term used to “refer to a variety of sophisticated statistical techniques that use one or more years of prior student test scores, as well as other data, to adjust for preexisting difference among students when calculating contributions [of educators or institutions] to student test performance” (NRC, 2010, p.1). Throughout, this article uses the term “VAM” to apply to the wide array of approaches that use student scores to assess educators or institutions.

2008, 2012; Braun, 2005; Harris, 2011; Hill, Kapitula & Umland, 2010), but more significant has been the rapid adoption of these approaches through requirements in recent state statutes and regulations (McGuinn, 2012; National Council on Teacher Quality (NCTQ), 2012). These programs were often responses to initiatives implemented by the Obama Administration, which called for VAM approaches as a condition for the award of funds through the Race to the Top program (RTTT) created by the American Recovery and Reinvestment Act (ARRA) (2009). The RTTT initiative included requirements for the use of student growth measures and the application of information from those data to make judgments about educator quality. A January 2012 report identified 24 states requiring the use of student achievement data as part of teacher evaluations following the adoption of RTTT (National Council on Teacher Quality, 2012).

The VAM approach and student test score information are being used to determine educator tenure and termination, merit pay bonuses for individuals or faculties, educator professional development needs, as well as having an impact on educators' professional reputations. For example, in the District of Columbia, 200 teachers were dismissed recently, based in part on VAM data (Hill et al., 2010). The media and public have developed an interest in these metrics as newsworthy indicators of the quality of local schools and teachers (*Los Angeles Times*, 2012; Song, 2012). In addition, there is growing interest in the use of student test scores to make determinations about the quality of the educator preparation programs teachers completed (Floden, 2012; Knight, Edmondson, Lloyd, Arbaugh, Nolan, Whitney, et al., 2012; Sawchuk, 2012).

While there has not yet been much litigation involving VAM, legal disputes are inevitable given potential high stakes individual and institutional consequences. This article sets out the policy contexts and statutory requirements associated with VAM. It describes the litigation concerning the use of student test scores to evaluate teachers and VAM that has arisen thus far. It provides a survey of the range of legal issues that might be relevant for policymakers because they might arise in the future in the implementation and utilization of VAM. It also discusses the social science controversies associated with whether it is appropriate to use VAM for high-stakes purposes and how those controversies could impact the outcome of legal disputes associated with VAM.

The Public Policy Context for Legal Claims on the Challenges and Uses of VAM

We all strive to have public schools that continually work to improve themselves and outcomes for their students. Any public official advocating the use of VAM as a policy tool acts with particular presumptions about education, testing and social science, as well as the appropriate role of law in our society. Like education policymakers, judges, labor arbitrators or hearing officers deciding disputes concerning the use of student test scores will be acting as concerned citizens just like the rest of us. The overall context in which judges apply the appropriate legal standards to scrutinize a program is important for policymakers to understand.

VAM initiatives are consistent with a highly publicized press from the business community and many politicians to make government services more like private business, data-driven to measure productivity and accountability (Kupermintz, 2003). VAM approaches are in part a response to concerns that the current system of selecting and compensating teachers based their education and credentials is insufficient for insuring teacher quality (Corcoran, 2011; Gordon, Kane & Staiger, 2006; Hanushek & Rivkin, 2012; Harris, 2011). There have been increasing expressions of concern that teacher evaluation practices are not robust and do not improve practice (Kennedy, 2010). In the contemporary public policy context, much of the support for the use of student test scores for educator evaluation comes from a concern that the current system for evaluation is

ineffective and that the current legal protections for teachers are too cumbersome for schools seeking to terminate teachers (Harris, 2009, 2011).

In addition to responding to public discussions about education quality, any judge assessing legal claims will be heavily influenced by the tradition of judicial deference to education policymakers in most types of education law disputes. Judges have regularly declared that determinations about the provision of education programs should for the most part be left in the hands of professionals and elected officials (Dagley & Veir, 2002; *Massachusetts Federation of Teachers, AFT, AFL-CIO v. Board of Education*, 2002; Zirkel, 2003). For example, in past court cases over issues concerning teacher evaluations, terminations, and licensure, the overwhelming majority of the cases were won by school officials and few of the cases decided in court actually involved disputes over the quality of teachers' pedagogic practices (Pullin, 2010; Zirkel, 2003). However, as this article will discuss, this does not mean that education policymakers will always prevail in disputes over the use of VAM. Social science evidence might play a significant role in future legal disputes because of state and federal statutory requirements concerning the evidence needed to support programs.

The Federal and State Requirements for VAM

The 2002 No Child Left Behind Act (NCLB) implemented what was probably the most significant set of federal requirements ever imposed on states and local public schools. The cornerstone of NCLB was heavy reliance on state systems of high stakes student testing and the use of student test data to determine school quality through measures of adequate yearly progress in improved student scores (AYP). However, there arose concerns about the lack of provisions in NCLB for educator accountability for student growth (Braun, 2005; Harris, 2011). The RTTT added to the existing NCLB requirements that state programs be adapted to include provisions for judging teacher effectiveness based on rates of growth in student test scores (U.S. Department of Education, 2010). These federal requirements paralleled in part a long-standing approach in several local districts and in the Tennessee state statutes, which adopted a VAM approach in the early 1990s as part of a state accountability system to measure the effects of teachers, schools, and school districts on student achievement (Braun, 2005; Harris, 2011; Kupermintz, 2003; McGuinn, 2012).

States adopted their own variations when they created VAM provisions in their statutes and regulations. For example, the Colorado statute mandates districts use VAM to assess or evaluate educator performance (Colorado Revised Statute Annotated § 22-9-102 (2), 2011). The Illinois statute uses VAM to identify incompetent or underperforming teachers or to incentivize educators to perform more effectively (Illinois Compiled Statutes Annotated 5/24A-20 (a)(10), 2010). Florida's statute includes educator evaluation but also seeks to identify educators' professional development needs (Florida Statutes Annotated § 1012.34, 2011), as does the Louisiana statute (Louisiana Revised Statutes Annotated § 17:3881, 2011).

States vary somewhat in the weight afforded VAM as a component of their educator evaluation system, with most coupling VAM data with other indicators, such as objective observations of teaching (Corcoran, 2010; McGuinn, 2012). One researcher identified five states that implemented statutes or regulations where more than half of a teacher's evaluation would be based on VAM data (Corcoran, 2010). Whether VAM is a small part or a major part of an educator evaluation system, the use of student test scores to evaluate educators for high-stakes consequences could lead to litigation.

Previous Court Challenges to the Use of Student Test Scores to Evaluate Educators

Given the widespread increase in state adoptions of VAM approaches to gathering information about educators, what has been the history of litigation over these types of initiatives? As discussed below, the use of student test scores to make decisions about teachers was challenged unsuccessfully in courts several times before the advent of VAM. Each outcome can be attributable at least in part to heavy judicial deference to state and local education policymakers and the allure of using test scores to make decisions about education quality.

There are several different provisions of the U.S. Constitution that can apply to the use of VAM to make high-stakes decisions about individual educators. The Due Process and Equal Protection Clauses of the Fifth and Fourteenth Amendments have in the past been important considerations by courts (Pullin, 2001, 2010; Superfine, 2008;). The Due Process Clause regulates decision-making in government by requiring proper procedures to promote accuracy in decision-making (procedural due process); it also requires government decisions that aren't arbitrary and seem fair under the circumstances (substantive due process) (*Bd. of Regents v. Roth*, 1972; *Cleveland Bd. of Education v. Loudermill*, 1985; Pullin, 2001). The Equal Protection Clause focuses upon the goals and methods used in government decision-making when public officials sort individuals into groups and treat them differently on the basis of that classification. In Equal Protection challenges to government programs, unless race/ethnicity or gender discrimination issues arise, courts most often ask whether the government is attempting to further a legitimate governmental interest and whether it is using a rational or reasonable means for doing so (Pullin, 2001). Certainly there is a legitimate interest in insuring the competence of educators, but whether or not a reasonable or rational way of doing so is being implemented is a separate question. Courts have applied these types of constitutional principles previously when making determinations in the small number of disputes that have arisen over the use of student test scores to evaluate educators (Pullin, 2001, 2010).

A 1973 federal appellate court decision upheld a teacher dismissal based on student scores (*Scheelhaase v. Woodbury Cent. Comm. Sch. Dist.*, 1973). A long time teacher was dismissed due to below average performance of her students on two standardized basic skills achievement tests. The teacher argued that the district violated her substantive due process by rights by acting arbitrarily in misusing student test scores. Although the trial court held that the teacher should be reinstated, having determined that a teacher's professional competence could not be based solely on her student's achievements, the Eighth Circuit of Appeals reversed the trial court. The appellate court held that state laws determine the standards by which teachers are to be hired and fired, thus empowering a local board of education to act with its best discretion. Therefore, the court found, districts have the authority to terminate a teacher as long as they were seeking to further the best interests of the educational system and acted in good faith, even in situations where others would conclude that they are acting unwisely or wrongly. The court emphasized "such matters as the competence of teachers and the standards of its measurement are not, without more, matters of constitutional dimensions. They are peculiarly appropriate to state and local administration" (*Scheelhaase v. Woodbury Cent. Comm. Sch. Dist.*, 488 F. 2d at 244).

In a 1987 Missouri case (*St. Louis Teachers Union, Local 420, American Federation of Teachers, AFL-CIO v. Board of Educ. of the City of St. Louis*), the teachers' union challenged a school district teacher evaluation program on grounds that the Equal Protection and Due Process Clauses of the U.S. Constitution were violated. Under the program, unsatisfactory student standardized test scores could subject English language, communications, and mathematics teachers to additional evaluation review. The teachers argued that in evaluating teachers based on their students' test scores, the district acted arbitrarily, capriciously, and irrationally because the student test was not designed for use as a teacher evaluation instrument. The federal district court judge found that insuring teacher competency and improving the quality of education are legitimate state objectives, and since the defendant board only evaluated teachers on the test scores (reading, language, math) that

corresponded to the subjects each teacher taught, the district did not violate the Equal Protection Clause by reviewing only certain teachers. The court also found that the district's classification of teachers based on their students' test results was rational and that subjecting some teachers to further evaluation was rationally related to state interests. The case was later settled out of court (Pullin, 2010).

In a more recent case (*Massachusetts Federation of Teachers, AFT, AFL-CIO v. Board of Education*, 2002), the two state teacher unions sued the Massachusetts State Board of Education in an effort to prevent implementation of a new state regulation that would require diagnostic testing of math teachers in certain low performing schools as a prerequisite for renewal of contracts for teachers in those schools. The low performing schools were chosen based on student performance on the state's student accountability testing program, the MCAS. Analyzing the case in light of the provisions of the Massachusetts Education Reform Act of 1993, the state's highest court concluded that the "Board has broad authority to establish such policies as are necessary to fulfill the purposes of the Act and to promulgate regulations that encourage innovation, flexibility, and accountability in schools and school districts" (*Massachusetts Federation of Teachers, AFT, AFL-CIO v. Board of Education*, 436 Mass. At 766). The court determined that the State Board had considerable discretion, specifically in mathematics, where the state's students had historically lagged. While the teachers argued that the regulation would violate their rights under the Equal Protection Clause by treating math teachers in low performing schools differently from math teachers in other schools, the court stated that the regulations were rationally related to the furtherance of a legitimate state interest in providing a high quality public education to every child through assessment of the subject matter knowledge of mathematics teachers. To the judges, there was no problem in singling out only teachers in certain subjects in certain schools, as statewide education reform is a large project and implementation understandably began with a particular focus.

Recent Court Cases Concerning the Use of Student Test Scores and VAM

The use of student test scores to evaluate educators as part of the use of VAM is a relatively new phenomenon across the country, so there haven't been many legal disputes thus far. However, the legal implications of VAM use have provoked some legal activity of which policymakers should be aware as they consider implementation.

Disputes over the public's right to know VAM results

Some of the first attention paid to VAM by the courts resulted from efforts by the media to obtain public release of individual teachers' student test scores or VAM results. These disputes involving major newspapers in New York City (*Mulgrew v. Board of Education*, 2011) and Los Angeles (Los Angeles Times, 2012; Song, 2012) brought into direct conflict educators' privacy interests in protecting their professional reputations and the media's interest in reporting on government programs. Since public school teachers serve in the public interest, the media asserted it was in the public's interest to know individual teacher scores so that citizens could assess the quality of their schools and parents could know more about their children's teachers. State statutes, such as those in New York, generally have privacy laws protecting individual private information and other statutory provisions allowing access to government information to which the public should have access given the goal of promoting government transparency and accountability (*Mulgrew*, 2011). These two sets of goals can stand in stark contrast in the context of providing access to individual teachers' student test scores or VAM data.

In California, the *Los Angeles Times* obtained data from state student achievement tests and hired an economist to conduct a VAM analysis based on student test scores. It then published rank-ordered lists of names of Los Angeles teachers. The *Los Angeles Times* has subsequently maintained a website explaining its approach and allowing users to pull down individual teacher data by name (*Los Angeles Times*, 2012). It is now pursuing litigation to seek access to the school district's own calculations of teachers' VAM data (Song, 2012).

In New York City, the teachers' union brought litigation in an effort to bar the *New York Times* and other media from publishing individual teacher data (*Mulgrew v. Board of Education of the City School District of the City of New York*, 2011). The information at issue in New York was individual teacher data from a VAM pilot program in the New York City schools. The issue of whether individuals could assert a right to the privacy of their own school records had never been explicitly addressed by the courts in interpreting the New York statute, which contained several specific exemptions to disclosure, but didn't directly address VAM data. The New York state appellate court noted in discussing the VAM data the tension between individual privacy and the public's interest in disclosure, but determined that the presumption in the law is in favor of the public right to know (*Mulgrew v. Bd. of Educ.*, 2011). The court found in part that the VAM data should be subjected to disclosure because they were objective statistical data about individuals, unlike teacher observation data a previous New York court had found to be private and not subject to disclosure because it was opinion about teacher performance. The court allowed release of VAM data because release of public sector job-performance information did not, in the court's view, constitute an unwarranted invasion of privacy (*Mulgrew v. Bd. of Educ.*, 2011). Subsequently, New York Governor Cuomo and the state legislature, in response to concerns of the teacher unions, initiated a revision of state law to limit access to a teacher's evaluation data to the parents or guardians of students in that teacher's class (Associated Press, June 21, 2012).

Disputes over how much weight to give VAM data

In another recent New York case, the teachers' union used litigation to challenge the state's formula for taking student growth data into account as part of teacher evaluations. In a successful effort to obtain almost \$700 million in federal RTTT funding, the state had collaborated with teacher unions to revise its teacher evaluation statute (Corcoran, 2010). However, the new regulations the state formulated to implement the statute were challenged by the union. The teachers' union successfully convinced a state court that the state's reliance on student test scores was too heavy and contravened a statutory obligation to utilize multiple measures of performance. The union also successfully argued that the state assigned disproportionate weight to student test scores and violated an obligation for local collective bargaining to define some components of the local evaluation system (*New York State United Teachers Association v. Board of Regents*, 2011).

Disputes over cheating on student tests

Given the high-stakes consequences of using VAM as an indicator of educator quality, there were in some places efforts to subvert the system. Students for centuries have mastered the art of cheating on tests (Cizek, 1999). As the stakes grow higher, incentives for cheating increase. Given the increasing attention to the use of student test scores to determine educator or institutional quality, it is perhaps not surprising that these initiatives have led to some inappropriate teacher and administrator behavior. States have acted accordingly, suspending or revoking licenses and, in one case successfully incarcerating a superintendent in a criminal case (Herold, 2012; Herold, & Mezzacappa, 2012; Otterman, 2011; Samuels, 2012; Severson, 2011; Vogell, 2011; Zubrzycki, 2012). It is not surprising that courts tend to support education officials in these types of actions to address cheating. So, for example, a Georgia appellate court found that it was appropriate for the

state to temporarily suspend the license of a kindergarten teacher who changed some of her student's incorrect answers on the Iowa Tests of Basic Skills (*Prof'l Standards Comm'n v. Denham*, 2001; see also *Scoggins v. Bd. Of Educ.*, 1988).

States have utilized forensic investigations by current or former criminal law investigators to scrutinize unusual score gains and licensure denials and criminal charges could be the result if educators are found to have a role in cheating (Zubrzycki, 2012). Sometimes accusations of cheating are based not on direct evidence, but on scientific or statistical analysis of possible indicators of cheating behaviors (Cizek, 1999). The methods used to detect cheating have largely not been regulated by policymakers, although some states have recently added governmental agencies to address the cheating problem given the increasingly high states consequences for educators or the institutions in which they work (Otterman, 2011; Waldman, 2012).

The Federal and State Requirements for Evidence to Support Education Programs

The cases and controversies described above do not provide many examples of how judges might look at the use of student test scores and VAM. Where there has been litigation, particularly concerning consequences for individual educators, public officials generally prevail. However, public officials may not prevail when they fail to follow requirements set forth in state or federal statutes, as was the case in the New York dispute over the weighting to apply in evaluations. And, there are recent state and federal statutory provisions about the quality of the evidence needed to substantiate the use of these education programs that could dramatically alter the outcomes of future litigation. This article argues that these provisions should be considered by policymakers seeking to ensure that they are meeting their goals for improving education. In addition to its student achievement goals, NCLB has requirements concerning the quality of evidence used to support or implement its programs. At numerous points, the NCLB sets out requirements for rigorous, scientifically based research to justify programs enacted under that law (20 U.S.C. § 7801). NCLB also requires adherence to professional and technical standards (20 U.S.C. § 6311(D)). Under NCLB, educational agencies are required to use assessments that are "valid and reliable, and consistent with relevant, nationally recognized professional and technical standards" and also "of adequate technical quality for each purpose required under [the] Act" (20 USCA § 6311).

The requirements states have adopted for the implementation of VAM are also important to consider. Some of the state statutes incorporate standards on the scientific quality of evidence to support VAM use. For example, statutes in Washington and Maryland call for evaluation models that are research-based. Washington state law calls for a "set of articulated teacher knowledge, skill, and performance standards for effective teaching that are evidence-based, measurable, meaningful, and documented in high quality research as being associated with improved student learning" (Washington Revised Code Annotated § 28^a.410.270, 2010). Maryland law requires that the "State Board shall adopt regulations that establish general standards for performance evaluations for certificated teachers and principals that include observations, clear standards, rigor, and claims and evidence of observed instruction" (Maryland Code Annotated, Education § 6-202, 2011).

The Colorado statute on educator evaluation calls for a "professionally sound and credible system" (Colorado Revised Statutes Annotated § 22-9-102 (2)) with "fair, transparent, timely, rigorous, and valid methods" for evaluation (§ 22-9-105.5(a)). The Illinois, Michigan and Utah statutes all call for school district evaluation systems that are "valid and reliable" (Illinois Compiled Statutes Annotated 5/24^a-20, 2010; Michigan Compiled Laws Annotated § 380.1249, 2010), and Utah Code Annotated § 53^a-10-106, 2010). These types of statutory requirements could be taken

into account in legal disputes in which the social science evidence concerning the defensibility of using VAM will be an important consideration for state or federal judges and hearing officers.

The Conflicting Social Science Perspectives on VAM

The state and federal statutory and regulatory language requiring evidence-based approaches and validity and reliability evidence to support programs seems clear. However, there are differences of perspective among social scientists about VAM and the defensibility of using it to make high-stakes decisions about educators. This article will not delve into all of that controversy here (see Harris, 2011, for an overview on VAM and NRC, 2010, enumerating some of the differing perspectives on VAM), but will highlight a couple of the most critical issues concerning the social science controversies associated with VAM about which policymakers should be aware.

A VAM methodology is a multi-faceted approach utilizing student test score and student demographic data and the statistical manipulation of that data to estimate the quality of educators and institutions. As a result, there are potentially multiple layers of legal scrutiny of VAM data metrics to assess the quality of the underlying student test scores, to assess the utilization of those scores through statistical modeling, and to assess the use of other statistical data on students and schools used in VAM. In many respects, it is the hybrid nature of VAM that leads to social science controversy. This controversy likely will impact litigation concerning VAM and its use when the social science evidence may be a critical factor in determining the legal outcome.

While there has been, since the RTTT initiative began, rapid increases in the use of student test results to promote accountability and make judgments about the effectiveness of individual educators or education institutions, there is not any clear, research-based consensus in the social science literature that these practices will work. In fact, most prominent education researchers, particularly those focused on the science of educational testing (psychometrics), have significant concerns about these uses of student test scores and the lack of evidence to support them, pointing to significant errors of measurement in VAM scores, the considerable instability of VAM scores over time and the lack of data to support the inferences drawn from the student scores as well (Amrein-Beardsley, 2008, 2012; Baker, Darling-Hammond, Haertel, et al., 2010; Newton, Darling-Hammond, Haertel & Thomas, 2010). One sociologist of education, who embraces the promise of VAM approaches, also reported considerable variability in VAM metrics resulting from the underlying psychometrics of the student tests and the varying statistical approaches taken to calculate VAM metrics (Ready, 2013). These types of concerns raise considerable doubts about the validity and reliability of VAM and its use.

On the other hand, a group of social scientists are strongly supportive of VAM and its use. The most prominent social science proponents of VAM are not traditional education researchers or testing specialists, but rather are economists and statisticians. They focus little, if at all, on the scientific quality of the underlying student test scores or the appropriateness of their use in VAM metrics, relying instead upon statistical models designed to reflect both the change in test scores over time (student growth) and the contexts in which the student scores are generated (see, for example, Gordon, Kane & Staiger, 2006; Hanushek & Rivkin, 2012; Harris, 2011). Their approach is focused upon the statistical or econometric methodology of VAM and a felt imperative to address the perceived shortcomings of current educator evaluation systems (Hanushek & Rivkin, 2012; Gordon, Kane & Staiger, 2006).

There are a small number of economists who have taken positions against VAM or high stakes uses of VAM (Lang, 2010; Rothstein, 2009, 2010). For example, economist Kevin Lang assessed the contrasting social science perspectives on VAM and concluded:

the economics and statistics communities have ...develop[ed] value-added measures that carry a scientific aura. However, economists have largely failed to recognize many of the problems with such [student test] measures. These problems are sufficiently important that they should preclude any automatic link between these measures and rewards or sanctions (2010, p. 168).

Several reports to Congress from cross-disciplinary groups of prominent social scientists convened by the prestigious National Research Council (NRC) at the National Academy of Sciences caution that there is limited defensibility of many high stakes educational testing practices. A 1999 NRC report, focusing primarily upon the impact of high-stakes tests on students, cautions against the use of a student test scores for high stakes decisions (NRC, 1999). Three later reports from the same organization also issue strong cautions on testing educators (NRC, 2001, 2008, 2012). Recently, NRC conducted a workshop to examine VAM practices and concluded that there are many technical challenges associated with the models and that years of research hadn't overcome the challenges (NRC, 2010). This report stated "persistent concerns about precision and [measurement] bias militate against employing value-added indicators as the principal basis for high-stakes decisions" (NRC, 2010, p. 59).

Professional and Technical Standards Related to VAM

Education researchers and other social scientists raising concerns about VAM approaches base many of their concerns on the failure of these programs to meet professional and technical standards. There have long been professional standards of practice concerning testing in education and employment (American Educational Research Association, American Psychological Association, National Council for Measurement in Education, 1999 (AERA/APA/NCME Test Standards); American Educational Research Association (AERA), 2002; Society of Industrial and Organizational Psychologists (SIOP), 2003). These are presumably the types of "professional and technical standards" NCLB requires to substantiate the use of an educational program (20 U.S.C. 6311). Judges or other decision-makers applying state statutes calling for valid and reliable educator evaluation processes might also turn to these professional standards for guidance.

Professional standards on testing state clearly that it is essential that a test or assessment is both valid and reliable for the purposes for which it is being used (AERA/APA/NCME, 1999; AERA, 2002; SIOP, 2003). This would presumably require validity and reliability evidence for both the underlying student scores as well as the use of those scores in VAM metrics. Validation of the use of a score is "an evaluation of the extent to which the proposed interpretations and uses [of a score] are plausible and appropriate" in an effort to meet the "scientific and social requirement that public claims and decisions be justified" (Kane, 2006, p. 17). According to professional standards, validation for each separate use of a test must be established: "When test scores are used or interpreted in more than one way, each intended interpretation must be validated" (AERA/APA/NCME, 1999, p.9). As the AERA position statement says:

Tests valid for one use may be invalid for another. Each separate use of a high-stakes test, for individual certification, for school evaluation...or for other uses requires a separate evaluation of the strengths and limitations of both the testing program and the test itself" (AERA, 2002, n.p.).

In order for the inferences made from a test or assessment to be valid, the test must also have sufficient reliability (AERA/APA/NCME, 1999). The reliability of a test or assessment is the precision of what it measures as well as whether the measure would be repeatable and dependable across time or across different versions or contexts (Haertel, 2006). A reliable test or assessment provides consistency of the measurement over time and would be free of significant errors of

measurement.

The AERA/APA/NCME Standards (1999) also call for evidence of the fairness of the uses of tests or assessments. This requires consideration not only of the properties of a test but also the ways tests results are reported and used and “the factors that are validly or erroneously thought to account for patterns of test performance” (p.73).

In addition to the attention to professional and technical standards mandated by NCLB, validity and reliability requirements for tests and test use have also been incorporated in federal regulations concerning claims of discrimination in employment (Equal Employment Opportunities Commission, 2010) and in the legal standards used in a wide range of court cases involving education and employment testing (*Gillespie v. State of Wis.*, 1985; *Griggs v. Duke Power Co.*, 1971; *Guardians Association v. Civil Service Commission*, 1979).

Potential Legal Challenges to the Use of VAM to Evaluate Educators

There is a high likelihood of legal challenges to the use of VAM when it is used for evaluation for high-stakes consequences like salary differentiation, termination, or damage to professional reputation. Policymakers will need to consider the social science controversies over VAM and its scientific defensibility. Given state and federal statutory mandates for accountability data based on valid and reliable approaches, the social science evidence will potentially be important to judges in ways that it was not in past court cases. In the limited number of past court cases, judges tended to support the use of student test scores to make decisions about individual educators. Given the new federal and state statutory requirements on the quality of evidence required to support education programs, judges could view very differently the use of student test scores and VAM metrics to assess educators.

Potential constitutional issues on the use of VAM for educator evaluation

The previous cases on the use of student test scores to evaluate educators discussed earlier afford an illustration of how courts might react in the future to constitutional challenges to the use of VAM. Yet, given how significantly VAM methodology differs from the straightforward use of student test scores to judge educators, these previous cases do not illuminate all the possibilities for litigation outcomes.

The earlier cases on use of student test scores seemed to turn in large part on the public policy phenomenon of the government’s interest in the momentum of education reform, the persuasive appeal of the meaning of student test scores. Those perspectives, when coupled with the unwillingness of judges to second-guess education officials have meant that the constitutional analysis is weighted toward outcomes in favor of the government. The outcomes of those earlier cases were, therefore, not completely surprising. Yet these judicial determinations of the rationality and constitutionality of state approaches to determining teacher quality and the fairness of treatment of teachers stand in contrast to the scientific standards of the education research community as well as the requirements of NCLB and of many state statutes concerning the quality of evidence about educator competence.

Would contemporary judges view the use of VAM in the same way as prior cases on the use of student test scores to determine teacher quality? The same Due Process and Equal Protection arguments might apply but the outcomes might be different, depending upon the policy perspective of the judges and the extent to which they are willing to dig deeply into the scientific issues concerning VAM methodology. Any future constitutional challenges under the Due Process or Equal Protection Clauses of the Fourteenth Amendment could incorporate considerations of the validity and reliability and fairness issues on the use of the student scores as well as the VAM itself.

Given the scientific issues associated with VAM methodologies, it is possible that the use of VAM to make a high-stakes decision about an educator would not even survive a rational basis review under Equal Protection analysis. At least two illustrations of this type of problem can be provided.

Constitutional problems with score attributions. One major potential constitutional problem with VAM programs arises from the limited grades and subject matters covered in most state student testing programs. Given the limitations on most student testing programs, only a minority of teachers teach the subjects or grades covered by the tests and the VAM metrics. States report that a majority of their teachers work in areas not covered by student tests (McGuinn, 2012). The VAM metrics are a particular concern for teachers of students with disabilities or whose mother language was not English and are still acquiring English language proficiency (English language learners); these teachers serve populations particularly at risk of low achievement and often teach across grade levels and subject areas (Council for Exceptional Children, 2012; Holdheide, Goe, Croft & Reschly, 2012).

In addition to special education or English learner teachers, for any teacher who works in a subject, area, or grade level not covered by the student tests, like science, social studies, physical education or art teachers, there is no student test score data available. In order to obtain “student test scores” for teachers not teaching a tested subject or grade level, statistical attributions are made, usually from the mean student test scores of other teachers in the building or district (Braun, 2005; Kupermintz, 2003). This means that an individual teacher’s VAM performance is based on the test scores of other teachers. Sometimes, however, rather than having statisticians pick the scores for attribution, other approaches can be taken to address the problem of teachers without student scores. The potential fairness and accuracy issues associated with these approaches were illustrated recently in press reports about Tennessee, where, to address the problem of missing scores for teachers of untested grades and subjects, apparently the state allowed teachers to pick for themselves the teacher whose student scores they want to be judged upon by attribution (Winerip, 2011).

Any effort to base a VAM judgment about the performance of one teacher on the basis of the performance of other teachers raises significant constitutional issues, especially if the situation is like the one purported to occur in Tennessee. The attribution of student scores approach, either based on a statistical algorithm or a teacher’s own choice, seems to present significant constitutional problems. This could be a basis for VAM to fail judicial review, even given the tradition of strong deference to education policymakers, when judges consider whether the mechanism is a reasonable or rational approach.

Constitutional problems with curriculum and testing changes. Another set of potential constitutional issues arises from the speed with which VAM programs have been implemented and the challenging contexts for implementation. At the same time states are implementing new VAM approaches, most are also trying to significantly raise their curriculum standards through the voluntary national Common Core Curriculum Standards and to revise substantially the student assessments associated with those standards (McGuinn, 2012). The curriculum standards substantially change the content of curriculum, making it more challenging for students and their teachers. New assessment approaches to assess the new curriculum could also significantly impact the nature of the tests students take (McGuinn, 2012).

Previous litigation on behalf of students established the precedent that advanced notice and timely implementation of a high school graduation testing requirement is necessary for students to have adequate opportunity to prepare for the test, for school districts to develop and implement curriculum and remedial programs and opportunities for the state to correct any deficiencies in the test (*Debra P. v. Turlington*, 1981). Educators might mount successful Due Process and Equal Protection claims to mirror those successfully used by students, asserting that their property and liberty interests in their continued employment and professional reputations are at risk (*Bd. of Regents*

v. Roth, 1972). This could be an issue at any point in time, but is a particular issue at present due to the rapid implementation of the new curriculum standards and assessments linked to that curriculum (McGuinn, 2012). Educators may be able to successfully assert Due Process and Equal Protection violations as a result of the fact that they did not have a fair opportunity to prepare and implement the curriculum and instruction required to meet the new requirements.

Potential issues under civil rights statutes on the use of VAM for educator evaluation

If an employment test, selection process, or evaluation has disparate results for protected groups, such as racial or ethnic minorities, the nondiscrimination requirements of federal employment statutes will apply. Title VI of the Civil Rights Act of 1964 (42 U.S.C. 2000d) bars discrimination in education programs receiving federal financial support. The more widely used basis for challenging discrimination in employment testing, selection, and evaluation is Title VII of the Civil Rights Act of 1964 (42 U.S.C. 2000e). It is in the application of Title VII that courts have in the past been most heavily involved in issues of the validity and reliability of inferences based on tests and evaluations it is this statute that may be used by educators to challenge the use of VAM.

Title VII specifically permits the use of “professionally developed ability tests” for employment decisions, provided that their “administration or action upon the results is not designed, intended, or used to discriminate because of race, color, religion, sex, or national origin” (42 U.S.C. § 2000e-2(h); see also *Griggs v. Duke Power Co.*, 1971). Traditionally, an objective test that has a disparate impact on protected groups requires a demonstration by the employer that the test is “job-related with regard to the position in question and consistent with business necessity” (42 U.S.C. § 2000e-2(k)(1)(A)(i)(2010); EEOC, 29 C.F.R. § 1607.5; *Griggs v. Duke Power*, 1971). This requires evidence that the test actually measures skills, knowledge, or ability required. There is also broad coverage under Title VII and the Equal Employment Opportunity Commission’s (EEOC) *Uniform Guidelines* for Title VII enforcement (29 C.F.R. § 1607.2(B); 29 C.F.R. § 1607.2(C)) concerning not only use of standardized objective tests, but all employee selection devices, including “performance evaluations” (29 C.F.R. § 1607.2(C) for “successful performance on the job” (*Contreras v. City of Los Angeles*, 1981).

There is a long history of court cases concerning allegations of discrimination in employment testing, focused on paper and pencil standardized selection tests or hands-on performance tests. VAM doesn’t fit neatly into these past cases, given its hybrid nature, but it clearly fits into the requirements concerning tests, selection processes, and evaluations for decision-making in the employment context. Education researchers have described VAM metrics as assessments that fit the types of measures covered by Title VII requirements (Hill, Kapitula & Umland, 2010).

In previous cases involving federal statutes and regulations on employment discrimination, considerable attention has been paid to the technical quality of test-based approaches to decision-making (Lindeman & Grossman, 1996). Test validity is always of primary importance in an assessment of an employment test. This obligation arises out of the federal statutes, the Uniform Guidelines of the Equal Employment Opportunities Commission, and also the relevant standards of practice for the testing profession applied often through expert testimony in legal disputes (29 C.F.R. §§ 1607.14 and .15; see also *U.S. v. State of S.C.*, 1977; *Washington v. Davis*, 1976). There should be sufficient evidence to support the inferences made from a test. The test should measure what it purports to measure and “generally, validity is defined as the degree to which a certain inference from a test is appropriate and meaningful” (see, for example, *Richardson v. Lamar County Board of Education*, 1989, p. 806).

Given the concerns expressed about validity and reliability of VAM inferences raised by education researchers, the employment testing requirements of the federal civil rights laws could

present a considerable challenge to efforts to defend the use of VAM in contexts where the result of VAM implementation has a disparate impact on minority educators.

Potential legal issues arising from a commercial law perspective on the use of VAM for educator evaluation

In addition to these central legal issues arising from federal statutes barring discrimination in employment that require evidence of validity and reliability, there are other possibilities for legal issues that might arise in educators' challenges to the use of VAM. Since support for VAM is based in large part on efforts to make public elementary and secondary education more business-like (Kupermintz, 2003), it is perfectly reasonable to expect that legal claims might arise based on precedents arising from the commercial and common law contexts. From a public policy perspective, these approaches are perhaps easy to anticipate: VAM arises from a business perspective on the functioning of schools and commercial vendors provide VAM data. When VAM causes harm to educators, it might be expected that they would argue that VAM is a bad product and its use is bad business. Some recent activity in and out of court in related teacher licensure litigation provides an illustration.

To obtain licensure in many states, educators are obligated to enter into a contractual relationship with a testing company to take the exam(s) mandated by the states in which they seek licensure (Melnick & Pullin, 2000). A federal multi-district class action case was brought on behalf of over 30,000 future teachers denied licensure due to a scoring error on PRAXIS, an Educational Testing Service (ETS) teacher licensure test. The future teachers asserted claims under four theories: (1) breach of contract for failure to fairly administer the test, correctly score the test and issue a correct score report; (2) negligence in failure to exercise reasonable care in the design, administration and scoring of the licensure test; (3) negligent misrepresentation of performance; and (4) violation of § 2 of the U.S. Sherman Anti-Trust Act for anticompetitive monopolistic practices in the teacher licensure testing market. The argument brought by the educators was, in short, that the test scoring error was the result of bad business practices. There was a quick, over \$11 million dollar settlement in favor of the plaintiffs (*In re Educational Testing Service PRAXIS Principles of Learning and Teaching: Grades 7-12 Litigation*, 2006, 2007).

Given the heavy involvement of vendors in both student tests and the creation of VAM metrics, new legal challenges similar to these, based on a business perspective on the use of tests and VAM, are possible. Given the validity and reliability issues that education researchers have raised about VAM, commercial claims similar to those used in the ETS litigation could well result if there is any question about the quality of the business practices and products of the vendors providing VAM or its underlying student test scores.

Potential Legal Issues on the Use of VAM by Students

While educators may wish to challenge the use of VAM data on grounds that it is scientifically indefensible under any of the legal provisions just discussed, students and parents and citizens acting on their behalf may chose the opposite strategy, embracing VAM data at face value and using low VAM scores as evidence to substantiate their own legal claims that schools are not meeting their obligations to provide education. This has not yet occurred, but as been at least contemplated by one prominent litigator who has pursued a long set of legal claims that the state of New York is failing to provide sufficient funding for education in violation of state laws (Rebell, 2011/2012).

Potential student civil rights claims

Since the mid 20th century, the nation has implemented policies intended to insure that minority, low income, female students and students with disabilities receive the education they deserve, free of discrimination. The Equal Protection and Due Process Clauses of the U.S. Constitution, Title VI of the Civil Rights Act, Title IX, the Equal Educational Opportunities Act, and the Individuals with Disabilities Education Act were all designed to improve access to educational opportunity. In addition, many state constitutions contain requirements that the state must create and fund a system of public education that provides students access to adequate, appropriate, or meaningful educational opportunity (Rebell, 2011/2012; Pullin, 2007). If students and their advocates feel that any of these legal provisions for access to educational opportunity are being denied, litigation can be mounted challenging the education provided by the state or local schools.

In disputes over alleged denial of educational opportunity to students, social science evidence is often utilized as important evidence relative to whether a denial of educational opportunity has occurred. Students and their advocates might seek to use VAM data and the low VAM scores of the educators who serve them as evidence these legal rights have been violated. In light of the limited social science of the validity and reliability of VAM data, would it be, or should it be, regarded as helpful evidence in these legal disputes?

Potential student federal statutory issues on access to educational opportunity

The Equal Educational Opportunities Act (EEOA) obligates schools receiving federal assistance to prohibit discrimination on the basis of race, color or national origin in public schools. The U.S. Supreme Court recently highlighted the obligation under the EEOA for states and localities to take “appropriate action” to provide and improve educational opportunities for these educationally disadvantaged populations (*Horne v. Flores*, 2009). Lower court judges applying the EEOA have described a three-part standard defining their obligations in enforcing the statute that focuses largely on the soundness of the educational theory behind a program, the adequacy of the implementation of the program, and evidence that the program improves educational achievement (*Castenada v. Pickard*, 1981). There has long been a strong argument that the use of student test score data as a measure of educational improvement falls far short of legal mandates that schools provide appropriate educational opportunity (Pullin, 2007; Rebell & Wolff, 2008).

Legal challenges to the use of student test scores as a measure of whether legal obligations to provide education have arisen under other federal statutes. Students have successfully asserted that Title VI of the Civil Rights Act of 1964 (42 U.S.C. § 2000d, 2010) bars programs and practices with a disproportionate impact on racial and ethnic minorities or perpetuates the effects of past unlawful discrimination (*Debra P. v. Turlington*, 1981). Some of the data used in VAM involve adjustments based upon the demographic characteristics of students and schools (Braun, 2005; Harris, 2011; NRC, 2010). Factors such as race/ethnicity, student socioeconomic status (usually measured by eligibility for free/reduced price lunch, a status that can switch dramatically from time to time, especially in the current economy), disability status, or English language proficiency can be taken into account. At first consideration, taking these factors into account makes sense given all that social science has demonstrated about the variability in achievement associated with differences in student backgrounds. However, the result of these statistical adjustments may mean that, in fact, the very nondiscrimination obligations imposed by civil rights statutes are now being violated by a metric that in essence grants latitude for lower levels of performance for minority students, students with disabilities and students from minority and low socio-economic status families.

Potential student use of VAM data in state constitutional right to education claims

Since the 1980's, students in low property wealth school districts have brought litigation under state constitutional provisions concerning the operation and funding of public schools. Most

recently, these cases have focused on whether appropriate or adequate education is being provided by a state. These often long-running and complicated cases seek to define the nature of a state's obligation to educate, determine how to assess whether adequate funding is being provided, and then measure whether the system is achieving the level of educational success the state constitution contemplated (Pullin, 2007). Some courts have looked at issues of student performance as an outcome measure for the success of the system, the adequacy of education (Pullin, 2007). Some state courts have taken issues of teacher quality into account in these cases. For example, in New York, teacher evaluations were one consideration in litigation over whether that state met its obligation under the state's constitution to provide a sound basic education (*Campaign for Fiscal Equity, Inc. v. State of New York*, 2002). See also a similar approach considering in part issues of quality in a case involving the North Carolina constitutional provisions on education (*Hoke County Board of Education v. State of North Carolina*, 2004).

An example of the role of test data as evidence in disputes over whether a state is meeting its obligations under the state's constitutional arose in Massachusetts. Here the constitutional provision was an obligation on the part of state officials to "cherish" education, a term when added to the state's constitution in the Eighteenth Century meant a duty to provide education (*Hancock v. Driscoll*, 2005; *McDuffy v. Secretary of Executive Office of Education*, 1993). The state's highest court ruled that the constitutional duty to cherish education created an obligation on the part of state officials to provide sufficient funding to prepare educated citizens to function successfully in society (*McDuffy v. Secretary of Executive Office of Education*, 1993). Several years after the state funded and implemented a massive education reform law, students from low wealth districts came back to the courts to argue that state was still failing to meet its obligations. Some of the evidence they presented were dramatic variations from district to district in student test scores on the statewide achievement test. Both the special judge appointed to present fact-finding in the dispute and the state's highest court judges appeared to take at face value the utility of student test performance as an indicator of the success of the state's school reform statute in addressing the constitutional obligation to educate (*Hancock*, 2005; Pullin, 2007). The final decision by the court focused on the statute's use of "uniform, objective performance and accountability measures for every public school student, teacher, administrator, school, and district in Massachusetts" (822 N.E. 2d 1134 at 1138). The continuing upward trajectory in student test scores was the evidence the court used to find the constitutional obligation to educate was being met (822 N.E. 2d 1134, 1150).

Given the importance of state constitutional disputes over the funding and functioning of a state's public schools, VAM data could be seen as an important source of evidence in addition to student test scores. Particularly in an era of rapidly dwindling public resources to support schools, the prospects for increased adequacy litigation are high. Advocates for students in these lawsuits might be well advised to carefully consider whether the validity and reliability of VAM data are sufficient to be relied upon in making their case to the courts. Similarly, policymakers pondering whether they were meeting their state constitutional obligations to educate should engage in the same considerations.

Conclusion

The large number of states that have adopted provisions for the use of value-added models (VAM) in state law coupled with the Race to the Top initiatives of the American Recovery and Reinvestment Act illustrate the widespread appeal of VAM. Many state and federal statutory provisions embrace these efforts to quantify student achievement growth over time and to attach high stakes consequences for educators and institutions when there is a failure to perform at the desired level. Whether judges will find VAM data useful or not is another matter, even in the context

of traditional judicial restraint in reviewing the decisions of education officials, particularly given the state and federal mandates for valid and reliable approaches to education reform.

In the broad contemporary public policy context for education reform, the desire for accountability and transparency in government, coupled with heavily financed criticisms of public school teachers and their unions, may mean that VAM initiatives will prevail. The concerns of education researchers about VAM, coupled with legal obligations for the validity and reliability of education and evaluation programs should require judges and education policymakers to take a closer look for future decision-making. At the same time, the social science research community should be generating substantial new and persuasive evidence about VAM and the validity and reliability of all of its potential uses. For public policymakers, there are strong reasons to suggest that high-stakes implementation of VAM is, at best, premature and, as a result, the potential for successful legal challenge to its use is high. The use of VAM as a policy tool for meaningful education improvement has considerable limitations, whether or not some judges might consider it legally defensible.

References

- American Educational Research Association, American Psychological Association, National Council on Measurement in Education (AERA/ APA/ NCME) (1999). *Standards for Educational and Psychological Testing*. Washington, D.C.: American Educational Research Association.
- American Educational Research Association (2002), *Position Statement of the American Educational Research Association on High-Stakes Testing in PreK-12 Education*. Washington D.C.: authors.
- American Recovery and Reinvestment Act (2009), Public Law 111-5, 123 Stat 115.
- Amrein-Beardsley, A. (2008). Methodological Concerns About the Education Value-Added Assessment System, *Educational Researcher*, 37: 2, pp. 65–75.
- Amrein-Beardsley, A., & Collins, C. (2012). The SAS Education Value-Added Assessment System (SAS® EVAAS®) in the Houston Independent School District (HISD): Intended and Unintended Consequences. *Education Policy Analysis Archives*, 20, 12. Retrieved from <http://epaa.asu.edu/ojs/article/view/1096>.
- Associated Press (2012, June 21). *NY legislature passes teacher evaluation disclosure bill*. As reported in *Education Week Online* June 25, 2012.
- Baker, E., Barton, P., Darling-Hammond, L., Haertel, E., Ladd, H., Linn, R., Ravitch, D., Rothstein, R., Shavelson, R. & Shepard, L. (2010). *Problems with the Use of Student Test Scores to Evaluate Teachers*. Economic Policy Institute. Retrieved at http://epi.3cdn.net/b9667271ee6c154195_t9m6iij8k.pdf
- Bd. of Regents v. Roth*, 408 U.S. 564 (1972).
- Braun, H. I. (2005). *Using Student Progress to Evaluate teachers: A Primer on Value-Added Models*. Princeton, N.J.: Policy Information Center, Educational Testing Service
- Campaign for Fiscal Equity, Inc. v. State of New York*, 744 N.Y.S.130 (2002).
- Castenada v. Pickard*, 648 F. 2d 989 (5th Cir. 1981).
- Civil Rights Act of 1964, 42 U.S.C. § 2000d (West 2010).
- Cizek, G. (1999). *Cheating on Tests: How To Do It, Detect It, and Prevent It*. Mahwah, N.J.: Lawrence Erlbaum.
- Cleveland Bd. of Education v. Loudermill*, 470 U.S. 532 (1985)
- Colorado Revised Statutes Annotated § 22-9 (2011).
- Contreras v. City of Los Angeles*, 656 F. 2d 1267 (9th Cir. 1981).

- Corcoran, S. P. (2010). *Can teachers be evaluated by their students' test scores? Should they be? The use of value-added measures of teacher effectiveness in policy and practice*. Providence, R.I.: Annenberg Institute for School Reform at Brown University. Retrieved at www.annenberginstitute.org.
- Council for Exceptional Children. (2012). *Position Statement on Special Education Teacher Evaluation*. Arlington, VA: Council for Exceptional Children.
- Dagley, D. and Veir, C. (2002). Subverting the academic abstention doctrine in teacher evaluation: How school reform legislation defeats itself, 2002 *Brigham Young University Education Law Journal* 123.
- Debra P. v. Turlington*, 644 F.2d 397, 404; (5th Cir. 1981), *aff'd*, 730 F. 2d. 1405 (11th Cir. 1983)
- Equal Educational Opportunities Act, 20 U.S.C. 1703 (2010).
- Equal Employment Opportunities Commission (EEOC), *Uniform Guidelines on Employee Selection Procedures*, 29 Code of Federal Regulations Part 1600
- Floden, R. E. (2012). Teacher value added as a measure of program quality: Interpret with caution. *Journal of Teacher Education*, 63(5), 356-360.
- Florida Statutes Annotated § 1012.34 2011.
- Gillespie v. State of Wis.*, 771 F. 2d. 1035 (7th Cir. 1985).
- Gordon, R., Kane, T. J., & Staiger, D. O. (2006). *Identifying effective teachers using performance on the job*. Washington, D.C.: The Brookings Institution.
- Guardians Association v. Civil Service Commission*, 630 F. 2d. 79 (2^d Cir. 1979).
- Griggs v. Duke Power Co.*, 401 U.S. 424 (1971).
- Haertel, E. H. (2006). Reliability. In R. L. Brennan (Ed.), *Educational Measurement* (4th ed., pp. 65-110). Westport, CT: American Council on Education and Praeger.
- Hanushek, E. A., & Rivkin, S. G. (2012). The distribution of teacher quality and implications for policy. *Annual Review of Economics*, 4, 131-157.
- Hancock v. Driscoll*, 2004 WL 877984 (Mass Super. 2004).
- Hancock v. Driscoll*, 443 Mass. 428, 822 N.E. 2d 1134 (2005).
- Harris, D. N. (2009). The policy issues and policy validity of value-added and other teacher quality measures. In D. H. Gitomer (Ed.), *Measurement issues and assessment for teaching quality* (pp. 99-130). Thousand Oaks, CA: Sage.
- Harris, D. (2011). *Value-added Measures in Education: What Every Educator Needs to Know*. Cambridge MA: Harvard Education Press.
- Herold, B. (2012). Philadelphia principal describes fallout of test-score inflation. *Education Week/NewsWorks*, from www.edweek.org/ew/articles/2012/11/28/13
- Herold, B., & Mezzacappa, D. (2012). New evidence: Philadelphia test cheating likely worse than suspected. *Education Week*, from http://www.edweek.org/ew/articles/2012/08/17/01pnbk_erasures.h32.html?tkn=PTTFR7TlsS37MTiwjAhbmWRjLybVwnxw%2FILI&print=1
- Hill, H. C., Kapitula, L., & Umland, K. (2010). A validity argument approach to evaluating teacher value-added scores. *American Educational Research Journal*, 48(3), 794-831.
- Hoke County Board of Education v. State of North Carolina*, 599 S.E.2d 365, 382-384 (N.C., 2004).
- Holdheide, L. R., Goe, L., Croft, A., & Reschly, D. J. (2012). *Challenges in Evaluating Special Education Teachers and English Language Learner Specialists*. Chicago, IL: National Comprehensive Center for Teacher Quality.
- Horne v. Flores*, 129 S. Ct. 2579 (2009).
- Illinois Compiled Statutes Annotated 5/24A-20 (West 2010).
- In re Educational Testing Service PRAXIS Principles of Learning and Teaching: Grades 7-12 Litigation*, 447 F. Supp. 2d 612 (E.D. La. 2006), 517 F. Supp. 2d 832 (E.D. La. 2007).

- Individuals with Disabilities Education Act, 20 U.S.C. 1600 et seq.
- Joint Committee on Standards for Education Evaluation (1994). *The Personnel Evaluation Standards* (The Joint Committee on Standards for Educational Evaluation & Sanders, J. R. eds), 2nd ed.
- Kane, M. (2006). Validation. In R. L. Brennan (Ed.), *Educational Measurement* (4th ed., pp. 17-64). Westport, CT: American Council on Education and Praeger.
- Kennedy, M. (ed.) (2002). *Teacher Assessment and the Quest for Teacher Quality*. San Francisco, CA: Jossey-Bass.
- Knight, S. L., Edmondson, J., Lloyd, G. M., Arbaugh, F., Jr., J. N., Whitney, A. E., et al. (2012). Examining the complexity of assessment and accountability in teacher education. *Journal of Teacher Education* 63(5), 301-303.
- Kupermintz, H. (2003). Teacher effects and teacher effectiveness: A validity investigation of the Tennessee Value Added Assessment System. *Educational Evaluation and Policy Analysis*, 25(3), 287-298.
- Lang, K. (2010). Measurement matters: Perspectives on education policy from an economist and school board member. *Journal of Economic Perspectives*, 24: 67, 168.
- Lindemann, B. and Grossman, P. (1996). *Employment Discrimination Law* (3d. ed.) Bureau of National Affairs (BNA).
- Los Angeles Times (2012a). Value-added in education. Retrieved at projects.latimes.com/value-added.
- Louisiana Revised Statutes Annotated § 17:3881 (2011).
- Massachusetts Federation of Teachers, AFT, AFL-CIO v. Board of Education, 436 Mass. 763 (2002).
- Maryland Code Annotated, Education § 6-202 (2011).
- McDuffy v. Secretary of Executive Office of Education, 415 Mass. 545 (1993).
- McGuinn, P. (2012). *The State of Teacher Evaluation Reform: State Agency Capacity and the Implementation of New Teacher-evaluation Systems*. Washington, D.C.: Center for American Progress.
- Michigan Compiled Laws Annotated § 380.1249 (2010).
- Melendez v. Illinois Bell Telephone, 79 F. 3d 661 (7th Cir 1996).
- Melnick, S. and Pullin, D. (2000, September/October). Can you take dictation? Prescribing teacher quality through testing. *Journal of Teacher Education*, 51(4).
- Mulgrew v. Board of Educ. of City School Dist. of City of New York, 31 Misc.3d 296, 919 N.Y.S.2d 786, 265 Ed. Law Rep. 1206, 2011 N.Y. Slip Op. 21030 (N.Y.Sup. Jan 10, 2011), aff'd 87 A.D.3d 506, 928 N.Y.S.2d 701, 270 Ed. Law Rep. 817, 40 Media L. Rep. 1505, 2011 N.Y. Slip Op. 06328 (N.Y.A.D. 1 Dept. Aug 25, 2011).
- National Council on Teacher Quality (NCTQ) (2012, January). *2011 State Teacher Policy Yearbook*. Washington, D.C.: author. Retrieved at www.nctq.org
- National Research Council (1999). *High Stakes: Testing for Tracking, Promotion, and Graduation* (Robert M. Hauser & Jay P. Heubert eds., National Academies Press, 1999).
- National Research Council (2001). *Testing Teacher Candidates: The Role of Licensure Tests in Improving Teacher Quality*. (Karen J. Mitchell, David Z. Robinson, Barbara S. Plake & Kaeli T. Knowles eds. Washington, D.C.: National Academies Press.
- National Research Council (2008). *Assessing Accomplished Teaching*. (Milton D. Hakel, Judith Anderson Koenig & Stuart W. Elliot eds., Washington, D.C.: National Academies Press 2008).
- National Research Council (2010). *Getting Value out of Value-added*. Washington D.C.: National Academies Press.

- National Research Council (2012). *Incentives and Test-based Accountability in Education* (M Hout and S. Elliot, eds. Washington, DC: National Academies Press.
- New York State United Teachers Association v. Board of Regents of the University of the State of New York*, 333 Misc. 3d 989, 929 N.Y.S.2d 699 (Supreme Court, Albany County, 2011)
- Newton, X., Darling-Hammond, L., Haertel, E., and Thomas, E. (2010, Sept.). Value-Added Modeling of Teacher Effectiveness: An Exploration of Stability across Models and Contexts. *Education policy analysis archives* retrieved at <http://epaa.asu.edu/ojs/article/view/810>
- No Child Left Behind Act of 2001 (2002), 20 U.S.C.A. 6301-7941
- Otterman, S. (2011). Regents move to improve test security amid complaints about academic cheating. *The New York Times*. Retrieved November 28, 2011, from www.nytimes.com
- Professional Standards Commission v. Denham*, (2001)
- Pullin, D. (2001). (2001, July). Key questions in implementing teacher testing and licensing. *Journal of Law and Education*. 30, 383-429.
- Pullin, D. (2007). Ensuring an adequate education: Opportunity to learn, law and social science. *Boston College Third World Law Journal*. 27:1, 83-130.
- Pullin, D. (2010). Judging teachers: The law of teacher dismissals. *Teacher assessment and the quest for teacher quality*. (M. Kennedy, ed.). San Francisco, CA: Jossey-Bass.
- Ready, D. D. (2013). Associations between student achievement and student learning: Implications for value-added school accountability models. *Educational Policy*, 27(1), 92-120.
- Rebell, M. (2011-2012). Safeguarding the right to a sound basic education in times of fiscal constraint. 75 *Albany Law Review* 1855.
- Rebell, M. and Wolff, J. (2008). *Moving Every Child Ahead: From NCLB Hype to Meaningful Educational Opportunity*. New York: Teachers College Press.
- Richardson v. Lamar Cnty Bd. of Educ.*, 729 F. Supp. 806, 820 (M.D. Ala. 1989).
- Rothstein, J. (2010). Teacher quality in educational production: Tracking, decay, and student Achievement, *Quarterly Journal of Economics* 125: 175
- Rothstein, J. (2009). Student sorting and bias in value added estimation: Selection on observables and unobservables, *Education finance and policy* 4: 537.
- Sawchuk, S. (2012). 'Value added' proves beneficial to teacher prep. *Education Week*, pp. 1, 20, from www.edweek.org/ew/articles/2012/02/22/21louisiana_ep-2.h31.html?tkn=VRVF2i06pG8p1lEuc2op6%2Frlw9Yc0%2BpQE%2BgG&intc=es&print=1
- Samuels, C. A. (2012). Four educators resign in Georgia cheating probe. *Education Week*, from <http://www.edweek.org/ew/articles/2012/01/11/15cheat.h31.html?tkn=VPSFRGdZ27CpsT0iJIn8K9dLmrMhLuV1ETAx&print=1>
- Scheelhaase v. Woodbury Central Community School District*, 488 F. 2d 237 (8th Cir. 1973), cert. den. 417 U.S. 969 (1974).
- Scoggins v. Bd. of Educ.*, 853 F. 2d 1472 (8th Cir. 1988).
- Severson, K. (2011, October 14). Georgia: Licenses suspended in cheating scandal. *The New York Times* 19.
- Society of Industrial and Organizational Psychologists (2003). *Principles for the validation and use of personnel selection procedures* (4th Ed.). Bowling Green, OH:autho. Retrieved at www.siop.org
- Song, J. (2012). Times sues L.A. Unified for teacher ratings. Retrieved November 1, 2012 from <http://latimesblogs.latimes.com/lanow/2012/10/times-sues-la-unified-for-teacher-ratings.html>

- St. Louis Teachers Union, Local 420, Am. Fed'n of Teachers, AFL-CIO v. Board of Educ. of the City of St. Louis*, 652 F. Supp. 425 (E.D. Mo. 1987).
- Superfine, B. M. (2008). *The courts and standards-based education reform*. New York, NY: Oxford University Press.
- Title VI of the Civil Rights Act of 1964, 42 U.S.C. 2000d.
- Title VII of the Civil Rights Act of 1964, 42 U.S.C. 2000e.
- U.S. v. State of South Carolina*, 445 F. Supp. 1094 (D.S.C. 1977) *aff'd mem. sub nom National Education Association v. South Carolina*, 434 U.S. 1026 (U.S.S.C. 1978).
- U.S. Department of Education (2009). 74 Fed. Reg. 59692 (Nov. 18, 2009).
- Utah Code Annotated § 53A-10-106 (2010).
- Vogell, H. (2011, July 26). Investigation into APS cheating finds unethical behavior across every level. *Atlanta Journal-Constitution*, retrieved from <http://www.ajc.com/news/news/local/investigation-into-aps-cheating-finds-unethical-be/nQJHG/>
- Waldman, S. (2012). N.Y. State acts to derail standardized test cheats. *Education Week/Times Union*. Retrieved November 23, 2012, from www.edweek.org
- Washington Revised Code Annotated § 28A.410 (2010).
- Washington v. Davis*, 426 U.S. 229 (1976).
- Winerip, M. (2011, November 6). In Tennessee, following the rules for evaluations off a cliff, *New York Times*, retrieved at www.nytimes.com/2011/11/07/education/tennessees-rules-on-teacher-evaluations-bring-frustration.
- Zirkel, P. (2003). Legal Boundaries for Performance Evaluation of Public School Professional Personnel, 172 *Education Law Reporter*.
- Zubrzycki, J. (2012). Cheating scandal lands ex-superintendent in prison. *Education Week*, from <http://www.edweek.org/ew/articles/2012/10/17/08jail.h32.html?tkn=LYOFL1bLjkE6vKS8mMBk0BNCPIhgBKRPQLMm&print=1>

About the Author

Diana Pullin

Boston College, Lynch School of Education and the School of Law

diana.pullin@bc.edu

Diana Pullin, JD., Ph.D. is Professor of Education Law and Public Policy at Boston College. The focus of her work is the impact of law on education practice and the impact of social science on the law. She has served as legal counsel for students, educators, and school systems in many different types of education disputes, particularly over high stakes uses of testing and represented the students in *Debra P. v. Turlington*, the landmark litigation over high school graduation testing. She has published numerous books, chapters, and articles on education law and public policy, educational and employment testing, educator quality, and individuals with disabilities. Professional standards of practice have also been a focus of her work; she is one of the co-authors of the 1999 AERA/APA/NCME *Standards on Educational and Psychological Testing* and she served as well for a number of years as a member of the Joint Committee on Standards for Educational Evaluation. She currently serves on the Board on Testing and Assessment of the National Research Council of the National Academy of Sciences and is Associate Editor of the interdisciplinary journal *Educational Policy*. She is a Fellow of the American Educational Research Association and a National Associate of the National Academy of Sciences.

About the Guest Editor and Assistant Guest Editors

Guest Editor

Dr. Audrey Amrein-Beardsley

Arizona State University

audrey.beardsley@asu.edu

Dr. Amrein-Beardsley is currently an Associate Professor in the Mary Lou Fulton Teachers College at Arizona State University. Audrey's research interests include educational policy, research methods, and more specifically, high-stakes tests and value-added measurements and systems. In addition, she researches aspects of teacher quality and teacher education. She is also the creator and host of a show titled Inside the Academy during which she interviews some of the top educational researchers in the academy. For more information please see: <http://insidetheacademy.asu.edu>.

Dr. Clarin Collins

Virginia G. Piper Charitable Trust

clarin.collins@asu.edu

Clarin Collins recently completed her Ph.D. in Educational Policy and Evaluation from Arizona State University, with an emphasis in research methods. Via her dissertation, she examined teachers' understandings of and experiences with the SAS Education Value-Added Assessment System (EVAAS) in the Houston Independent School District where it is used to evaluate teacher effectiveness. Clarin is currently a Research and Evaluation Officer at the Virginia G. Piper Charitable Trust in Phoenix.

Assistant Guest Editor

Dr. Sarah Polasky

Arizona State University

sarah.polasky@asu.edu

Dr. Sarah Polasky is the Value-Added Specialist for the Arizona Ready-for-Rigor Project, a Teacher Incentive Fund Grant, within the Mary Lou Fulton Teachers College. Her current research interests include the development and implementation of value-added measurements and systems using high-stakes test data, assessment in early childhood education, the use of alternative achievement (e.g., district benchmarks, formative assessments) and non-achievement (i.e., developmental) data for value-added analysis, as well as the impact of socioemotional and neurological development of young children on their short- and long-term academic achievement.

Assistant Guest Editor

Edward F. Sloat

Mary Lou Fulton Teachers College, Arizona State University; Dysart Unified School District, Surprise, Arizona

esloat@asu.edu

Mr. Sloat is currently employed as the Director of Research and Accountability at Dysart Unified School District located in Surprise, Arizona and a doctoral student in the Leadership and Innovation Program within the Mary Lou Fulton Teachers College, Arizona State University. Mr. Sloat has served as Deputy Associate Superintendent for Research and Evaluation within the Arizona Department of Education, the Director of Research, Planning, and Assessment for the Peoria (Arizona) Unified School District, and as Director of Research and Assessment at the Glendale (Arizona) Elementary School District. He regularly contributes to state technical and policy working/advisory groups concerning assessment design and accountability systems and is past President of the Arizona Education Research Organization. Mr. Sloat holds a Master's Degree in Applied Economics from the University of Arizona, concentrating in econometric methods and management information systems. His academic interests focus on value-added modeling, education accountability and evaluation systems, data-driven instructional planning, applications of measurement theory, and research methods.

SPECIAL ISSUE

Value-Added: What America's Policymakers Need to Know and Understand

education policy analysis archives

Volume 21 Number 6

January 29th, 2013

ISSN 1068-2341



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-sa/3.0/>. All other uses must be approved by the author(s) or **EPAA**. **EPAA** is published by the Mary Lou Fulton Institute and Graduate School of Education at Arizona State University. Articles are indexed in CIRC (Clasificación Integrada de Revistas Científicas, Spain), DIALNET (Spain), [Directory of Open Access Journals](#), EBSCO

Education Research Complete, ERIC, Education Full Text (H.W. Wilson), QUALIS A2 (Brazil), SCImago Journal Rank; SCOPUS, SOKOLAR (China).

Please contribute commentaries at <http://epaa.info/wordpress/> and send errata notes to Gustavo E. Fischman fischman@asu.edu

Join EPAA's Facebook community at <https://www.facebook.com/EPAAAAPE> and **Twitter feed** @epaa_aape.

education policy analysis archives
editorial board

Editor **Gustavo E. Fischman** (Arizona State University)

Associate Editors: **David R. Garcia** (Arizona State University), **Stephen Lawton** (Arizona State University)

Rick Mintrop, (University of California, Berkeley) **Jeanne M. Powers** (Arizona State University)

Jessica Allen University of Colorado, Boulder

Gary Anderson New York University

Michael W. Apple University of Wisconsin, Madison

Angela Arzubiaga Arizona State University

David C. Berliner Arizona State University

Robert Bickel Marshall University

Henry Braun Boston College

Eric Camburn University of Wisconsin, Madison

Wendy C. Chi* University of Colorado, Boulder

Casey Cobb University of Connecticut

Arnold Danzig Arizona State University

Antonia Darder University of Illinois, Urbana-Champaign

Linda Darling-Hammond Stanford University

Chad d'Entremont Strategies for Children

John Diamond Harvard University

Tara Donahue Learning Point Associates

Sherman Dorn University of South Florida

Christopher Joseph Frey Bowling Green State University

Melissa Lynn Freeman* Adams State College

Amy Garrett Dikkers University of Minnesota

Gene V Glass Arizona State University

Ronald Glass University of California, Santa Cruz

Harvey Goldstein Bristol University

Jacob P. K. Gross Indiana University

Eric M. Haas WestEd

Kimberly Joy Howard* University of Southern California

Aimee Howley Ohio University

Craig Howley Ohio University

Steve Klees University of Maryland

Jackyung Lee SUNY Buffalo

Christopher Lubienski University of Illinois, Urbana-Champaign

Sarah Lubienski University of Illinois, Urbana-Champaign

Samuel R. Lucas University of California, Berkeley

Maria Martinez-Coslo University of Texas, Arlington

William Mathis University of Colorado, Boulder

Tristan McCowan Institute of Education, London

Heinrich Mintrop University of California, Berkeley

Michele S. Moses University of Colorado, Boulder

Julianne Moss University of Melbourne

Sharon Nichols University of Texas, San Antonio

Noga O'Connor University of Iowa

João Paraskveva University of Massachusetts, Dartmouth

Laurence Parker University of Illinois, Urbana-Champaign

Susan L. Robertson Bristol University

John Rogers University of California, Los Angeles

A. G. Rud Purdue University

Felicia C. Sanders The Pennsylvania State University

Janelle Scott University of California, Berkeley

Kimberly Scott Arizona State University

Dorothy Shipps Baruch College/CUNY

Maria Teresa Tatto Michigan State University

Larisa Warhol University of Connecticut

Cally Waite Social Science Research Council

John Weathers University of Colorado, Colorado Springs

Kevin Welner University of Colorado, Boulder

Ed Wiley University of Colorado, Boulder

Terrence G. Wiley Arizona State University

John Willinsky Stanford University

Kyo Yamashiro University of California, Los Angeles

* Members of the New Scholars Board

archivos analíticos de políticas educativas consejo editorial

Editor: **Gustavo E. Fischman** (Arizona State University)

Editores. Asociados **Alejandro Canales** (UNAM) y **Jesús Romero Morante** (Universidad de Cantabria)

Armando Alcántara Santuario Instituto de Investigaciones sobre la Universidad y la Educación, UNAM México

Claudio Almonacid Universidad Metropolitana de Ciencias de la Educación, Chile

Pilar Arnaiz Sánchez Universidad de Murcia, España

Xavier Besalú Costa Universitat de Girona, España

Jose Joaquín Brunner Universidad Diego Portales, Chile

Damián Canales Sánchez Instituto Nacional para la Evaluación de la Educación, México

María Caridad García Universidad Católica del Norte, Chile

Raimundo Cuesta Fernández IES Fray Luis de León, España

Marco Antonio Delgado Fuentes Universidad Iberoamericana, México

Inés Dussel FLACSO, Argentina

Rafael Feito Alonso Universidad Complutense de Madrid, España

Pedro Flores Crespo Universidad Iberoamericana, México

Verónica García Martínez Universidad Juárez Autónoma de Tabasco, México

Francisco F. García Pérez Universidad de Sevilla, España

Edna Luna Serrano Universidad Autónoma de Baja California, México

Alma Maldonado Departamento de Investigaciones Educativas, Centro de Investigación y de Estudios Avanzados, México

Alejandro Márquez Jiménez Instituto de Investigaciones sobre la Universidad y la Educación, UNAM México

José Felipe Martínez Fernández University of California Los Angeles, USA

Fanni Muñoz Pontificia Universidad Católica de Perú

Imanol Ordorika Instituto de Investigaciones Económicas – UNAM, México

Maria Cristina Parra Sandoval Universidad de Zulia, Venezuela

Miguel A. Pereyra Universidad de Granada, España

Monica Pini Universidad Nacional de San Martín, Argentina

Paula Razquin UNESCO, Francia

Ignacio Rivas Flores Universidad de Málaga, España

Daniel Schugurensky Universidad de Toronto-Ontario Institute of Studies in Education, Canadá

Orlando Pulido Chaves Universidad Pedagógica Nacional, Colombia

José Gregorio Rodríguez Universidad Nacional de Colombia

Miriam Rodríguez Vargas Universidad Autónoma de Tamaulipas, México

Mario Rueda Beltrán Instituto de Investigaciones sobre la Universidad y la Educación, UNAM México

José Luis San Fabián Maroto Universidad de Oviedo, España

Yengny Marisol Silva Laya Universidad Iberoamericana, México

Aida Terrón Bañuelos Universidad de Oviedo, España

Jurjo Torres Santomé Universidad de la Coruña, España

Antoni Verger Planells University of Amsterdam, Holanda

Mario Yapu Universidad Para la Investigación Estratégica, Bolivia

arquivos analíticos de políticas educativas
conselho editorial

Editor: **Gustavo E. Fischman** (Arizona State University)
Editores Associados: **Rosa Maria Bueno Fisher** e **Luis A. Gandin**
(Universidade Federal do Rio Grande do Sul)

Dalila Andrade de Oliveira Universidade Federal de Minas Gerais, Brasil
Paulo Carrano Universidade Federal Fluminense, Brasil
Alicia Maria Catalano de Bonamino Pontifícia Universidade Católica-Rio, Brasil
Fabiana de Amorim Marcello Universidade Luterana do Brasil, Canoas, Brasil
Alexandre Fernandez Vaz Universidade Federal de Santa Catarina, Brasil
Gaudêncio Frigotto Universidade do Estado do Rio de Janeiro, Brasil
Alfredo M Gomes Universidade Federal de Pernambuco, Brasil
Petronilha Beatriz Gonçalves e Silva Universidade Federal de São Carlos, Brasil
Nadja Herman Pontifícia Universidade Católica –Rio Grande do Sul, Brasil
José Machado Pais Instituto de Ciências Sociais da Universidade de Lisboa, Portugal
Wenceslao Machado de Oliveira Jr. Universidade Estadual de Campinas, Brasil

Jefferson Mainardes Universidade Estadual de Ponta Grossa, Brasil
Luciano Mendes de Faria Filho Universidade Federal de Minas Gerais, Brasil
Lia Raquel Moreira Oliveira Universidade do Minho, Portugal
Belmira Oliveira Bueno Universidade de São Paulo, Brasil
Antônio Teodoro Universidade Lusófona, Portugal
Pia L. Wong California State University Sacramento, U.S.A
Sandra Regina Sales Universidade Federal Rural do Rio de Janeiro, Brasil
Elba Siqueira Sá Barreto [Fundação Carlos Chagas](#), Brasil
Manuela Terrasêca Universidade do Porto, Portugal
Robert Verhine Universidade Federal da Bahia, Brasil
Antônio A. S. Zuin Universidade Federal de São Carlos, Brasil