



Education Policy Analysis
Archives/Archivos Analíticos de Políticas
Educativas

ISSN: 1068-2341

EPAA@asu.edu

Arizona State University
Estados Unidos

Pizmony-Levy, Oren; Woolsey, Ashley
Politics of Education and Teachers' Support for High-Stakes Teacher Accountability
Policies
Education Policy Analysis Archives/Archivos Analíticos de Políticas Educativas, vol. 25,
2017, pp. 1-28
Arizona State University
Arizona, Estados Unidos

Available in: <http://www.redalyc.org/articulo.oa?id=275050047118>

- 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



Politics of Education and Teachers' Support for High-Stakes Teacher Accountability Policies

Oren Pizmony-Levy



Ashley Woolsey

Teachers College, Columbia University
United States

Citation: Pizmony-Levy, O., & Woolsey, A. (2017). Politics of education and teachers' support for high-stakes teacher accountability policies. *Education Policy Analysis Archives*, 25(87).
<http://dx.doi.org/10.14507/epaa.25.2892> This article is part of the special issue, *Global Perspectives on High-Stakes Teacher Accountability Policies*, guest edited by Jessica Holloway, Tore Bernt Sorensen, and Antoni Verger.

Abstract: Although educators are at the center of contentious high-stakes teacher accountability policies, we know very little about their attitudes toward these policies. This research gap is unfortunate because teachers are considered key actors in successful implementation of educational reforms. To what extent do the politics that accompany the introduction of high-stakes teacher accountability policies affect teachers' support for the policies themselves? To address this gap, we used data from an experimental survey of teachers in New Jersey ($n=444$), where a new reform—Teacher Effectiveness and Accountability for Children of New Jersey Act (TEACHNJ)—was signed into law in 2012 and implemented shortly thereafter. The cornerstone of the reform is a new

evaluation system that ties student performance on standardized tests to teachers' evaluations. We found that the majority of teachers in our study oppose the new evaluation system. Teachers' attitudes were shaped by the politics of the key actors advocating for the policy, perceptions of implementation efforts, and beliefs in the potential outcome of the policy. Open-ended responses indicated that teachers question the validity of the evaluation system and are concerned about the negative intended and unintended consequences of the system. We conclude this paper by discussing the implications of these findings for policy studies and policymaking.

Keywords: Accountability; education reform; politics of education; teachers; teacher evaluation system; teachers' attitudes; experimental design; New Jersey

Política de educación y profesores en apoyo a las políticas de rendición de cuentas

Resumen: Aunque los educadores están en el centro de las polémicas políticas de responsabilidad de los maestros, sabemos muy poco acerca de sus actitudes hacia estas políticas. Esta brecha de investigación es desafortunada porque los maestros son considerados actores clave en la implementación exitosa de las reformas educativas. ¿Hasta qué punto la política que acompaña a la introducción de políticas de responsabilidad de maestros de alto riesgo afecta el apoyo de los profesores a las políticas mismas? Para abordar esta laguna, utilizamos datos de una encuesta experimental de maestros en New Jersey ($n = 444$), donde una nueva reforma -La Ley de Eficacia y Rendición de Cuentas de los Maestros de Niños de New Jersey (TEACHNJ) poco después. La piedra angular de la reforma es un nuevo sistema de evaluación que vincula el desempeño de los estudiantes en las pruebas estandarizadas con las evaluaciones de los maestros. Encontramos que la mayoría de los maestros en nuestro estudio se oponen al nuevo sistema de evaluación. Las actitudes de los maestros fueron moldeadas por la política de los actores claves que abogaban por la política, las percepciones de los esfuerzos de implementación y las creencias en el resultado potencial de la política. Las respuestas abiertas indican que los docentes cuestionan la validez del sistema de evaluación y están preocupados por las consecuencias negativas y no intencionales del sistema. Concluimos este trabajo discutiendo las implicaciones de estos hallazgos para los estudios de políticas y la formulación de políticas.

Palabras clave: Rendición de cuentas; reforma educativa; política educativa; maestros; sistema de evaluación docente; actitudes de los maestros; diseño experimental; New Jersey

Política educativa e professores em apoio às políticas de prestação de contas

Resumo: Embora os educadores estejam no centro das controvertidas políticas de responsabilidade dos professores, sabemos muito pouco sobre suas atitudes em relação a essas políticas. Esta lacuna de pesquisa é lamentável porque os professores são considerados protagonistas da implementação bem-sucedida de reformas educacionais. Em que medida a política que acompanha a introdução de políticas de responsabilidade de professores de alto risco afeta o apoio dos professores às próprias políticas? Para abordar esta lacuna, utilizamos dados de uma Pesquisa de Professores Experimentais em New Jersey ($n = 444$), onde uma nova reforma - Lei de Responsabilidade e Responsabilidade dos Professores de Nova Jersey (TEACHNJ) pouco depois. A pedra angular da reforma é um novo sistema de avaliação que relaciona o desempenho do aluno em testes padronizados com avaliações de professores. Achamos que a maioria dos professores em nosso estudo se opõe ao novo sistema de avaliação. As atitudes dos professores foram moldadas pela política dos principais atores que defendiam políticas, percepções de

esforços de implementação e crenças sobre o resultado potencial da política. As respostas abertas indicam que os professores questionam a validade do sistema de avaliação e estão preocupados com as consequências negativas e não intencionais do sistema. Concluímos este artigo discutindo as implicações desses achados para estudos de política e elaboração de políticas.

Palavras-chave: Prestação de contas; reforma educacional; política educacional; professores; sistema de avaliação de professores; atitudes docentes; design experimental; New Jersey

Introduction

High-stakes teacher accountability policies are the new frontier in the application of private-sector discourses and practices to the field of education. Often called New Public Management [NPM], these approaches rely on privatization, choice, and accountability to address educational problems (Gewirtz & Ball, 2000). Major components of high-stakes teacher accountability policies include (a) teacher evaluation that is based on how much academic growth students experience over the course of the school year, and (b) teacher evaluation that has implications for professional development, compensation and benefits, and tenure (see reports by Educational Testing Service [Braun, 2005] and RAND Corporation [Steele, Hamilton & Stecher, 2010], and see review by Baker, Oluwole & Green, 2013).

To some degree, these high-stakes teacher accountability policies draw on opportunities created by international large-scale assessments (ILSA) as well as national assessments (Elmore, Abelman, & Fuhrman, 1996; Holloway-Libell & Collins, 2014). In many countries, poor performance on ILSA was blamed on teachers and teacher preparation programs (Figazzolo, 2009; Pizmony-Levy, under review). In turn, this interpretation generated a discursive opportunity structure (Koopmans, 1999) that is open for high-stakes teacher accountability policies as a panacea.¹ Furthermore, the creation of an accountability infrastructure with an abundance of student-level data created the illusion that high-stakes teacher accountability policies are not only needed, but also possible (e.g., through recycling and reapportioning data collected for another purpose).

High-stakes teacher accountability policies are contentious. On the one hand, policymakers and practice communities (e.g., testing agencies) advocate enthusiastically for the implementation of teacher evaluation systems that can bring about improvements in instructional practices (e.g., Tucker & Stronge, 2005; Weisberg et al., 2009). On the other hand, scholars still debate whether these evaluation systems are scientifically and technically sound (AERA *Statement on Use of Value-Added Models (VAM) for the Evaluation of Educators and Educator Preparation Programs*, 2015). This is not only a debate among the “elite,” but the broader American public is also split about the question of basing part of teachers’ evaluations and salaries on how much their students learn. The 2016 Education Next survey shows that slightly more than half of Americans (52%) favored this approach, with more support among Republicans (57%) than Democrats (50%) (Peterson, Henderson, West & Barrows, 2016). Further, Phi Delta Kappa (PDK) surveys show a significant decrease in public support for state policies requiring that teacher evaluations include how well a teacher’s students perform on standardized tests. The 2015 PDK survey shows slightly more than two-fifths of Americans (43%) favored this approach (Richardson & Bushaw, 2015).

¹ Koopmans (1999) introduces the term “discursive opportunity structure” to identify ideas in the larger political culture that are believed to “make sense” or to be “realistic” and “legitimate.”

The opt-out movement in the US, in which parents refuse to have their school-aged children take federally mandated educational assessments, provides additional evidence to the contentious nature of high-stakes teacher accountability policies (Mittra, Mann & Hlavacik, 2016). A recent survey of activists in the opt-out movement estimated that teachers represent about half of the movement's social base (Pizmony-Levy & Green Saraisky, 2016). When asked to rank the main reason for their participation in the movement, more than one-third of activists (36.9%) indicated that they "oppose using students' performance on standardized tests to evaluate teachers." Moreover, teachers were more likely than others to mention this reason as a motivation for activism (44% versus 27%).

Although educators are at the center of high-stakes teacher accountability policies, we know very little about how they actually feel about these policies (Ballou & Podgursky, 1993; Goldhaber, DeArmond, & DeBurgomaster, 2010). The 2016 Education Next survey, for example, shows that less than one-fifth of teachers (19%) favored policies related to high-stakes teacher accountability. Furthermore, there is little empirical analysis about the factors that shape teachers' support of these policies. Do teachers oppose these policies because they disagree with their principles? Or do teachers oppose these policies because of their experience with the implementation efforts? In fact, we have very little systematic knowledge about what shapes teachers' views toward any kind of school reform (for an exception, see Addi-Raccah, 2012).

Understanding educators' views of high-stakes teacher accountability policies is important for at least two reasons. First, teachers are the centerpiece of educational change. Working on the front lines of the education system, teachers translate abstract ideas (articulated in policy documents) into reality (Lipsky, 2010; Spillane, 2004). Second, because teachers are central to the implementation and mediation of top-down policies, the extent to which they support a reform may affect its outcome (Fullan, 2001).

In this article, we describe a study of teachers' engagement with high-stakes teacher accountability policies. We conducted a survey of teachers in the state of New Jersey ($n=444$), where a new teacher evaluation policy was launched in 2012. The survey asked teachers about their (1) endorsement of underlying principles of the policy, (2) self-reported knowledge of key concepts in the policy, (3) perception of the potential impact of the policy, (4) evaluation of the implementation efforts, and (5) support for the policy. The survey also included an embedded survey experiment to assess the impact of the political environment on teachers' support for the new teacher evaluation policy. Half of the respondents read a description of the policy that mentioned New Jersey Governor Chris Christie (who was the key political actor behind the policy) and the other half read a description of the policy that mentioned the New Jersey Legislature. An open-ended question asked teachers to reflect on the issues raised in the survey. We use these data to address three related questions:

Research questions 1: What are teachers' views toward the new teacher evaluation policy?

Research question 2: Does framing the new teacher evaluation policy in the context of education politics affect teachers' support for the policy?

Research question 3: What factors explain teachers' support (or lack thereof) for the new teacher evaluation policy?

Conceptual Framework

Scholars have agreed that teachers' engagement is key for successful policy implementation (Fullan, 2001; Lipsky, 2010; Spillane, 2004). It is reasonable to assume that teachers' engagement would be even more important in the context of high-stake teacher accountability policies. After all, these policies pose sanctions and rewards on teachers based on student performance on standardized tests. To date, however, scholars have mostly overlooked teachers' views toward these policies (for exceptions see Bogart, 2013; Bridich, 2015; Collins, 2014; for research on teachers' support of merit pay, see Ballou & Podgursky, 1993; Goldhaber, DeArmond & DeBurgomaster, 2010).

Our study is informed by two lines of research. The first is concerned with teachers' engagement with policy and policy implementation. Scholars have pointed to four factors that enhance teachers' engagement with top-down educational change: policy "buy-in" and consensus over the principles of policy (Datnow, 2000; Datnow & Castellano, 2000; Fullan, 1991, 1993; Kim & Youngs, 2016; Sarason, 1990, 1996), good experience with implementation efforts (method, type, and pace; Desimone, 2002), knowledge and understanding of the policy (Spillane, Reiser, & Reimer, 2002), and beliefs in the efficacy of the policy (Desimone, 2002; Datnow & Castellano, 2000). Most of the research that has been done is based on qualitative methods, which limits the possibility to test these four factors simultaneously. In this study, we use a quantitative research design to address this gap.

The second line of research focuses on teachers' views toward NPM inspired policies and reforms. Recent research on teachers' engagement with new evaluation and accountability policies shows a complex pattern. Bogart (2013), for example, found that although teachers in Tennessee doubt the impact of new evaluation systems on schools and classrooms, they also reported an increase in planning time and in preparation of lessons that focus on higher-order thinking skills. Bridich (2015) found that the majority of teachers and administrators in one school district in Colorado believed that the new teacher evaluation system can improve teaching and learning. But, only one-in-seven respondents actually supported the legislation behind the education reform. Collins (2014) found that teachers in the Southern region of the United States not only see little value in new assessments, but they also report on negative unintended consequences of these assessments.

A review of the factors related to teachers' attitudes toward NPM suggest that only a few personal and organizational characteristics have been studied (Belfield & Wooten, 2003). Hess, Maranto, & Milliman (2000), for example, found that more experienced teachers, those who identified as Democrats, those who majored in education, or those who had never worked in a competitive educational environment were more likely to oppose public school choice. Further, drawing from a micro-political approach, Addi-Raccah (2012) found that Israeli teachers holding internal leadership positions or those who reported feeling empowered by their principals were more supportive of privatization.

A common theme in both lines of research is the role of politics and power in shaping teachers' engagement with policies. Given the heated debate surrounding high-stake teacher accountability policies, we hypothesize that the political environment in which these policies are discussed and introduced will affect teachers' support. Indeed, Mortimore and Mortimore's (1998) analysis of the British education system concluded that the public discourse characterized with "anti-teacher" sentiments and scandalization of student performance on standardized tests had negative consequences on policies intended to improve teaching and learning.

In the context of our case, Governor Christie, known for his harsh, real “New Jersey talk,” has repeatedly and publicly disparaged any organized teachers efforts to critique the policy. Therefore, in order to further Mortimore and Mortimore’s (1998) analysis we examine the potential impact of the political environment on the perception of the TEACHNJ policy. To do this we draw from the literature discussing frames (Benford & Snow, 2000) and framing effects (for summaries, see Levy 2002; Levy 2003). Frames serve as thought organizers and can alter an individual’s response in a survey even though the items that are framed differently may be logically equivalent. Equivalency framing usually involves “casting the same information in either a positive or negative light” (Druckman, 2004, p. 671). For example, a frame that mentions 95% employment as opposed to one that refers to 5% unemployment, although substantively equivalent in content, might elicit divergent opinions as to the state of the labor force and economy (example taken from Druckman, 2004). Equivalency frames involve alternative descriptions of a qualitatively similar phenomenon (Druckman, 2001), such as the frames used when describing a new teacher evaluation system as either “proposed by the Legislature” or “proposed by the Governor Christie.”

In this article, we systematically examine teachers’ views toward the high-stakes teacher accountability policy (TEACHNJ) and its implementation platform (Achieve NJ). We show a gap between teachers’ endorsement of policy principles (i.e., policy alignment) and teachers’ view of the specific policy. Drawing on an embedded survey experiment, we demonstrate the important role of the political environment on teachers’ support for a new teacher evaluation policy. We also demonstrate the role that teachers’ beliefs and attitudes play in the efficacy of the policy and its implementation. Finally, we use open-ended comments to shed light on potential reasons behind the statistical patterns. The next section provides background for the study setting: the state of New Jersey.

Study Setting: New Jersey

With over 8.5 million people, New Jersey is the country’s most densely populated state per square mile, with 17.3% of the population being school-age children 5 to 18 years old (U.S. Census Bureau, 2010). As of 2008, the New Jersey educational system was the 11th largest educational system in the United States (National Education Association, 2011). Various reports have suggested that New Jersey offers high-quality public education. The Editorial Projects in Education Research Center (2010), for example, ranked New Jersey seventh overall in the United States based on students’ chance for success, overall K-12 achievement, standards, assessments and accountability, and school finances. New Jersey has traditionally outranked its peers with its above average Advanced Placement exam test scores (O’Neill, 2011). New Jersey has consistently outperformed 90% of the states in fourth- and eighth-grade reading and mathematics assessments as determined by the National Assessment of Educational Progress (NAEP; National Center for Educational Statistics, 2011).

This positive view of the New Jersey educational system was challenged in 2009 with the release of a 12th-grade assessment by NAEP. While New Jersey outscored all other participating states (only 11 states volunteered to participate in this assessment) in mathematics, the state matched the national reading scores. In addition, this assessment pointed—again—to a wide achievement gap between whites and minority students. The release of the report sparked a local policy “shock.” In an interview with local media, then Acting Education Commissioner Rochelle Hendricks said, “While these results show that our seniors score well when compared to other participating states in math, the bigger picture is clear in that we must do better” (Mascarenhas, 2010).

Despite New Jersey's comparatively successful educational system, the 2009 gubernatorial candidate Christopher Christie, a Republican, ran on a platform arguing that "Nationally, academic achievement has been flat for many years... New Jersey is suffering from the same stagnation when we should be pushing for greater student achievement and academic excellence" (Rundquist, 2009).² With that belief, upon taking office in January 2010, Governor Christie began a series of education reforms, including "merit pay, changes to tenure, completing a statewide data system that tracks student achievement, forming of a teacher evaluation task force, creating the designation of master teacher and allowing alternate route certification for principals" (Fleisher, 2010). On the issue of tenure and merit, Governor Christie was quoted in the same Fleisher (2010) article: "Tenure has become a job guarantee regardless of performance or success. Tenure has become the sclerosis that coats the veins of our school system [...] Any type of compensation that allows for anything but merit—gone." This, coupled with contentious interactions with his Democratic opponent who supported the New Jersey National Education Association (NJEA), created a hostile political environment for education reform. Governor Christie has consistently criticized the NJEA, calling the union "a group of political thugs" (Blackburn, 2011) and refusing to meet with the NJEA, arguing that "Frequently, the leadership of the NJEA has been a strong advocate for the status quo, whether the status quo is succeeding or failing" (Springer, 2009). The NJEA has not hesitated in launching media campaigns against the governor. In the same news article, Fleisher quoted NJEA President, Barbara Keshishian, who said, "He is proposing reforms that are not based on good educational research or practice. What he proposes—an over-reliance on student test scores to make critical decisions from compensation to employment—is fatally flawed" (Fleisher, 2010).

In his first year in office, Governor Christie reduced public employee pensions and benefits to cut the state deficit. In addition, the governor placed a 2% cap on property taxes, which in turn capped school budgets that are funded through town property taxes. At the same time, 475 million dollars were cut in school aid, and further budgets were redirected to expand school voucher programs and support charter schools, forcing districts to reduce budgets and staff. All these reforms were introduced with little—if any—consultation with the NJEA. These actions and the overall approach were interpreted by NJEA as a personal attack from Christie against educators and public education (NJEA, 2013).

Subsequently, Governor Christie introduced the flagship of his educational reform: Teacher Effectiveness and Accountability for Children of New Jersey Act (TEACHNJ). TEACHNJ's theory of action is that improving educator effectiveness through evaluation and feedback will improve student performance (New Jersey Department of Education, 2010). The cornerstone of the TEACHNJ reform is a new teacher evaluation system, which is currently in its fifth year of implementation in schools throughout the state. Despite the reform's positive language of improving teacher effectiveness and student learning, the NJEA has met it with resistance and animosity (NJEA, 2013). Further, although the Act passed the state legislature unanimously, public discourse highlighted the important role of the governor in advancing this agenda (see for example, Rizzo, 2012). The question remains: To what extent did the politics of introducing and advocating for TEACHNJ affect teachers' support?

TEACHNJ was developed as part of a federal grant called Race To The Top (RTTT). The funding received by state from the RTTT grant was based on the ability of the New Jersey Department of Education (NJDOE) to meet the deadlines established in its proposal, which placed TEACHNJ implementation on a fast track. The newly developed evaluation system consists of four categories of effectiveness (ineffective, partially effective, effective, and highly effective) based on

² For a critical review of education politics in New Jersey during Governor Christie's term, see Murphy, Strothers, & Lugg (2017).

multiple measures that are used to determine tenure. This drastically altered the former system of evaluation and tenure in the state, which used a binary measure of effectiveness, based on a single measure of educator practice and only required observations of non-tenured staff. Prior to TEACHNJ, teachers with tenure were not required to be observed by administrators and the removal of ineffective teachers often took years and hundreds of thousands of dollars, which discouraged firings. The new law made the revocation of tenure easier with automatic firing after two consecutive years of ineffective ratings and a streamlined tenure due process. TEACHNJ features numerous changes in evaluation structure, expectation, and support and is part of an educational reform movement in New Jersey focused on increasing student performance through effective educator evaluations that provide recognition, feedback, and support to education professionals (New Jersey Department of Education, 2013).

To implement TEACHNJ, the state developed AchieveNJ—a teacher evaluation and support system. Under AchieveNJ the state DOE has made significant efforts to provide communication and support to teachers through the creation of informational guides, presentations, forms, templates, etc. AchieveNJ was designed according to the following guiding principles:

1. Educator effectiveness can and should be measured to ensure our students have the best teachers in the classroom.
 2. Evaluations should always be based on multiple measures that include both learning outcomes and effective practice.
 3. Timely feedback and high-quality professional development, tied to evaluations, are essential to help educators improve.
 4. Evaluation and support systems should be developed with significant input from educators.
 5. Tenure and other forms of recognition should be based on effectiveness.
- (NJDOE, 2013)

As we discuss in the data and methods section, we draw on these statements to evaluate teachers' alignment with the underlying principles of the policy.

The new teacher evaluation system, mandated by TEACHNJ, is based on up to three components of assessment used to evaluate teacher performance. First, the median *Student Growth Percentile* (SGP) represents the median score of the growth that the individual students make from one year to the next on a statewide assessment. New Jersey adopted the Partnership for Assessment of Readiness for College and Careers PARCC. The SGP only applies to fourth- to eighth-grade language arts teachers and fourth- to seventh-grade mathematics teachers (hereafter “tested subject teachers”).³ This design might lead to differences in support for TEACHNJ between those subject teachers who were tested and those who are teaching non-tested subjects or non-tested grades. Second, *Student Growth Objectives* (SGO) are academic goals with measurable assessments designed by the teacher, with assistance from the principal and/or the supervisor. All teachers are required to develop SGOs. Third, *Teacher Practice* is based on observations of instruction. The state of New Jersey has approved numerous observation instruments (e.g., the Charlotte Danielson’s Framework for Teaching, and the McREL Teacher Evaluation System). Taken together, these three components of the evaluation system are weighted and calculated to reflect a teacher’s overall effectiveness on a scale from one to four, ineffective (1), partially ineffective (2), effective (3), and highly effective (4).

³ The SGP model is not a VAM by traditional standards and definitions, mainly because the SGP model does not use as many sophisticated control variables.

(NJDOE, 2013). The goal of this evaluation system is to redefine the teaching profession based on quantitative terms with the hopes of improving student outcomes by improving teacher effectiveness.

Data and Methods

Our analysis is based on the New Jersey Teacher Engagement with Education Reform Study, an original survey conducted by the second author (see Woolsey, 2014). The primary goal of this survey was to collect information from teachers about their evaluation of the implementation of TEACHNJ and to document their professional experiences in the context of a large education reform. The survey was conducted in the winter of 2014, during the first academic year after the implementation of TEACHNJ. Because of restricted resources, the study is based on a non-probability sample of 481 teachers.

To capture the wide swath of teachers across the state, data collection efforts included three recruitment strategies. First, we asked all public-school superintendents in the state of New Jersey (with correct contact information in the state DOE directory) to share the link to the survey with teachers. This strategy yielded the majority of survey responses (78%). Second, we asked all building administrators and supervisors in the Fair Lawn School District, where the second author has taught for the past eight years, to share the survey link with teachers. Third, we posted a link to the survey on the second author's Facebook page, a common social media platform, with a request to complete and share the survey. Following the initial contact, focal contacts—public-school superintendents as well as Fair Lawn School District building administrators, and supervisors—received one reminder about the survey.

The survey instrument was designed as a web-based, self-administered questionnaire. It included questions gauging respondents' attitudes toward TEACHNJ and its implementation schema, AchieveNJ. We developed these items based on a thorough examination of documents published by the New Jersey Department of Education about TEACHNJ and AchieveNJ (e.g., policy documents and training materials). The survey also contains detailed information on respondents' socio-demographic and professional characteristics. Most of these items were adapted from the General Social Survey and the Trends in International Mathematics and Science Study (TIMSS) Teacher Questionnaire. To evaluate the reliability of the instrument, the second author administered the survey to five teachers and conducted a short follow-up cognitive interview about the survey format and item wordings. The instrument was revised to incorporate their feedback.

Despite the richness of the data, two caveats should be mentioned. First, the study does not involve a random selection of respondents. This design might have implications on (a) the extent to which our sample is representative of the teacher population, and (b) the extent to which we can generalize patterns about the teacher population. Comparing the sample to the general population of New Jersey public-school teachers, we find that our sample includes more women (81% vs. 73%) and more white teachers (93% versus 85%). While our sample includes teachers from all 21 counties in the state, it overrepresents some of the counties and underrepresents other counties. To ameliorate this issue, we calculated and applied sample weights based on gender, race/ethnicity, and county.⁴ Although the generalizability is compromised, the study still sheds light on important and understudied perspectives on new teacher evaluation systems.

Second, the survey instrument did not include information about union affiliation. This is an important control variable because teacher unions had a key role in the public debate over the

⁴ In additional analysis (available upon request), we used the unweighted data. Overall, patterns are very similar.

education reform put forward by Governor Christie. Affiliated members could be more responsive to political framing and messaging by the teacher unions. According to a report by the Thomas B. Fordham Institute (Northern, Scull & Zeehandelaar, 2012) 97.1% of the teachers in New Jersey are unionized. Thus, we are less concerned about the omission of union affiliation as a control variable.⁵

Measures

Support for AchieveNJ. The dependent variable in this study is the level of support for the new teacher evaluation system. As discussed earlier, teacher support or “buy-in” is crucial for successful implementation of any educational reform effort (Akiba, 2013; Fullan, 1993; Spillane, 2004). The item, which was the final question in the survey, included five response categories: (1) strongly oppose, (2) oppose, (3) I’m not sure; (4) favor, and (5) strongly favor.

Teachers’ support for AchieveNJ can be shaped by various factors. The following measures capture the possible effects of the political environment, support for the underlying principles of the reform, self-reported policy knowledge, evaluation of implementation efforts, and perceived policy impact.

Political environment. To test whether teachers’ support was affected by the political environment in which the new teacher evaluation system was proposed and implemented, we used an embedded survey experiment. This design is useful when scholars seek to infer causal relationships from a cross-sectional survey (i.e., the survey experiment clearly distinguishes cause and effect). For example, Pizmony-Levy and his associates used survey experiments to examine whether terminological changes in the debate over same-sex marriage (Pizmony-Levy & Ponce, 2012) and in the debate over the Opt Out movement (Pizmony-Levy & Cosman, 2017) affected public support for these causes. In another study, O’Brien and Pizmony-Levy (2016) used a survey experiment to examine whether a description of a university professor as an engaged scholar affected students’ perception of the professor.

We manipulated the question stem by varying the terms describing the proposer of the new system, AchieveNJ. We used “Governor Christie” to signal a politicized sponsorship and “NJ Legislature” to signal a more neutral sponsorship. Recall that the TEACHNJ bill was proposed by Governor Christie and unanimously passed into law by the NJ Legislature. This manipulation yielded two versions of the item: “Do you oppose or favor AchieveNJ, the new teacher evaluation system, as proposed by [the NJ Legislature / Governor Christie]?” Each version was randomly assigned to respondents. In the analysis to follow, the version of the question is captured by a dummy variable coded 1 for the “Governor Christie” frame and 0 for the “NJ Legislature.” If mentioning the name of Governor Christie reduced teachers’ support for AchieveNJ in the context of a survey, it is possible to infer that political environment may have shaped teachers’ attitudes in the real-world.

Reform principles. Scholars have demonstrated that teachers’ endorsement of reform principles is important for their support of the policy (e.g., Deci, 2009). Thus, the survey asked respondents to indicate the extent to which they agree or disagree with five statements describing the underlying principles and assumptions of the new teacher evaluation system. These statements were derived from the AchieveNJ communication materials (e.g., “evaluation and support system should be developed with significant input from educators” and “tenure should be based on effectiveness,” etc.). These items included four response categories: (1) strongly disagree, (2) disagree, (3) agree, and (4) strongly agree. Drawing on these items, we calculated an index where

⁵ Collins (2014) found no statistically significant difference between the responses of union members and non-union members with regards to attitudes toward SAS Education Value-Added Assessment System.

higher values reflect higher levels of alignment with the policy intention of the new teacher evaluation system (Cronbach's $\alpha = .79$).

Self-reported policy knowledge. Successful implementation of policy requires that stakeholders “on the ground” demonstrate knowledge of different aspects of the reform (e.g., Healey & DeStefano, 1997). To assess policy knowledge, the survey asked respondents to indicate the extent to which they are knowledgeable about nine key concepts related to the reform. These concepts included the law itself (TEACHNJ), the implementation platform (AchieveNJ), and state-mandated components of the reform (e.g., student growth objectives, teacher observation instrument, etc.). These items included four response categories: (1) not at all knowledgeable, (2) not very knowledgeable, (3) somewhat knowledgeable, and (4) very knowledgeable. Drawing on these items, we calculated an index where higher values reflect higher levels of perceived knowledge about key concepts in the new teacher evaluation system (Cronbach's $\alpha = .83$).

Evaluation of implementation. Another factor that might play a role in teachers' support is their interaction with and evaluation of the implementation process. To measure teachers' evaluation of the implementation efforts of the new teacher evaluation system, the survey asked respondents to indicate the extent to which they agree or disagree with five statements describing different facets of policy implementation: communication, pacing, resources, support, and monitoring. These items included four response categories: (1) strongly disagree, (2) disagree, (3) agree, and (4) strongly agree. Drawing on these items, we calculated an index where higher values reflect a more positive view of the implementation process (Cronbach's $\alpha = .89$).

Perceived policy impact. Finally, we determined whether teachers' belief in the efficacy of the new teacher evaluation system is associated with their support for the system. The survey asked respondents to indicate the extent to which they agree or disagree with the following two statements: “The new evaluation system will improve teaching” and “The new evaluation system will improve student learning.” The item included four response categories: (1) strongly disagree, (2) disagree, (3) agree, and (4) strongly agree. These survey items are strongly correlated ($r = .88, p < .001$).

Controls

In all multivariate analysis models, we controlled for a series of sociodemographic and professional characteristics, including gender, race/ethnicity, age, urban/rural status of school community, tenure, and tested subject teachers.⁶ In the 2013-14 academic year, when data for this study were collected, the tested subjects included mathematics and English/language arts in grades four through eight. Table 1 presents definitions, metrics, and descriptive statistics for all variables included in the analysis.

⁶ We used the variable age instead of years of experience because the former had fewer missing cases. The two variables are highly correlated ($r = .75, p < .001$). Supplementary analysis with the variable years of experience produced similar results.

Table 1

Definitions, Metrics, and Descriptive Statistics of Variables in Study (n=444)

Variable	Definition/Metric	Mean	SD
<i>Dependent Variable</i>			
Support for AchieveNJ	Do you oppose or favor AchieveNJ, the new teacher evaluation system, as proposed by [the NJ Legislature/Governor Christie]? The item included five response categories: (1) strongly oppose, (2) oppose, (3) I'm not sure, (4) favor, and (5) strongly favor.	2.18	1.01
<i>Experimental Conditions</i>			
<i>Proposer Frame</i>			
Governor Christie	0 = other, 1 = Governor Christie	.51	-
NJ Legislature	0 = other, 1 = NJ Legislature	.49	-
<i>Independent Variables</i>			
Policy alignment	Respondent's endorsement of underlining principles of the new teacher evaluation system. The index is based on five items. Higher values reflect higher levels of alignment with the policy principles.	3.38	.55
Perceived knowledge	Respondent's self-reported knowledge of key concepts in the new teacher evaluation system. The index is based on nine items. Higher values reflect higher level of knowledge.	2.55	.63
Perceived impact	Respondent's expectations of the impact of new teacher evaluation system on teaching and learning. Average of two items. Higher values reflect greater expected impact.	1.97	.76
Evaluation of implementation	Respondent's evaluation of implementation efforts. The index is based on five items. Higher values reflect more positive evaluation of implementation efforts.	1.92	.62
Women	Gender of respondent: 0 = men, 1 = women	.79	-
Non-white	Race of respondent: 0 = white, 1 = non-white	.07	-
Age	Respondent's age in years	43.65	10.50
<i>School community</i>			
Urban	0 = no, 1 = yes	.14	-
Suburban	0 = no, 1 = yes	.74	-
Rural	0 = no, 1 = yes	.12	-
Tenure status	0 = no, 1 = yes	.84	-
Assessment target	0 = no, 1 = yes	.30	-

Note: Data is weighed based on gender, race/ethnicity, and county.

Open-ended Responses

The final section of the survey instrument included an open-ended question that prompted respondents to share their thoughts on issues raised in the survey. This qualitative dimension allowed us to preserve the teachers' voices and emphases, and provided a window into how teachers make sense of the new evaluation system. Furthermore, we used the responses to open-ended questions to help corroborate statistical findings. Approximately two-fifths (39.3%) of the sample responded to the open-ended item. Analyses of cases missing open-ended responses revealed several differences in the likelihood of responding. Older and tenured teachers who work in urban contexts were more likely than others to respond to the open-ended item. We found no differences based on sex, race/ethnicity, teaching a tested subject, and experimental condition.

Analytical Technique

After removing missing data, the final sample included 444 teachers. We used Ordinary Least Squares (OLS) regression modeling to examine teachers' support for AchieveNJ. Model 1 includes the randomized condition from the survey experiment. Model 2 adds sociodemographic characteristics (gender, race/ethnicity, and age) and professional characteristics (school community, teaching experience, tenure status, and tested subject teachers). The next four models introduce different measures of engagement with the reform: policy alignment (model 3), self-reported knowledge (model 4), perceived impact (model 5), and evaluation of implementation (model 6). Finally, model 7 introduces simultaneously all measures of engagement with the reform.

Because the sample includes teachers nested in counties, the usual regression assumption of the stochastic independence of error terms underlying tests of statistical significance is violated. For example, teachers' experience with the actual implementation of AchieveNJ could be influenced by the local context (e.g., resources, reception, etc.). Teachers working in the same county or district will have more similar experience than teachers working in other counties. Therefore, in all regression analyses, we corrected for this non-independence of observations using the cluster procedure available in Stata to compute robust standard errors. We present results with standard errors clustered for county (rather than school district) because of missing data and sample size of each cluster. Analysis with standard errors clustered for school district yielded the same results and are available upon request.

We entered all the open-ended responses into Atlas.ti, a computer program that facilitates qualitative analysis. Coding began with the creation of open categories that emerged through multiple readings of the data rather than from an existing theoretical perspective (Creswell, 2007). Attending to the structure of the comments, for example, we noticed that many respondents simultaneously supported and criticized AchieveNJ. This became a category. Additional reading of the responses revealed doubts about the validity of the new evaluation system. Data fitting this broad description became another category.

Results

Descriptive Statistics

This section presents frequency distribution and descriptive statistics for the factors that might shape teachers' support for AchieveNJ. In our analysis, we began by exploring teachers' support for the underlying principles of TEACHNJ and AchieveNJ. That is, we examined the alignment of teachers' views with the principles behind the policy reform. As seen in figure 1 below, a large majority of teachers (more than four-fifths) indicated that they agree with the five statements describing the new teacher evaluation system. For example, almost all teachers (96.5%) endorsed the

notion that “timely feedback and high-quality professional development are essential to help educators improve.” Similarly, almost all teachers (95%) endorsed the notion that “evaluation and support systems should be developed with significant input from educators.” Even the link between tenure and effectiveness was validated by more than four-fifths of the sample (82.5%), although at a notably lower percentage than other items.

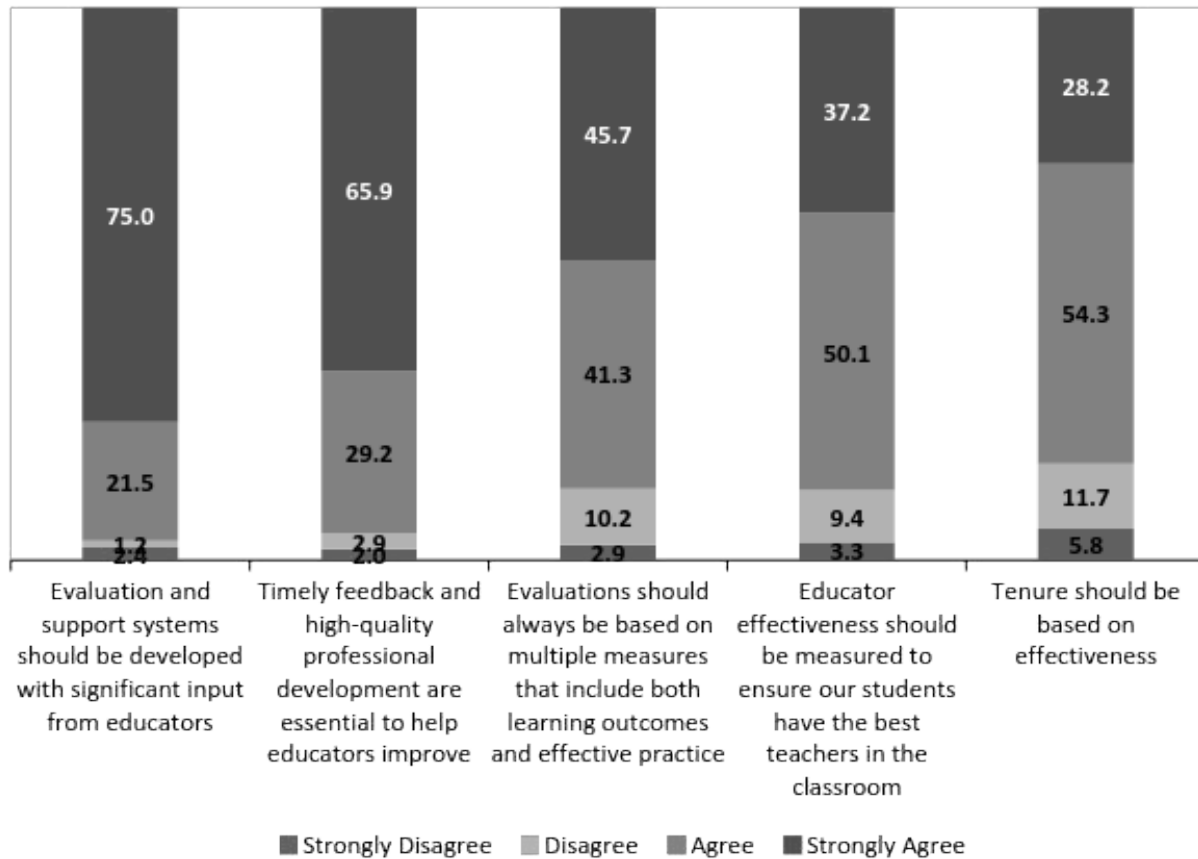


Figure 1. Attitudes toward underlining rationales of the new teacher evaluation system

The majority of teachers (94.9%) participated in trainings that were offered as part of the implementation of TEACHNJ and AchieveNJ. When we examined their self-reported knowledge, however, we found much variation across key concepts. That is, the teachers in this study were more knowledgeable of the most immediate and concrete aspects of the new evaluation system and less knowledgeable of the law TEACHNJ and its implementation platform AchieveNJ. For example, using a rating scale, a majority of teachers said they were knowledgeable (combination of “very knowledgeable” and “somewhat knowledgeable”) about “student growth objectives” (84.6%), the “teacher evaluation system” (82.5%), and the “teacher observation instrument” (71.9%). In contrast, less than half of teachers said they were knowledgeable about “AchieveNJ” (46.8%), “TEACHNJ” (43.1%), the “school improvement panel” (34.4%), and the “District Evaluation Advisory Committee” (29.1%).

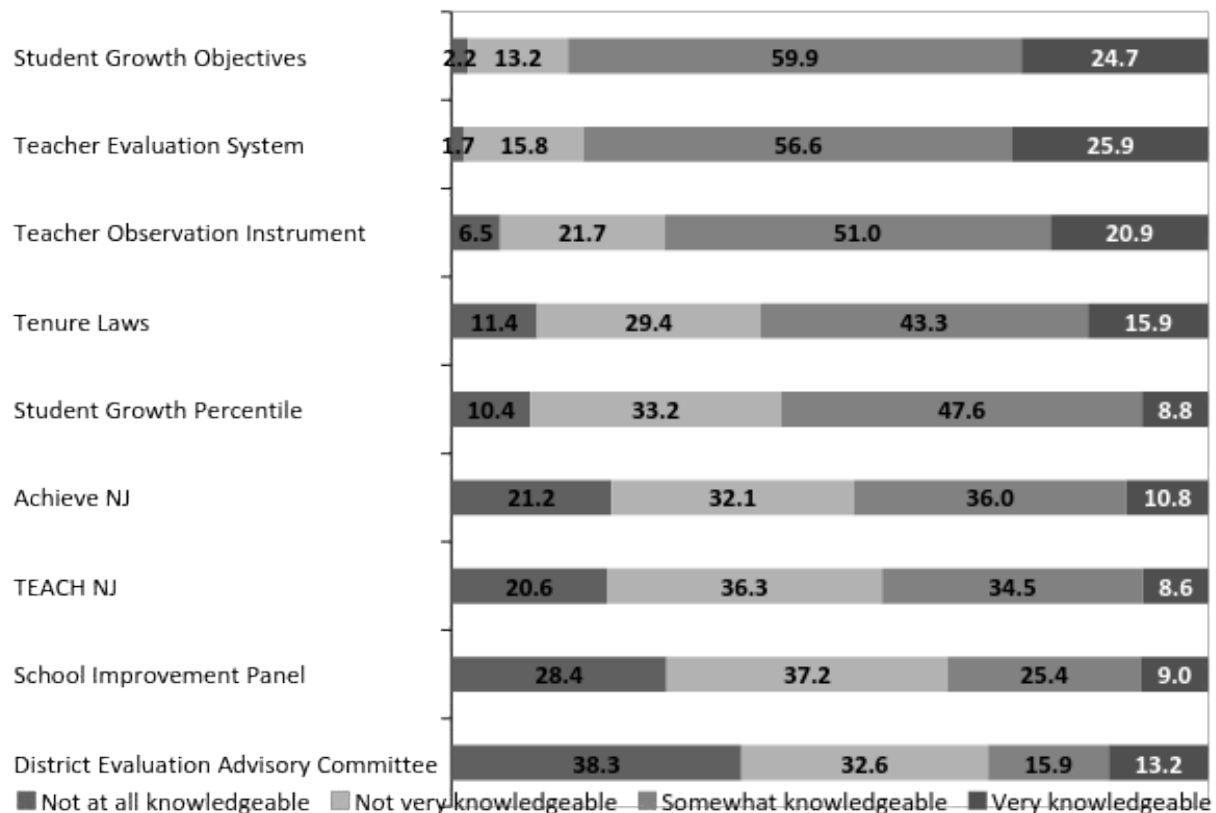


Figure 2. Teachers' knowledge of the new teacher evaluation system

Next, we examined teachers' evaluation of implementation efforts of the new evaluation system. As seen in Figure 3, teachers are critical of these efforts. A clear majority (more than two-thirds) of the sample disagreed to some extent with statements describing the implementation efforts. Slightly more than four-fifths (81.8%) disagreed with the statement "the rollout of the new teacher evaluation system, TEACHNJ, was done at a comfortable pace that allowed for effective implementation," and slightly more than three-quarters (77.9%) disagreed with the statement, "The new laws and changes in teacher evaluation are effectively communicated by NJ Department of Education."

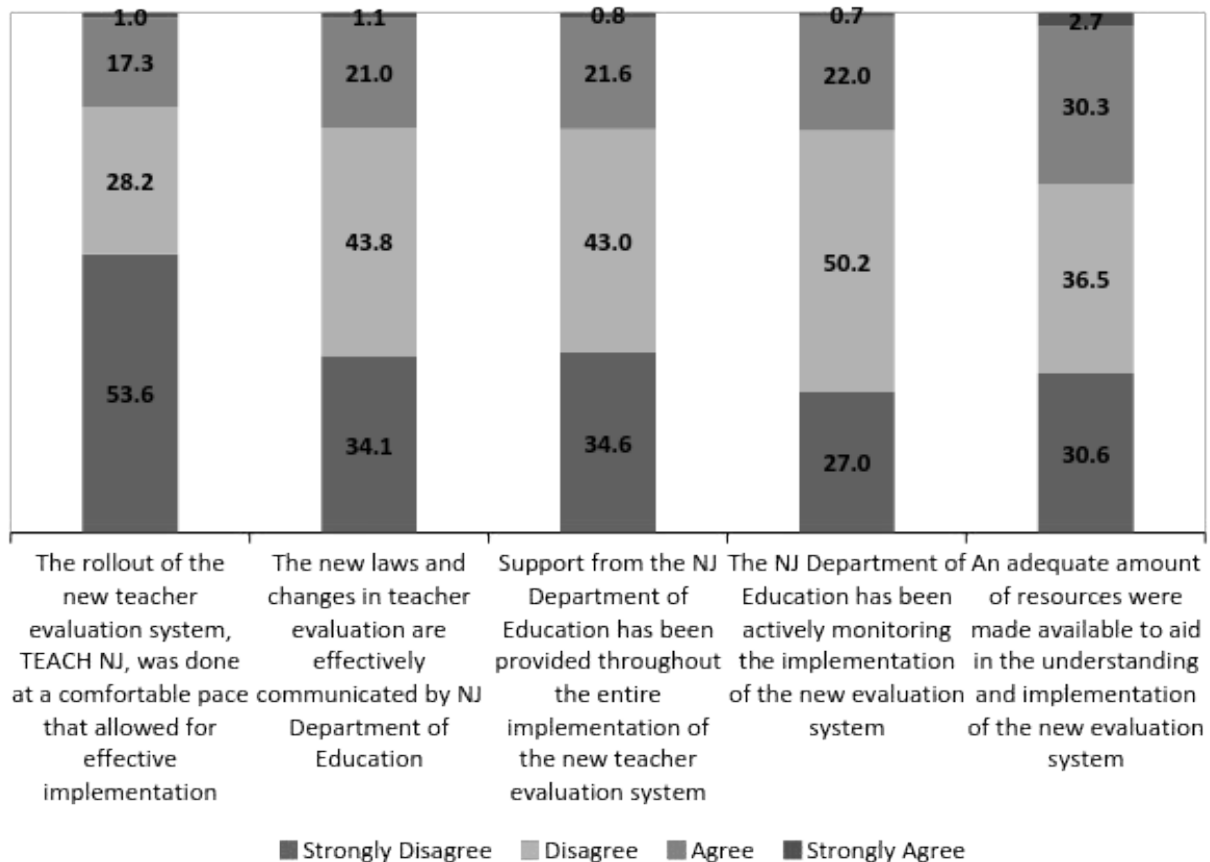


Figure 3. Attitudes toward the implementation efforts of the new teacher evaluation system

Teachers' experience with implementation efforts could be shaped by local control at the individual county, school district or school. Thus, we conducted a one-way ANOVA to compare the average of evaluation of implementation scale among respondents from different counties and school districts. The analysis for counties did not produce a significant result. The analysis for districts, however, produced a statically significant results ($F_{(75, 382)}=1.50, p<.01$).

Not only were the teachers in this study critical of the implementation efforts, but the majority of respondents reported pessimistic views of the impact of the new evaluation system on teaching and learning. About three-quarters (74.0%) of respondents said it will not improve teaching (30.1% strongly disagreed and 43.9% disagreed with a positive statement). Slightly more than four-fifths (81.1%) of respondents said the new system will not improve student learning (30.1% strongly disagreed and 51.0% disagreed with a positive statement).

Finally, we examined the distribution of the dependent variable: Support for AchieveNJ. As expected, two-thirds of respondents (66.5%) indicated they oppose AchieveNJ (29.5% strongly opposed and 37% opposed). Only one-in-ten respondents (9.5%) said they favor AchieveNJ. The remainder, one-fourth of the sample (24%) said that they are not sure whether they favor or oppose AchieveNJ.

Multivariate Analysis

Earlier we noted that teachers' support for educational reforms is often shaped by the political environment in which the reform was introduced. Furthermore, we found that the process

in which TEACHNJ and AchieveNJ were developed was highly contentious. Consequently, we now turn to discuss the effects of different sponsor frames on support for AchieveNJ, as illustrated in Table 2.

Table 2

Unstandardized OLS Regression Coefficients of Support for AchieveNJ (n=444)

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
Sponsor:	-.356***	-.376***	-.365***	-.380***	-.341***	-.407***	-.383***
Gov. Christie	(.078)	(.067)	(.075)	(.066)	(.058)	(.051)	(.052)
Women		.096 (.115)	.061 (.107)	.090 (.123)	.024 (.097)	-.041 (.135)	-.059 (.113)
Non-white		.051 (.144)	.089 (.130)	.028 (.140)	-.053 (.171)	.077 (.125)	.036 (.140)
Age		-.007 (.005)	-.007 (.004)	-.008 (.004)	-.005 (.004)	-.006 (.003)	-.005 (.003)
School community							
Rural		.273 (.198)	.246 (.182)	.255 (.205)	.148 (.193)	.148 (.155)	.106 (.156)
Suburban		.256 (.168)	.238 (.155)	.260 (.173)	.154 (.157)	.145 (.122)	.114 (.124)
Tenured		-.223 (.182)	-.202 (.183)	-.206 (.171)	-.053 (.139)	-.015 (.122)	.031 (.115)
Assessment target		.081 (.099)	.075 (.095)	.033 (.073)	.159* (.064)	.180* (.078)	.192** (.054)
Policy alignment			.368* (.157)				.106 (.127)
Policy knowledge				.352** (.100)			.028 (.083)
Evaluation of implementation					.676*** (.077)		.297** (.088)
Perceived policy impact						.816*** (.045)	.673*** (.049)
Intercept	2.351***	2.535***	1.312*	1.677***	1.101***	.913**	.145
adj. R ²	.029	.043	.071	.082	.215	.366	.394
AIC	1260.259	1260.765	1248.364	1242.987	1173.432	1078.926	1061.669
BIC	1268.451	1297.627	1289.322	1283.945	1214.390	1119.884	1114.915

Notes: Reference category for sponsor is New Jersey Legislature and for school community is urban. Standard errors in parentheses are clustered for county.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Model 1 presents the bivariate correlation between the sponsor framing of the new evaluation system and the respondent's support for the system. We found that when the system was presented as an initiative of Governor Chris Christie (rather than an initiative of the New Jersey Legislature), the new evaluation system received less support from teachers (model 1; $b = -.356$, $p < .001$). This pattern persisted after we controlled for sociodemographic and professional characteristics (model 2). This should not be a surprise given the fact that the sponsor framing was randomly assigned to respondents. The significant effect of sponsor framing holds across all other models, regardless of the controls added to the model.

Model 2 introduces the control variables. Sociodemographic characteristics and most professional characteristics had no significant effect on support for the new evaluation system. The only exception was the coefficient for tested subject teachers (those teaching mathematics and English/language arts in grades four through eight). In the first two models (models 2 and 3), the coefficient's effect size is very small and not significant. However, in the final three models, the coefficient increases and becomes statistically significant (model 6; $b = .191$, $p < .01$). In other words, two variables – teachers' evaluation of implementation efforts and teachers' expectations of policy impact – suppress the relationship effect of tested subject teachers on the outcome variable. Compared to other teachers, tested subject teachers showed more support for AchieveNJ, but they were also more critical about the implementation efforts and the impact of the new evaluation system.⁷ Thus, when including these engagement measures in the model, the effect size of tested subject teachers increases and becomes significant.

Finally, we examined the effect of four engagement measures on support for the new teacher evaluation system. Not surprisingly, teachers' endorsement of the underlying principles of the new teacher evaluation system had no significant effect on support for the new system (model 3). In contrast, teachers' self-reported policy knowledge had positive and significant effects on support for the new evaluation system (model 4). Teachers who perceived themselves as knowledgeable of various components of the reform expressed greater support for the new system. This coefficient, however, reduced in size and loses significance when we accounted for other engagement measures (model 7). This might be a result of the strong correlation (or overlap) between self-reported policy knowledge and evaluation of implementation ($r = .366$, $p < .001$).

Both teachers' evaluation of implementation efforts and teachers' expectations of policy impact had positive and significant effects on support for the new evaluation system (models 5 and 6). Teachers who expressed a more positive view of the implementation efforts also showed more support for the new evaluation system. Teachers who perceived the new policy as efficacious also showed more support for the new evaluation system. This pattern was consistent, though the coefficients' effect size shrunk, after we accounted for other engagement measures (model 7). This might be a result of the strong correlation (or overlap) between the two variables ($r = .486$, $p < .001$). These variables accounted for much of the variation in the outcome variable, as reflected in the adjusted R-squared.

Open-ended Responses

As mentioned above, we used open-ended response data to further examine teachers' views toward the new teacher evaluation policy (i.e., TEACHNJ) and its implementation platform (i.e., AchieveNJ). We identified five key themes in the data. All themes cut across the two experimental

⁷ In order to investigate this pattern, we estimated a regression model for the evaluation of implementation with sociodemographic and professional characteristics as predictors. The effect of tested subject teachers was negative, but not significant ($b = -.114$, $p = .167$). We estimated a similar model for policy impact. The effect of tested subject teachers was negative and marginally significant ($b = -.122$, $p = 0.06$).

conditions.⁸ First, similar to the pattern we discussed in the quantitative analysis, many teachers *simultaneously* expressed both support for change *and* critique of the current direction of policy. For example, a Latina English teacher with 11 years of teaching experience wrote, "Although I do believe that there needs to be a change in education, I don't feel that implementing a new evaluation system is the answer. We should be more focused on engaging students and families, once this is done, we then can focus on the teacher's accountability." Echoing the concern that there was a problem that needed to be addressed, a White male science teacher with 28 years of teaching experience pointed to problems with both old tenure laws, yet argued that the new rules add a new set of concerns:

"Although there was abuse of the old tenure laws, I have seen with the new ones administrators using the new rules as an intimidation tool. It seems to me that you are at the whim of evaluators as to your job status. If they take a dislike to you for any reason the new system is sufficiently complicated that it would be easy to mask the real reason for bringing tenure charges and dismissing a teacher."

A second theme that emerged highlighted concerns about the negative consequences of the new teacher evaluation policy. One-in-five teachers (20.9%) mentioned these sentiments in their responses. This theme maps well on responses to survey items about the perceived impact of the policy.⁹ For example, a relatively new White male science teacher, with five years of teaching experience, stated that he is now discouraged about how he will continue to grow as a teacher and is considering a career change. He suggested that the new policy negatively affects teacher's morale:

"The new teaching evaluation system, with random unannounced visits that seem to count, while providing little feedback has given a sense of "Big Brother" to the classroom. It has taken the fun out of trying new things and sitting back and reflecting to see if it works [...] the haphazard implementation of all of these new systems have turned what was a profession that I loved when I first started 5 years [ago] into something that is turning into a job. [...] I have recently stopped working on my master's and am currently considering a change in profession due to the direction I have seen our school system take in the past few years."

Other teachers addressed common narratives of teaching to the test and wasting resources (e.g., time). A Latina teacher who has been teaching physical and health education for 25 years wrote, "Too much time is spent teaching to the test. This is just a small measurement of student learning." This concern was echoed by a White female science teacher who believes that the new policy does not allow her to use the skills that she has gone to school to learn and has been honing over the last 10 years. She commented, "[the test] takes too much time away from planning for the students [...]" Time is of the essence. The more time teachers have to plan the more exciting learning classes can be."

The third theme concerned the DOE's poor implementation efforts. One-in-10 teachers (10.4 %) addressed this issue in their responses. Many teachers commented about the quick pace of

⁸ We conducted a chi-square analysis to determine whether having the theme in the open-ended responses (yes coded 1 and no coded 0) is associated with a specific experimental condition. The analysis produced no significant pattern.

⁹ We conducted a chi-square analysis to determine whether having the theme in the open-ended responses (yes coded 1 and no coded 0) is associated with response to survey items about the expected impact of the policy on teaching and learning. The analysis produced a significant X^2 values ($X^2=19.49, p<.001$; $X^2=15.50, p<.001$), indicating dependency between the open-ended responses and the survey items.

the implementation, which did not allow teachers to fully engage. For example, a multiracial female midcareer computer science teacher noted:

“I do believe that overall this new system will be beneficial. I am concerned that it was rolled out a bit quickly, not giving educators and administrators time to understand it fully. Because of the speedy implementation, I do believe there will be a learning curve. A slower roll out of this process would have allowed everyone to find and correct the issues without having to hold to a flawed process.”

In addition to writing about the quick pace of the implementation, teachers also commented about poor training and preparation. For example, a White female teacher with nine years of experience reported that she has seen a number of different policy rollouts and has worked in a district that supports its teachers, in her estimation, well. However, the implementation of the new policies caught even the district off guard. As she reported, “Just to clarify, in my district, we were very prepared for the Charlotte Danielson model and started our training about 4 years ago. However, the training with the SGP's and SGO's was poorly done, not because our district did not provide resources, but because the state was slow in providing the district with needed answers.”

A fourth theme that emerged concerned the politics of introducing and advocating for the new evaluation system. While our quantitative analysis indicates that teachers' support is affected by the political environment, teachers' open-responses shed further light on possible reasons for this association. Some teachers indicated that they view the new policy as part of a political campaign against public education. For example, a White female foreign language teacher with over four decades of teaching experience, indicated that she is suspect of anything that Governor Christie proposes given her assessment of his motives. She wrote, “I would never support anything proposed by this governor who has decimated and vilified teachers to the extent that he has.” Similarly, a White male science teacher with 25 years of experience opined that this policy was meant to inflict damage. “This new system is the biggest waste of time and resources I have witnessed in the last 25 years,” the teacher wrote. “I guess [Governor] Christie is going to try to kill the public schools with this initiative.”

Other teachers criticized the fact that the new policy is not informed by relevant knowledge and professional experience. For example, a White female English teacher wrote, “I think the people in charge of the new evaluation system have never been in a classroom or have little to no background in education. I do not enjoy being seen as a number on a spread sheet.” Another White female elementary school teacher with 14 years of experience pointed to the de-professionalization of education policy and how teachers are going to be held to standards for which they are not trained. “Governor Christie just passed legislation on kids with Dyslexia [...] yet teachers are not trained in teaching these kids (I am not). Yet we will be held accountable for these students' growth. This is not fair.” She noted that basic educational theory was seemingly missing from the creation and consideration of the policy, writing, “What happened to Piaget's theory of childhood development? It seems to have gone out the window when they expect kids to all be at a certain reading, math level by a certain date. Everything we learned in college about child psychology seems to have vanished.”

Teachers also linked the new evaluation system to the broader phenomenon in which business-sector ideas are entering schools. A White male math teacher with over a quarter-century of teaching experience argued that the policies were created to support testing corporations rather than to improve education:

“The use of high stakes standardized test score [sic] in evaluating performance is what makes the new system an absolute farce. It ignores the overwhelming amount

of research that indicates that these measures are suspect at best. As far as I am concerned, this new system has nothing to do with children or their teachers and everything to do with the corporate reform movement and the testing-industrial complex and their attempt to privatize public education.”

Another White female elementary school teacher added, “They are trying to make teaching a ‘business,’ and it is not a business and will never be. Silly me, I thought we were teaching kids, not computers and guess what? Kids make mistakes and should make mistakes, because that’s how they learn from them. Kids are not motivated by standardized tests and here’s another clue, NEITHER ARE TEACHERS!!!!” (Emphasis added by respondent).

A fifth theme included questions and doubts about the validity of the new evaluation system. In contrast to the previous themes, this theme was not reflected in the survey instrument. Thus, this theme extends our insights regarding how teachers view the new evaluation policy. A female White science teacher with over a decade of experience argued that the important role of out-of-school factors in shaping students’ achievement was ignored by policymakers. “While I believe it is important for educators to be evaluated, and not just a one-time classroom visit, I do not believe that a teacher should be measured by their students’ performance,” she said. “Some students have isolated situations of which [sic] no teacher will make a difference.” Similarly, a White female special education teacher with 25 years of experience added that the learning environments for certain students were not taken into account:

“Most of my students are far below their grade level and need to be taught in a specialized manner in order for them to achieve. The evaluation system does not allow consideration to those of us who teach said children. We will never cover some of the items described in the distinguished areas. I feel the evaluation should be written for groups of educators [...] To group us all the same is not fair.”

In sum, the open-ended responses have suggested that New Jersey teachers recognize the need for change and endorse the principles of the new teacher evaluation system. However, the teachers in this study expressed concerns about different aspects of the evaluation system, including possible negative consequences, poor implementation, the politics that accompanied the policy process, and the validity of the system. These concerns, in turn, translate into clear opposition to the evaluation system.

Discussion

Through an analysis of a survey of New Jersey teachers, we examined teachers’ views of a high-stakes teacher accountability policy (TEACHNJ) and its implementation platform (AchieveNJ). Our study points to an interesting paradox. On the one hand, a majority of teachers are aligned with the principles of the policy. For example, they agreed that effectiveness should be measured to ensure students have the best teachers. Some even endorsed linking tenure to effectiveness. On the other hand, teachers were critical of the policy and the implementation efforts. The vast majority (75%) of teachers indicated that they believe that the new policy will not improve their teaching or students’ learning. And, only 10% of teachers said they support the new teacher evaluation system. The rest opposed or were undecided. The paradox is illustrated in the closed-ended items and the open-ended responses. Many of the open-ended responses had a similar structure: “I believe in teacher accountability, but...” In fact, Woody et al. (2004) echoed these patterns, based on research in four districts in California. Overall, participants in New Jersey and California saw the need for the

reform, but they did not necessarily agree with the efforts being carried out to implement the reform. In other words, teachers tend to agree that the idea of “teacher accountability” is now taken for granted, yet the debate about how teachers should be held accountable remains unsettled.

Our empirical findings also highlight the role of the political environment on teachers. As we describe in the study setting, the process in which TEACHNJ and AchieveNJ were introduced was highly contentious and political. Local media coverage reported this tension, regularly citing disparaging remarks by Governor Christie toward the local teachers’ union. Rather than asking teachers about whether the political environment affected them, we used an embedded experiment. This innovative approach enabled us to estimate the causal effect of a treatment in the context of a large survey. Respondents who read about AchieveNJ as an initiative of Governor Christie were significantly less supportive of the new teacher evaluation system than respondents who read about AchieveNJ as an initiative of the New Jersey Legislature. Open-ended responses shed light on the how teachers think about the politics of education. Some of the teachers accused Governor Christie for leading a campaign against public education. Others commented on the de-professionalization of education by individuals who craft policy solutions with limited expert knowledge or experience. Teachers also pointed to the growing role of corporations in the education and schools (e.g., testing agencies).

Finally, our empirical findings suggest that two additional factors are associated with teachers’ support for AchieveNJ. First, similar to previous studies (Desimone, 2002), we found that teachers holding positive views of the implementation efforts also reported higher levels of support for the policy. Both closed-ended items and open-ended responses point to the importance of timing and pace. Many teachers reported that TEACHNJ was implemented very quickly with teachers feeling rather unprepared. Second, as reported by Datnow & Castellano (2000), teachers who believed that the policy will have an impact on teaching and learning demonstrated higher levels of support. Open-ended responses suggest that teachers were concerned about the negative impact of the new policy on morale, motivation, and engagement.

Our study has three limitations that should be considered. First, the study is based on a small non-random sample of teachers, which may raise a question about generalizing about the entire population of teachers. As we discussed earlier, this sampling design is a result of limited resources and support. In an effort to address the sampling design, we calculated and applied sample weights (based on gender, race/ethnicity, and county) to adjust for demographic differences between the sample and the population. The relatively small sample ($n=444$) prevents us from applying more sophisticated analysis to the data (e.g., multi-level analysis of the nested nature of education systems). Second, the study is based on a sample of teachers in one state, New Jersey. It is possible that teachers in other contexts (states or countries) will have different views and attitudes towards high-stakes teacher accountability policies. Third, data for this study were collected during the first academic year after the implementation of TEACHNJ. It is possible that teachers’ views may have changed over time once they had more experience with the new policy. Evaluation of TEACHNJ (Callahan & Sadeghi, 2014), however, has shown that “while teachers indicated that they were observed more often, they also noted the value of the observation was diminished” (p. 56).

Given these limitations and in light of the findings of this study, there are at least two possible directions for future research. First, additional systematic research on teachers’ attitudes toward high-stakes teacher accountability policies is required. Quantitative studies based on large and random samples will address the limitations and expand our findings. These studies could introduce additional factors that may shape teachers’ views toward reforms, such as party identification, political ideology, leadership roles, sense of empowerment, and motivations to join the teaching profession. Furthermore, these studies could examine the role of organizational

affiliation (e.g., unions) and social networks (e.g., personal and professional networks). Qualitative studies based on in-depth interviews and observations will help scholars and stakeholders to better understand *how* teachers interpret and “make sense” of high-stakes teacher accountability policies. Additionally, these studies will provide further insights about *why* teachers support or oppose these policies.

Second, additional research would explore teachers' views of other applications of NPM models, such as school choice, privatization, corporatization, technology, alternative requirements and training programs (e.g., Teacher For America), and standardized testing (e.g., reliance on ILSA and national assessments to inform policy and practice). Comparing teachers' views across these different policies and practices will provide a more refined understanding of their engagement with these policies. Do teachers differentiate between these policies? If so, what factors shape support for one policy versus another policy?

In conclusion, teachers are important actors in the education system. This study demonstrated that teachers endorse the principles of new teacher evaluation policy, but they oppose the actual policy and its implementation platform for a variety of reasons. Teachers in this study view high-stakes teacher accountability policies (and perhaps broader contemporary education reform movements) as undermining teachers by challenging the building blocks of their professional status (e.g., expert knowledge, specialized training, and autonomy). In an era of multiple education reform movements, scholars can enrich the policy process by engaging teachers' voices. Careful analysis of teachers' views not only could shed light on policy implementation, but it could also yield innovative ideas for policies aimed at improving education quality.

References

- Addi-Raccah, A. (2012). Supporting or Opposing Privatization in Education: Teachers' Attitudes and Positions of Power in Israeli Schools. *Leadership and Policy in Schools*, 11(4), 391-417. <https://doi.org/10.1080/15700763.2012.654887>
- Akiba, M. (Ed.). (2013). *Teacher reforms around the world: Implementations and outcomes* (Vol. 19). Emerald Group Publishing. [https://doi.org/10.1108/S1479-3679\(2013\)19](https://doi.org/10.1108/S1479-3679(2013)19)
- American Educational Research Association. (2015). AERA statement on use of value-added models (VAM) for the evaluation of educators and educator preparation programs. *Educational Researcher*, 44(8), 448-452. <https://doi.org/10.3102/0013189X15618385>
- Baker, B. D., Oluwole, J., & Green, P. C. (2013). The legal consequences of mandating high stakes decisions based on low quality information: Teacher evaluation in the race-to-the-top era. *Education Evaluation and Policy Analysis Archives*, 21(5), 1-71. <https://doi.org/10.14507/epaa.v21n5.2013>
- Ballou, D., & Podgursky, M. (1993). Teachers' attitudes toward merit pay: Examining conventional wisdom. *Industrial & Labor Relations Review*, 47(1), 50-61. <https://doi.org/10.1177/001979399304700104>
- Blackburn, B. (April 6, 2011). New Jersey Governor Chris Christie Calls His State's Teachers Union 'Political Thugs'. Retrieved from: <http://abcnews.go.com/Politics/jersey-governor-chris-christie-calls-teachers-union-political/story?id=13310446>
- Bogart, C. D. (2013). Teacher evaluation and classroom practice: Teacher perceptions in northeast Tennessee (Doctoral dissertation).
- Braun, H. I. (2005). Using Student Progress to Evaluate Teachers: A Primer on Value-Added Models. Policy Information Perspective. Princeton, NJ: Educational Testing Service.

- Bridich, S. (2015). The invisible schism: Teachers' and administrators' differing perceptions of education reforms. *Education Policy Analysis Archives*, 24(87).
<http://dx.doi.org/10.14507/epaa.24.2192>
- Callahan, K., & Sadeghi, L. (2014). TEACHNJ: An evaluation of two years of implementation. *US-China Education Review A*, 4(10), 728–736.
- Collins, C. (2014). Houston, we have a problem: Teachers find no value in the SAS education value-added assessment system (EVAAS®). *Education Policy Analysis Archives*, 22.
<https://doi.org/10.14507/epaa.v22.1594>
- Creswell, J. W. (2007). *Qualitative inquiry and research design: Choosing among five approaches*. Thousand Oaks, CA: Sage.
- Datnow, A., & Castellano, M. (2000). Teachers' responses to Success for All: How beliefs, experiences, and adaptations shape implementation. *American Educational Research Journal*, 37(3), 775–799. <https://doi.org/10.3102/00028312037003775>
- Deci, E. L. (2009). Large-scale school reform as viewed from the self-determination theory perspective. *Theory and Research in Education*, 7(2), 244–252.
<https://doi.org/10.1177/1477878509104329>
- Desimone, L. (2002). How can comprehensive school reform models be successfully implemented? *Review of Educational Research*, 72(3), 433–479. <https://doi.org/10.3102/00346543072003433>
- Druckman, J. N. (2001). The Implications of Framing Effects for Citizen Competence. *Political Behavior* 23, 225–56. <https://doi.org/10.1023/A:1015006907312>
- Druckman, J. N. (2004). Political Preference Formation: Competition, Deliberation, and the (Ir)relevance of Framing Effects. *American Political Science Review*, 98, 671–86.
<https://doi.org/10.1017/S0003055404041413>
- Education Next. (2016). Results from the 2016 Education Next Poll. Retrieved from website:
<http://educationnext.org/files/2016ednextpoll.pdf>
- Elmore, R. F., Abelman, C. H., & Fuhrman, S. H. (1996). The new accountability in state education reform: From process to performance. *Holding schools accountable: Performance-based reform in education*, 65–98.
- Figazzolo, L. (2009). Impact of PISA 2006 on the education policy debate. Education International.
- Fleisher, L. (2010, September 28). Gov. Christie pushes for teacher merit pay based on students' performance. Retrieved from:
http://www.nj.com/news/index.ssf/2010/09/gov_christie_advocates_merit_p.html
- Fullan, M. (1993). *Change forces: Probing the depths of educational reform* (Vol. 10). Psychology Press.
- Fullan, M. (2001). *The new meaning of educational change*. New York, NY: Teachers College Press.
- Gewirtz, S., & Ball, S. (2000). From 'Welfarism' to 'New Managerialism': shifting discourses of school headship in the education marketplace. *Discourse*, 21(3), 253–268.
- Goldhaber, D., DeArmond, M., & DeBurgomaster, S. (2011). Teacher attitudes about compensation reform: Implications for reform implementation. *Industrial & Labor Relations Review*, 64(3), 441–463. <https://doi.org/10.1177/001979391106400302>
- Healey, F. H., & DeStefano, J. (1997). *Education reform support: A framework for scaling up school reform*. Policy Paper Series. Research Triangle Park, NC: Research Triangle Institute.
- Hess, F. M., Maranto, R., & Milliman, S. (2000). Resistance in the trenches: What shapes teachers' attitudes toward school choice? *Educational Policy*, 14(2), 195–213.
<https://doi.org/10.1177/0895904800014002001>
- Holloway-Libell, J., & Collins, C. (2014). VAM-based teacher evaluation policies: Ideological foundations, policy mechanisms, and implications. *InterActions: UCLA Journal of Education and Information Studies*, 10(1).

- Kim, J., & Youngs, P. (2016). Promoting instructional improvement or resistance? A comparative study of teachers' perceptions of teacher evaluation policy in Korea and the USA. *Compare: A Journal of Comparative and International Education*, 46(5), 723-744.
<https://doi.org/10.1080/03057925.2015.1057478>
- Koopmans, R. (1999, March). Political. Opportunity. Structure. Some splitting to balance the lumping. *Sociological Forum*, 14(1), 93-105. <https://doi.org/10.1023/A:1021644929537>
- Levy, J. S. 2002. Daniel Kahneman: Judgment, Decision, and Rationality. *PS: Political Science and Politics*, 35, 271-3. <https://doi.org/10.1023/A:1021644929537>
- Levy, J. S. 2003. Applications of Prospect Theory to Political Science. *Synthese*, 135, 215-41.
<https://doi.org/10.1023/A:1023413007698>
- Lipsky, M. (2010). *Street-level bureaucracy: Dilemmas of the individual in public services*. New York: Russell Sage Foundation.
- Mascarenhas, R. (2010, November 18). Report finds N.J. 12th graders are above-average in math, average in reading. Retrieved from:
http://www.nj.com/news/index.ssf/2010/11/report_finds_nj_12th_graders_a.html
- Mitra, D., Mann, B., & Hlavacik, M. (2016). Opting out: Parents creating contested spaces to challenge standardized tests. *Education Policy Analysis Archives*, 24(31), 1-20.
<https://doi.org/10.14507/epaa.24.2142>
- Mortimore, P., & Mortimore, J. (1998). The political and the professional in education: An unnecessary conflict? *Journal of Education for Teaching: International Research and Pedagogy*, 24(3), 205-219. <https://doi.org/10.1080/02607479819737>
- Murphy, J. P., Strothers, A. S., & Lugg, C. A. (2017). Jersey-Style Neo-Liberalism: Governor Christopher Christie, Crony Capitalism and the Politics of K-12 Education. *Peabody Journal of Education*, 92(1), 115-126.
- National Center for Educational Statistics. (2011). State Profiles. Washington, DC. Retrieved December 20, 2016, from: <https://nces.ed.gov/nationsreportcard/states/>
- National Education Association. (2011). Rankings of the States 2010 and Estimates of School Statistics 2011. Report. Retrieved from:
http://www.nea.org/assets/docs/HE/NEA_Rankings_and_Estimates010711.pdf
- New Jersey Department of Education, (2010). *New Jersey's application for the Race To The Top program*. NJDOE. Trenton, NJ.
- New Jersey Department of Education. (2013a). *A guide to the TEACHNJ Act*. Retrieved from website: <http://www.state.nj.us/education/AchieveNJ/TEACHNJ.Guide.pdf>
- New Jersey Department of Education. (2013b). *AchieveNJ: Educator evaluation and support in New Jersey*. Retrieved from website: <http://www.state.nj.us/education/AchieveNJ/intro/guide.pdf>
- New Jersey Education Association. (2013, July 11). *NJEA repeats its call for gradual rollout of New Jersey*. Retrieved from website: <http://www.state.nj.us/education/AchieveNJ/intro/guide.pdf>
- Northern, A. M., Scull, J., & Zeehandelaar, D. (2012). How strong are US teacher unions? A state-by-state comparison. Washington D.C.: Thomas B. Fordham Institute.
- O'Brien, T. L., & Pizmony-Levy, O. (2016). Going Public, Gaining Credibility Student Perceptions of Publicly Engaged Scholars. *Sociological Perspectives*, 59(2), 246-269.
<https://doi.org/10.1177/0731121415586634>
- O'Neill, E. (2011, September 4). NJEA claims New Jersey public school students have best Advanced Placement scores in nation. Retrieved from: <http://www.politifact.com/new-jersey/statements/2011/sep/04/new-jersey-education-association/njea-claims-new-jersey-public-school-students-have/>

- Peterson, P. E., Henderson, M. B., West, M. R., & Barrows, S. (2016). Education Next — Program on Education Policy and Governance. Retrieved from: <http://educationnext.org/files/2016ednextpoll.pdf>
- Pizmony-Levy, O. (under review). Compare Globally, Interpret Locally: International Assessments and News Media in Israel.
- Pizmony-Levy, O., & Cosman, B. (2017). How Americans view the Opt Out movement. [Research Report]. New York: Teachers College, Columbia University.. <https://doi.org/10.7916/D8C53Z6K>
- Pizmony-Levy, O., & Ponce, A. (2012). Framing strategies and public support for the legalization of marriage between two people of the same sex. *Sociological Perspectives*, 56(2), 169-190. <https://doi.org/10.1525/sop.2013.56.2.169>
- Pizmony-Levy, O., & Saraisky, N. G. (2016). Who Opt's Out and Why? Results from a national survey on opting out of standardized tests. [Research Report]. New York: Teachers College, Columbia University. <https://doi.org/10.7916/D8K074GW>
- Richardson, J., & Bushaw, W. (2015). The 47th Annual PDK/Gallup Poll of the Public's Attitudes Toward the Public Schools. Bloomington, IN: Phi Delta Kappa International.
- Rizzo, S. (2012, June 25). Sweeping N.J. teacher's tenure bill passes Legislature, heads to Gov. Christie's desk. Retrieved from: http://www.nj.com/news/index.ssf/2012/06/sweeping_nj_teachers_tenure_bi.html
- Rundquist, J. (2009, October 7). Issues facing N.J. public schools, students tackled by gubernatorial candidates. Retrieved from: http://www.nj.com/news/index.ssf/2009/10/issues_facing_nj_public_school.html
- Sarason, S. B. (1990). *The Predictable Failure of Educational Reform: Can We Change Course before It's Too Late?* San Francisco: Jossey Bass.
- Sarason, S. B. (1996). *Revisiting the culture of school*. New York: Teachers College Press.
- Spillane, J. P. (2004). *Standards deviation: How schools misunderstand education policy*. Cambridge, Mass: Harvard University Press.
- Spillane, J. P., Reiser, B. J., & Reimer, T. (2002). Policy implementation and cognition: Reframing and refocusing implementation research. *Review of Educational Research*, 72(3), 387-431. <https://doi.org/10.3102/00346543072003387>
- Springer, J. (November 11, 2009). Christie tells NJEA “they need to get realistic”. Retrieved from: <http://www.bluejersey.com/tag/1220/page/16/>
- Steele, J. L., Hamilton, L. S., & Stecher, B. M. (2010). *Incorporating Student Performance Measures into Teacher Evaluation Systems*. Technical Report. RAND Corporation.
- The Editorial Projects in Education Research Center. (2010). *Quality Counts 2010*. Retrieved from <https://www.edweek.org/media/ew/qc/2010/QualityCounts2010>
- Tucker, P. D., & Stronge, J. H. (2005). Linking teacher evaluation and student learning. ASCD.
- U.S. Census Bureau. (2010). State & county quick facts. Retrieved December 10, 2016, from: <http://quickfacts.census.gov>.
- Weisberg, D., Sexton, S., Mulhern, J., Keeling, D., Schunck, J., Palcisco, A., & Morgan, K. (2009). The widget effect: Our national failure to acknowledge and act on differences in teacher effectiveness. Brooklyn, NY: The New Teacher Project.
- Woody, E. L., Russell, J., Park, S., Kafka, J., & Butts, M. (2004). *Voices from the Field: Educators Respond to Accountability*. Berkeley, CA: Policy Analysis for California Education.
- Woolsey, A. (2014). *Teacher Perceptions of Implementation Effectiveness of TEACH N.J.* Unpublished master's thesis, Teachers College Columbia University, New York, New York.

About the Authors

Oren Pizmony-Levy

Teachers College, Columbia University

pizmony-levy@tc.columbia.edu

Oren Pizmony-Levy is an Assistant Professor in the Department of International and Transcultural Studies. He holds a PhD in sociology and comparative and international education from Indiana University-Bloomington. His research and teaching focus on the intersection between education and social movements. One line of research examines the accountability movement and its role in the emergence of international assessments of student achievement (e.g., TIMSS, PIRLS, and PISA). Another line of research examines the Opt Out movement in the United States.

Ashley Woolsey

Teachers College, Columbia University

Ashley.woolsey@gmail.com

Ashley Woolsey completed her Masters in International and Comparative Education at Teachers College, Columbia University (MA' 14). Her research interests include policy planning and implementation. This article is based on data she collected for her thesis titled "Teacher Perceptions of Implementation Effectiveness of TEACH N.J." She currently works as a middle school social studies and special education teacher at Memorial Middle School in Fair Lawn, NJ. Ashley participates in various initiatives to improve education policy implementation and effectiveness at the local and state levels.

About the Guest Editors

Jessica Holloway

Deakin University

Jessica.holloway@deakin.edu.au

Jessica Holloway, Ph.D., is a post-doctoral research fellow at the Centre in Research for Educational Impact (REDI) at Deakin University. She draws on post-structural theory to understand contemporary modes of accountability and its production of new teacher and leader subjectivities. Her current project, entitled *Teacher Leaders and Democracy: An International Study*, looks at modes of distributive leadership in U.S. and Australian schools.

Tore Bernt Sørensen

University of Bristol

t.b.sorensen@bristol.ac.uk

Tore Bernt Sørensen completed his doctorate at the Graduate School of Education, University of Bristol, UK, in 2017 with the dissertation "*Work In Progress: The Political Construction Of The OECD Programme Teaching And Learning International Survey*". Tore's research centres on comparative studies of education governance in a global context. Tore has a background as teacher and teacher trainer in Denmark. Before starting his doctorate, he worked in the Analysis and Studies Unit of the European Commission's Directorate-General for Education and Culture.

Antoni Verger

Universitat Autònoma de Barcelona

Antoni.verger@uab.cat

Antoni Verger is associate professor at the Department of Sociology of the UAB. A former post-doctoral fellow at the University of Amsterdam, Antoni's research analyses the relationship between global governance institutions and education policy, with a focus on the study of public-private partnerships and accountability policies in education. Currently, he is coordinating the research project REFORMED - *Reforming Schools Globally: A Multiscalar Analysis of Autonomy and Accountability Policies in the Education Sector* (ERC StG, 2016–2021).

SPECIAL ISSUE

Global Perspectives on High-Stakes Teacher Accountability Policies

education policy analysis archives

Volume 25 Number 87

August 21, 2017

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 A1 (Brazil), SCImago Journal Rank; SCOPUS, Socolar (China).

Please send errata notes to Audrey Amrein-Beardsley at Audrey.beardsley@asu.edu

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