



EDUCATION

THE ARTS
CHILD POLICY
CIVIL JUSTICE
EDUCATION
ENERGY AND ENVIRONMENT
HEALTH AND HEALTH CARE
INTERNATIONAL AFFAIRS
NATIONAL SECURITY
POPULATION AND AGING
PUBLIC SAFETY
SCIENCE AND TECHNOLOGY
SUBSTANCE ABUSE
TERRORISM AND
HOMELAND SECURITY
TRANSPORTATION AND
INFRASTRUCTURE
WORKFORCE AND WORKPLACE

This PDF document was made available from www.rand.org as a public service of the RAND Corporation.

[Jump down to document](#) ▼

The RAND Corporation is a nonprofit research organization providing objective analysis and effective solutions that address the challenges facing the public and private sectors around the world.

Support RAND

[Purchase this document](#)

[Browse Books & Publications](#)

[Make a charitable contribution](#)

For More Information

Visit RAND at www.rand.org

Explore [RAND Education](#)

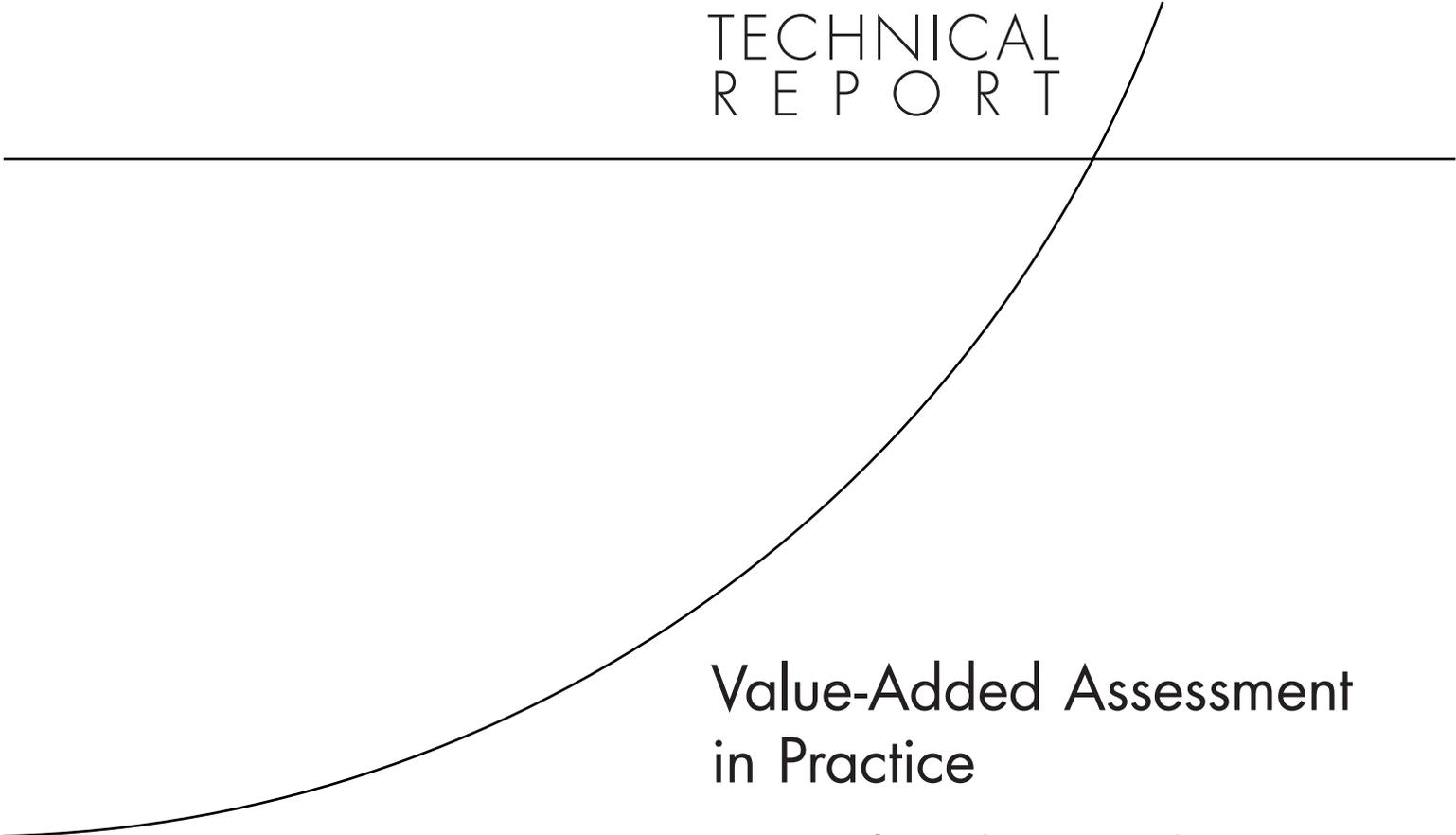
View [document details](#)

Limited Electronic Distribution Rights

This document and trademark(s) contained herein are protected by law as indicated in a notice appearing later in this work. This electronic representation of RAND intellectual property is provided for non-commercial use only. Unauthorized posting of RAND PDFs to a non-RAND Web site is prohibited. RAND PDFs are protected under copyright law. Permission is required from RAND to reproduce, or reuse in another form, any of our research documents for commercial use. For information on reprint and linking permissions, please see [RAND Permissions](#).

This product is part of the RAND Corporation technical report series. Reports may include research findings on a specific topic that is limited in scope; present discussions of the methodology employed in research; provide literature reviews, survey instruments, modeling exercises, guidelines for practitioners and research professionals, and supporting documentation; or deliver preliminary findings. All RAND reports undergo rigorous peer review to ensure that they meet high standards for research quality and objectivity.

TECHNICAL
REPORT



Value-Added Assessment in Practice

Lessons from the Pennsylvania
Value-Added Assessment System
Pilot Project

Daniel F. McCaffrey, Laura S. Hamilton

Supported by the Carnegie Corporation of New York,
the Ewing Marion Kauffman Foundation, the National Education Association,
and the Pennsylvania State Education Association

The research described in this report was conducted within RAND Education, a division of the RAND Corporation. It was funded by the Carnegie Corporation of New York, the Ewing Marion Kauffman Foundation, the National Education Association, and the Pennsylvania State Education Association. Additional funding came from the Connecticut Education Association, Education Minnesota, and the Ohio Education Association.

Library of Congress Cataloging-in-Publication Data

McCaffrey, Daniel F.

Value-added assessment in practice : lessons from the Pennsylvania Value-Added Assessment System pilot project / Daniel F. McCaffrey, Laura S. Hamilton.

p. cm.

Includes bibliographical references.

ISBN 978-0-8330-4236-1 (pbk. : alk. paper)

1. Educational tests and measurements—Pennsylvania. 2. Educational accountability—United States. 3. Educational indicators—United States. I. Hamilton, Laura S. II. Title.

LB3052.P46M33 2007

371.26'2—dc22

2007038195

The RAND Corporation is a nonprofit research organization providing objective analysis and effective solutions that address the challenges facing the public and private sectors around the world. RAND's publications do not necessarily reflect the opinions of its research clients and sponsors.

RAND® is a registered trademark.

© Copyright 2007 RAND Corporation

All rights reserved. No part of this book may be reproduced in any form by any electronic or mechanical means (including photocopying, recording, or information storage and retrieval) without permission in writing from RAND.

Published 2007 by the RAND Corporation
1776 Main Street, P.O. Box 2138, Santa Monica, CA 90407-2138
1200 South Hayes Street, Arlington, VA 22202-5050
4570 Fifth Avenue, Suite 600, Pittsburgh, PA 15213-2665
RAND URL: <http://www.rand.org>
To order RAND documents or to obtain additional information, contact
Distribution Services: Telephone: (310) 451-7002;
Fax: (310) 451-6915; Email: order@rand.org

Summary

Introduction

The use of student achievement data for decisionmaking is currently a focus of school and district reform efforts across the United States. Emphasis on data has grown as result of an increasing emphasis on using test scores to evaluate school performance, a use that is central to the No Child Left Behind¹ (NCLB) accountability provisions. Data use has also been facilitated by improved data systems and analysis tools. This technology has contributed to the growing use of value-added assessment (VAA) systems²—collections of complex statistical techniques that use multiple years of test-score data on students to try to estimate the causal effects of individual schools or teachers on student learning. The Tennessee Value-Added Assessment System (TVAAS) is the most widely known application of VAA in the United States, and efforts to extend or replicate this model are currently under way in other states and school districts.

VAA can be used to support external accountability and monitoring of school performance. It can also be used as a tool for promoting school improvement by providing data to help school personnel make decisions. To date, most VAA programs have emphasized the latter use. In these contexts, VAA is intended to contribute to better decisions about educational practice, which in turn should promote improved student achievement. This study is designed to evaluate the extent to which a VAA system achieves the goals of improving practice and student outcomes. It examines one recently adopted VAA system—the Pennsylvania Value-Added Assessment System, or PVAAS. Pennsylvania rolled out its system in four waves, which resulted in a quasi-experimental condition, with a subset of the school districts participating in PVAAS and a subset of possible comparison districts not in the program. This report describes an investigation of PVAAS that explores three related questions:

¹ Signed into law January 8, 2002, No Child Left Behind (NCLB) is the latest revision of the 1965 Elementary and Secondary Education Act. It establishes high learning standards for all students, including requirements that all students be proficient in reading and mathematics by 2013–2014. Among other provisions, it also requires that all students be tested against state standards in grades 3 to 8 and one high school grade in reading and mathematics and three times in their school career in science. The law mandates that schools be assessed on the basis of student test scores on their Adequate Yearly Progress toward the 2014 goals.

² Value-added assessment is sometimes referred to as value-added analysis, value-added modeling, or growth modeling. Because the Pennsylvania pilot program studied in this report is called the Pennsylvania Value-Added Assessment System, for consistency of terminology within the report, we use the term *value-added assessment* to refer to the value-added information created from test-score data and provided to schools.

1. What is the effect on student achievement of providing districts with information from a VAA system?
2. How does the use of data by educators whose districts participate in a VAA system differ from that of educators from nonparticipating districts?
3. How do educators respond to the VAA information they receive?

The first question is the causal question of primary interest. The second and third questions are intended to clarify the mechanisms through which provision of VAA information might affect practice and, ultimately, student achievement.

Methods

Because the pilot districts were not randomly chosen, the study matched the first two cohorts of PVAAS pilot districts to comparison districts by finding a sample with the smallest average distance between pilot and comparison districts in terms of demographic and historic test scores using an optimization algorithm. Overall, the matches for both cohorts were very similar in terms of over 100 variables describing the district's student demographics and historic achievement, district financing, and the populations living in each district as measured by the 2000 Census.

The effects of PVAAS on student achievement on the state's accountability test were measured by differences between students in the pilot and matched comparison districts. Several analyses test for effects, including models based on district aggregates, nonparametric methods, and mixed models with and without controls for student- and district-level variables.

The study surveyed all superintendents, 411 principals, and 2,379 teachers from the 93 study districts (47 PVAAS and 46 matched comparison districts) during the second half of the 2005–2006 school year. Over 85 percent of superintendents (or their designees), 58 percent of principals, and 44 percent of teachers responded to the survey. Responses from all educators are weighted by the inverse of the response probability, to account for differential nonresponse rates. Because many responding principals and teachers had little contact with PVAAS, we focus on educators who are “engaged” in the program (principals who saw the PVAAS reports and knew their school was participating in the program, or teachers who had heard of PVAAS and knew their school was participating in the pilot). We developed weights for comparison principals and teachers to match them to the engaged PVAAS samples on school and district variables.

Survey questions probed educators about their attitudes toward state tests and the state accountability system. They also asked educators about their use of test data for decisions and their training in the analysis and use of data. Items specifically for the educators in the pilot districts asked about PVAAS training, use of PVAAS data, and knowledge of the PVAAS methods. All survey instruments were reviewed by educators and state officials and were revised in response to their comments.

Study Limitations

A primary limitation of this study is the small number of pilot districts available for the study, which placed constraints on matching and limited our power for comparing pilot and comparison districts. Another limitation is the possibility that PVAAS and comparison districts

differed on unobserved variables. PVAAS districts needed to have district-wide testing, and we are unable to use such data when selecting comparison districts. For comparing student outcomes, matching on the extensive historical test-score data is likely to mitigate bias from such unobserved differences. However, educators in PVAAS districts tend to report greater emphasis on testing than their counterparts in comparison districts, and this is consistent with bias to preexisting difference in testing experiences. Also the study lacked longitudinal data on individual students. Although we matched on school-level test score trajectories, it is possible that individual students' achievement growth differed in pilot and comparison districts, and any such differences could not be controlled for by our matching procedure.

Low response rates among principals and teachers also limit our sample sizes and could introduce bias. Nonresponse weighting removes differences between respondents and nonrespondents on a large number of factors, but unobserved differences might remain after weighting. Another potential for bias from unobserved differences exist in the comparisons of engaged PVAAS principals and teachers and the educators from the comparison group. Although the engaged and comparison groups are similar on observed school and district variables, we lacked data on individual attributes, such as training or familiarity with data analysis, and remaining differences on such factors could bias comparisons. Finally, we studied the PVAAS districts in their initial years of the program participation. This design may not have provided enough time for school and district staff to learn to use the data effectively. Moreover, even if the use of PVAAS data is highly effective for students in schools and districts that are exposed to it over time, exposure might not have been sufficient in the PVAAS pilot.

Findings

PVAAS Did Not Affect Student Achievement

There were no statistically significant differences in student achievement between PVAAS pilot districts and matched comparison districts. In all comparisons across both cohorts, the differences in means between the pilot and comparison districts were generally small relative to the standard deviations in the scores, ranging from less than 1 percent to about 15 percent of a standard deviation. Moreover, for Cohort 2 districts, the differences between the scores for the PVAAS and comparison districts in the year before districts received their PVAAS reports were similar in direction and magnitude to the differences observed during the next two years. The results provide no evidence that participation in PVAAS affected student achievement.

District Administrators' Use of Achievement Data for Decisionmaking Was Similar in PVAAS and Comparison Districts

Analyses of the survey data suggest possible reasons for the lack of effects on achievement. At all three levels of the education system—district administrators, school principals, and classroom teachers—there was little evidence that use of achievement data differed between PVAAS and non-PVAAS districts, or that PVAAS information was being used in significant ways. Among district administrators, PVAAS participants were slightly more likely than nonparticipants to report that various forms of achievement data were useful for decisionmaking, but the only type of data for which the groups differed significantly was growth data. Both groups reported using data for a number of different decisions. Administrators in PVAAS districts were slightly more likely to support data analysis in their districts through provision of staff and professional

development, though the differences were not statistically significant. Although both groups reported receiving technical assistance with data analysis fairly rarely, among those who did receive it, PVAAS administrators were significantly more likely to rate it as useful. In addition, they were less likely than nonparticipants to view insufficient technology or lack of information about growth in achievement as hindrances to their ability to use data effectively.

District Administrators' Opinions of PVAAS Were Positive, But Use Was Limited

The opinions of PVAAS among administrators from pilot districts are generally favorable. A large majority (80 percent) stated that PVAAS provides accurate information about how the district is improving student achievement, compared with fewer than half who endorsed a similar statement about the Adequate Yearly Progress (AYP) measure. Similarly large majorities reported that it helps with communications with parents and helps school staff to see their efforts pay off. Slightly more than half reported that at least some schools in their districts look better with the PVAAS measure than they do using AYP status, so there is clearly a recognition that these sources of information can lead to different conclusions schools' performance. Three-quarters of administrators reported that PVAAS eliminates excuses for poor performance because it measures growth.

Despite their favorable opinions, administrators' use of PVAAS is not as widespread as might be expected. Ninety percent of administrators reported seeing the actual PVAAS reports, and about 70 percent of administrators reported giving teachers in their districts access to these reports. When asked about specific uses of PVAAS information, only a minority of administrators answered that they use PVAAS moderately or extensively in each case. PVAAS is most widely used for making curricular and professional development decisions and improvement planning. For these activities, administrators in about half the districts reported moderate or extensive use. For all decisions, reported use of PVAAS is substantially lower than for other data sources, particularly PSSA scores.

Many Principals Had Limited Experiences with PVAAS and Made Minimal Use of the Information It Provided

The most significant finding from the analysis of principal survey results is that 28 percent of principals in PVAAS districts do not know that their school is participating in the program, and another 14 percent have never seen the PVAAS reports. Because effects of PVAAS on principals' practices are likely to occur only if principals are knowledgeable of the program and the information it provides, we limited subsequent analyses to the 58 percent of principals who are engaged in PVAAS (i.e., they knew their schools were participating and had seen the reports).

Among principals engaged in PVAAS, attitudes about PVAAS are generally positive, though use of the information is somewhat limited. Nearly 80 percent feel PVAAS provides an accurate indication of how well their schools are improving achievement. A majority (60 percent) reported making changes to their leadership or school improvement plan based on PVAAS, and 56 responded that PVAAS helps motivate them. However, smaller percentages agreed or strongly agreed that PVAAS was discussed frequently at staff planning meeting (33 percent), caused the school to focus more on low-performing or high performing students (43 percent and 33 percent, respectively), was used to identify students at risk of not meeting the standards (42 percent), or helped with communications with parents (27 percent). A comparison with these principals' reports of use of other state and district test-score data indicates that PVAAS is not being used as extensively as these other data sources. In general, principals'

reports are consistent with those from district administrators, but principals are slightly less enthusiastic and reported slightly lower levels of understanding of the information compared with district administrators.

There Are Few Differences Between PVAAS and Comparison Principals

Principals' responses are important to understand because, as the instructional leaders of their schools, principals play an important role in ensuring that school staff use PVAAS information in ways that will contribute to improved student achievement. Our analyses compared the 58 percent of pilot principals who were engaged in PVAAS to the sample of comparison school principals that was weighted to match to the characteristics of the engaged principals' schools.

The survey results suggest a low level of engagement with PVAAS on the part of many principals, and few differences in the actions taken by principals participating in the PVAAS pilot program and their counterparts from nonparticipating schools. There were a few differences between the groups in their access to resources for data use: PVAAS principals were more likely than comparison principals to receive training on how to use test-score data for instructional planning and to receive information on data systems or guidance on selecting these systems. Other resources, such as professional development to help principals analyze data or to meet the needs of low-achieving students, were available to similar percentages of principals in both groups. Principals' perceptions of the factors that hindered their ability to use data were similar, with one exception: Over half of the comparison group principals (57 percent) reported that lack of data on student growth was a hindrance, but only 27 percent of the engaged pilot principals reported the same. This difference was much larger than any other differences between the groups on these items, and it parallels the finding for district administrators.

Teachers Are Not Engaged with PVAAS

To the extent that providing PVAAS information leads to improved student achievement, it is likely that this effect occurs in large part as a result of actions taken by teachers. As with principals, the most important finding from the teacher surveys is a relatively low level of engagement. Fewer than half of the surveyed teachers in PVAAS pilot districts reported that they had heard of PVAAS, and among those who had heard of the program, only half were aware of their schools' involvement in it. This lack of widespread knowledge of the program on the part of teachers provides one likely explanation for the lack of achievement differences between pilot and comparison districts.

Among the PVAAS teachers who were aware of the program and their school's involvement in it, there was wide variation in use of the information and level of understanding of it. For example, only a small minority understood that PVAAS was not part of schools' AYP calculations, and only about half expressed confidence in their understanding of the meaning of "a school effect" or in their ability to use PVAAS to guide their instruction. Comparisons of attitudes and practices related to data use suggest few differences between these PVAAS teachers and their counterparts in nonparticipating schools, though there is some evidence that the PVAAS teachers are more engaged with data use and test preparation in general.

Conclusions

This study was designed to shed light on the utility of a value-added assessment system for promoting educational improvement. The lack of effects of the program on student achievement are not surprising, given the limited implementation of the pilot program at the district, school, and classroom levels and the relatively short period of time during which the program was in place. The primary finding from the surveys is a lack of use, and in many cases even awareness, of PVAAS among educators—particularly teachers, the group most directly responsible for promoting student learning. This limited use is consistent with findings from research on other VAA systems, both in the United States and in England, where such systems have been in place for some time. The growing body of literature on VAA implementation suggests that providing educators with assistance to help them understand and use the data is likely to be one of the primary challenges associated with adopting such systems. In particular, the survey results from this study indicate a need for training focused on how to make use of the data rather than simply how to interpret it. Pennsylvania has revised its training materials since this study was conducted and is taking steps to improve the support provided to assist educators in their efforts to use the data for improving curriculum and instruction. These actions might increase the likelihood that PVAAS will contribute to more-effective decisionmaking, which in turn might eventually influence student outcomes.

The findings also suggest that one of the challenges to using PVAAS is the need for teachers and administrators to respond to accountability pressures imposed by NCLB. Although advocates of VAA systems often argue that the information these systems provide can be helpful for meeting NCLB goals, the immediate pressure to worry about whether students are proficient may cause educators to focus more of their attention on the state test scores and proficiency levels rather than on the PVAAS information. Combined with the results of other studies on VAA systems, the results of this study could be cautiously interpreted as suggesting that the lack of accountability attached to PVAAS might contribute to the relatively low use. The use of students' growth in achievement as the basis of accountability is increasing in some places, particularly through pay-for-performance measures, such as those promoted by the federal government's Teacher Incentive Fund and through the U.S. Department of Education's Growth Model Pilot program, which allows states to incorporate growth data into their AYP calculations. More generally, achievement growth data are becoming more widely available as a result of improved state and district data systems and analysis tools, and data-driven decision-making is being emphasized in many school and district reform efforts. Together, these trends are likely to increase educators' awareness of and interest in using VAA data.