

Catalyst for Change

PAY FOR PERFORMANCE IN DENVER
FINAL REPORT

ctac ctac ctac ctac ctac ctac ctac ctac ctac ctac ctac ctac JANUARY 2004

ABOUT CTAC:

The Community Training and Assistance Center is a national not-for-profit organization with a demonstrated record of success in urban communities. It focuses on developing leadership, planning and managerial expertise within community-based organizations, school systems, collaborative partnerships, state and municipal governments, and health and human service agencies. Since 1979, CTAC has provided assistance to hundreds of community-based organizations, coalitions and public institutions in the United States and several other countries.

CTAC's staff is comprised of nationally recognized executives, educators, policy makers and organizers who have extensive experience working with city, county and state agencies, educational institutions, federal legislative bodies, not-for-profit organizations, philanthropic institutions and the private sector.

© COMMUNITY TRAINING AND ASSISTANCE CENTER
BOSTON, MASSACHUSETTS
JANUARY 2004

ALL RIGHTS RESERVED

Catalyst for Change

PAY FOR PERFORMANCE IN DENVER
FINAL REPORT

ctac ctac ctac ctac ctac ctac ctac ctac ctac ctac ctac ctac ctac JANUARY 2004

Acknowledgements

The Community Training and Assistance Center wishes to thank the sponsors of this research study, the Denver Public Schools Board of Education and the Denver Classroom Teachers Association. Their commitment to bringing quality and accountability to public education is exemplary. During the four years of the pilot, they have been able to maintain a critical focus on the development and implementation of Pay for Performance.

CTAC would like to acknowledge the many people who have contributed both to the pilot and to making this study possible. The leadership of the Board of Education has been pivotal to the pilot. This has involved all board members: President Les Woodward, former President Elaine Gantz Berman, Lucia Guzman, Bruce Hoyt, Michelle Moss, Kevin Patterson, Theresa Peña, and previous board members Sue Edwards, Sharon Macdonald, James Mejia and Bennie Milliner. The leadership of the Association has also been essential, including President Rebecca Wissink, former President Andrea Giunta and Executive Director Bruce Dickinson. Thank you for being willing to share this professional challenge with CTAC; we have appreciated your support and confidence.

The CTAC research study provides a solid foundation for the ideas and desires of those who wish to create a new design for a public school system. We wish to thank those individuals who, during these past four years, have contributed so much to this effort. We have been assisted by many leaders of the Denver Public Schools, including Superintendent Jerry Wartgow and former superintendents Bernadette Seick, Chip Zullinger, Sharon Johnson and Irv Moskowitz. They have been uniformly helpful.

The coordination and implementation of the many facets of this study could not have been conducted without the assistance of Design Team Leader Brad Jupp, current Design Team members Cheryl Betz, Henry Roman and Shirley Scott, former members Pat Sandos and Rebecca Wissink, and assistant Ginger Doyle. Their drive, passion and sense of urgency are inspiring.

We would like to extend a special thank you to the unsung heroes of the pilot and study—the pilot school teachers, administrators and parents. They have made special efforts on behalf of students.

We have been assisted by many current and former members of the central administration. While there are too many such people to mention everyone individually, we would like to identify the following with our appreciation: Assistant Superintendents Richard Allen, Wayne Eckerling and Andre Pettigrew, Chief Academic Officer Sally Mentor Hay, and former administrators Beth Celva, Craig Cook, Larry Lindauer and Dianne Lefly. Key contributions have also been made by James McIntosh and Richard Machado of Assessment and Testing; Ethan Hemming of Educational Services; Joyce Fell, Peggy Flynn and Billy Gould of Human Resources; and Director Aaron Gray and Flor Amaro of Community Relations. They have all been generous with their time and attention.

There were two groups which have assumed additional responsibility for the progress of the pilot. They are the pilot Leadership Team and the Joint Task Force on Teacher Compensation. The participating individuals have dedicated an inordinate amount of their professional and personal time to understand and facilitate important elements of this project.

CTAC is grateful for all those listed above, and the many other people who helped to make this study a reality.

This study has been made possible through the generous support of Rose Community Foundation, The Broad Foundation, The Daniels Fund, The Sturm Family Foundation, Jay and Rose Phillips Family Foundation, The Denver Foundation, Donnell-Kay Foundation, and The Piton Foundation.

The findings, analyses and recommendations expressed in this study are those of the Community Training and Assistance Center.

© Community Training and Assistance Center, January 2004

Credits

This study was conducted and prepared by the Community Training and Assistance Center of Boston, Massachusetts.

Denver Project Team

Principal Study Authors

William J. Slotnik
Maribeth D. Smith

Contributing Authors

Roberta J. Glass
Barbara J. Helms, Ph.D.

Field Operations Director

Donald W. Ingwerson, Ph.D.

Team Members

Denise A. Bell
Lee Bray
Peggie L. Brown
Robin C. Burr, Ph.D.
Mary Ann Crayton
William M. Eglinton
Donald B. Gratz, Ph.D.
Mimi Howard
William C. Lannon
Juan J. Maya
Candy Miller

Marcia J. Plumleigh, Ph.D.

Lynn J. Stinnette
Martha Swartz
Julia D. Thomé
Barbara Zeno

Statistical Measurement Task Force

Robert H. Meyer, Ph.D.
John B. Willett, Ph.D.

Statistical Review

Kathleen Calore

Contents

4

Executive Summary

12

CHAPTER I

Overview

18

CHAPTER II

Pay for Performance

26

CHAPTER III

Research Design

42

CHAPTER IV

Objectives: The Nexus

64

CHAPTER V

The Teacher Perspective

82

CHAPTER VI

Quantitative Analyses

104

CHAPTER VII

Catalyst for Change

112

CHAPTER VIII

Organizational Alignment and System Quality

128

CHAPTER IX

Issues and Recommendations

136

CHAPTER X

National Implications

148

Appendix

Executive Summary

The Denver Public Schools (DPS) and the Denver Classroom Teachers Association (DCTA) jointly sponsored the Pay for Performance (PFP) pilot. This four-year pilot, conducted during the 1999–2003 school years, focused on developing a direct link between student achievement and teacher compensation. The sponsoring parties

committed to implementing the pilot and to studying the results of this initiative. *Catalyst for Change* is the final report of the results of this innovative collaboration on behalf of students and teachers.

The centerpiece of Pay for Performance in Denver has been the teacher objectives at the 16 pilot schools. Teachers developed two annual objectives based on student achievement that required the approval of the principal. Teachers received additional compensation if they met their objectives. The entire pilot was stewarded by a four-person Design Team that included district and union appointees. The pilot involved 13% of Denver's schools. These schools represented a cross section of the district's client and service base. In microcosm, the pilot schools presented the challenges of the broader district. Their experiences provided a foundation for a study of pilot impact.

The Community Training and Assistance Center (CTAC) was commissioned in November 1999 to conduct a comprehensive study of the impact of the pilot and to provide the technical assistance necessary to help assure a pilot of quality and integrity. The study's mid-point report, *Pathway to Results: Pay for Performance in Denver* was presented in December 2001. *Catalyst for Change* is the summative report.

The study has four core components. First, it examines the impact of the pilot on student achievement based on two independent assessments. Second, it examines teacher objectives: their substance, quality and relationship to student achievement. Third, the study considers the effect of a range of student, teacher, and school factors on the results of the pilot. Finally, it identifies the broader institutional factors that have affected implementation.

The data for this four-year study has been extensive. CTAC conducted surveys and examined responses from 2,870 teachers, parents, administrators and others involved in the pilot; conducted and reviewed 615 interviews; and organized and analyzed tens of thousands of student records plus teacher and demographic data for the pilot and control schools. The study also involved the careful review of artifact and documentary data and the detailed examination of 4,012 teacher objectives. Human resource records and school documentation were linked to school, teacher, and student information to create a comprehensive database. Finally, the study included hundreds of hours of observations, from classrooms to boardrooms, which contributed a strong sense of the context and the everyday work of the pilot.

The nature of a pilot is experimentation. Denver conducted the pilot in order to examine the effectiveness and impact of a new direction. By building around the objective setting process, the pilot design was straightforward and honored both teacher judgment and perspective. The implementation of the pilot, though, was necessarily more complex. As teachers were learning about developing and meeting measurable annual objectives, the schools and the district were learning about the necessary alignment of the curriculum, assessment, student data, human resources and other parts of the larger system with Pay for Performance. The alignment and strengthening of these functions proved complicated and challenging.

The pilot has demonstrated that the focus on student achievement and a teacher's contribution to such achievement can be a major trigger for change—if the initiative also addresses the district factors that shape the schools. The findings show promising results in many of the areas studied. However, the pilot's implementation also revealed areas of practice and policy that will need to be strengthened further in order to continue improving student achievement as the key elements of Pay for Performance are taken to scale in the district.

Highlighted below are CTAC's analyses, findings and recommendations. All of the recommendations are crafted to meet the standard of increasing the manageability, fairness and sustainability of Pay for Performance. The report also examines the national implications of Denver's Pay for Performance pilot. The issues are complex and multi-faceted, and are discussed in full detail in the chapters of the report.

A. Primary Findings

Impact on Student Achievement

- At all three academic levels—elementary, middle, and high school—higher mean student achievement in the pilot schools is positively associated with the highest quality objectives. Students whose teachers had excellent objectives, based on a four-level rubric developed by CTAC, achieved higher mean scores than students whose teachers' objectives were scored lower on the rubric. This holds true on most tests of the *Iowa Test of Basic Skills (ITBS)* and the *Colorado Student Assessment Program (CSAP)*.

At the elementary school level, students of teachers with excellent objectives (rubric level 4) had significantly higher mean normal curve equivalent (NCE) scores on the *ITBS* Reading, *ITBS* Language, and *CSAP* Math tests than students of teachers with lower rubric levels.

At the middle school level, students of teachers with level 4 objectives had significantly higher mean scores than students of teachers with level 3 and level 2 objectives on the *CSAP* Math test.

At the high school level, there were significantly higher mean scores on the *ITBS* Reading and *CSAP* Writing tests at Manual High School and on the *CSAP* Writing test at Thomas Jefferson High School for students whose teachers had level 4 objectives.

Six other tests (one elementary, three middle school, and two high school level) show a positive relationship between highest quality objectives and higher achievement level that is not statistically significant.

- Pilot students were compared to control students, estimating the change in mean NCE scores over time (e.g., from the baseline year through the end of the pilot) on three tests of the *ITBS* and three tests of the *CSAP*. Two-stage hierarchical linear modeling (HLM) was used to account for correlation between observations and to adjust for differences in school and student characteristics. The use of NCEs makes it possible to determine whether students are performing better than expected based on previous scores—attaining more than an expected year of growth—or not.

The effect of the pilot varies by level of school:

The pilot elementary school mean NCE scores declined on the *ITBS* Math, *CSAP* Reading and *CSAP* Math tests. The control school mean NCE scores declined on the *ITBS* Math test and increased on the *CSAP* Writing test. These results were statistically significant. The performance of the pilot students was lower than the controls on the tests except on the *ITBS* Language.

The pilot middle school students performed significantly higher than the controls on the *ITBS* Reading, *CSAP* Writing, and *CSAP* Math tests. The average NCE scores of middle school pilot students increased significantly over time (i.e., attained more than a year's expected growth) on the *ITBS* Reading, *CSAP* Writing, and *CSAP* Math tests. The controls experienced statistically significant declines in mean NCE scores on the three *ITBS* tests over the course of the pilot and statistically significant increases on the three *CSAP* tests. On the *ITBS* Language test, both the pilots and controls decreased by comparable amounts.

The high school pilots and controls experienced statistically significant increases in mean NCE scores over time on most tests. Manual High School students had significantly higher increases than the control school students on the *ITBS* Language and Math tests. Thomas Jefferson High School students performed significantly higher than the control school students on the *ITBS* Language, *ITBS* Math and the *CSAP* Reading tests and significantly lower on the *ITBS* Reading test.

- Meeting two objectives is positively associated with higher mean achievement scores.

At the elementary schools, the students of teachers who met two objectives had significantly higher mean NCE scores on all six tests than students of teachers who met one objective.

At the middle schools, meeting one or two objectives was associated with significantly higher mean NCE scores than meeting no objectives on the *ITBS* Reading and Language tests. Similar positive relationships were seen on the *ITBS* Math and *CSAP* Math tests but they are not statistically significant.

At the high schools, the students of teachers who met two objectives (at both Manual and Thomas Jefferson) had significantly higher mean NCE scores than students of teachers who met one objective or no objectives on the *ITBS* Reading test. Similar positive relationships are found on the *ITBS* Math, *CSAP* Writing, and *CSAP* Math tests at Manual High School and on *ITBS* Math, *CSAP* Reading, and *CSAP* Math at Thomas Jefferson High School but they are not statistically significant.

- Student achievement rises as length of teacher participation in the pilot rises.

Elementary students whose teacher had been in the pilot for two, three, and four years had mean *ITBS* Reading scores 0.8, 1.3, and 2.2 NCEs higher than students of one-year teachers. Elementary students of four-year teachers also had mean *ITBS* Math scores significantly higher than those of one-year teachers.

Middle school students of two-year teachers scored 2 NCEs higher on average and students of three-year teachers scored 3.2 NCEs higher than students of one-year teachers.

High school students of two-year pilot teachers scored higher on *ITBS* Reading, *ITBS* Math, and *CSAP* Reading at both pilot high schools, but the difference was only statistically significant for the Thomas Jefferson *ITBS* Reading exam.

- The pilot has been the catalyst for developing a fundamentally new compensation plan for teachers in Denver which is based, in part, on student achievement.

Impact of Objectives

- The percent of teachers who developed objectives that were rated at the two highest levels of the rubric increased steadily over the course of the pilot. The particular improvement in the objectives in the final year of the pilot is largely attributable to greater attention to learning content in the objectives. By the close of the pilot, 28% of the objectives were at level four (excellent) and 44% were at level three (adequate).
- There is a significant increase in the quality of the objectives as the number of years a classroom teacher participated in the pilot increases.
- Teachers met their objectives at a high rate. The data collected by the district over the four years of the pilot show that from 89% to 93% of the teachers met one or more objectives and were awarded additional compensation.

- As teachers gained more years of experience in the pilot, their ability to meet their objectives also increased significantly. One-year pilot participants met 89% of their objectives. The success rate rose to 98% for four-year participants.
- Educational background and years of experience in the Denver Public Schools are related to whether classroom teachers met their objectives.

Certified teachers met 92% of their objectives, while teachers participating in the Teacher-in-Residence alternative certification program met 83% of their objectives.

First year teachers met 86% of their objectives, while teachers with two or more years of experience met 92% of their objectives.

Teachers with 15 or more years of experience in the Denver Public Schools met their objectives at lower rates (85%) than teachers with fewer than four years (95%), four to 10 years (90%), or 11 to 14 years (95%) of experience.

- There are similarities between pilot school teacher objectives and control school teacher goals. However, the control school teachers make less use of baseline data. Some of the similarities are attributable to the pilot's objective setting protocol being in full or partial use in nearly one-third (32%) of the control school goals reviewed in 2002-2003.
- The objectives and their learning content are not included as part of the strategies in the school improvement plans.

Perceptions of Participants and Other Parties

- The pilot has significantly increased the school and district focus on student achievement. This focus has increased with each succeeding year of pilot implementation.
- Teachers indicate that they have greater access to student achievement data and that they use the data more effectively, particularly baseline data, to establish growth expectations, to focus earlier on students who may need more assistance and to monitor progress.
- Most pilot teachers do not attribute changes in their core classroom instructional practices to the pilot. Teachers indicate that they did not receive a mandate to make such changes.
- Most teachers feel that cooperation among teachers has improved or stayed the same at the pilot schools.
- Pilot teachers are less fearful of pay for performance than control school teachers. By the end of the pilot, pilot participants were more likely to offer suggestions for improvement than to indicate that pay for performance was not viable.
- Pilot teachers continued throughout the pilot to raise issues of fairness and trust in the objective setting and review process. However, they believe that it is possible to set fair objectives.
- The quality of interaction between the principals and teachers is pivotal to the implementation of Pay for Performance at the school sites. Teachers in the pilot schools believe that there are inconsistencies from school to school in how principals review and analyze progress on teacher objectives. Principals indicate that there is a lack of clarity regarding their role and authority and a need for targeted professional development.
- Parents indicate that a teacher's contribution to student achievement should be rewarded in financial terms.

- The vast majority of parents (94%) and teachers (93%) feel that more than one measure of student achievement should be used to determine teacher performance.
- Participants value the training they received, but express a need for more professional development based on the specific student achievement levels in the individual schools and classrooms and the instructional challenges of meeting objectives.

Institutional Factors

- The DPS/DCTA collaboration on behalf of student achievement has been significant. This collaboration has been pivotal to the development of Pay for Performance despite changes in district leadership and structure.
- Pay for Performance has enabled issues which have adversely affected district progress, sometimes for many years, to be put on center stage. Operating in a climate protected by external supporters and internal reformers, the pilot provided a vehicle for problems to be discussed, analyzed and acted upon. These actions have helped the district to develop an increased capacity to make mid-course corrections.
- Teachers and principals were provided with multiple opportunities through the study to influence the course of the pilot. For many, this was a marked and positive departure from past district practice.
- The Design Team contributed significantly to the progress of the pilot.
- District support systems were seriously challenged by the implementation of Pay for Performance. Many opportunities for change were identified and district action resulted. Challenges of organizational alignment still lie ahead for the district.
- The turnover in leadership positions during the course of the pilot, particularly at the level of the pilot school principals and the superintendency, contributed to some of the concerns related to trust and institutional priority that have affected the implementation of the pilot.
- The lack of an agreed-upon and aligned portfolio of district assessments for measuring student achievement meant that 166 identifiable assessments were used to measure progress in meeting objectives, and 256 teachers used generally referenced measurements, in the last year of the pilot.
- The task of linking student achievement results to specific teachers has proven more challenging than originally anticipated by the district. As pilot efforts go to a broader scale of implementation in the district, this type of data capacity will be greatly needed.
- Several factors, including the state and national high stakes testing environment and the district's experiences with pay for performance for administrators, adversely affected the climate for implementing the pilot.

B. Recommendations

Issue One: Alignment

Since the purpose of the district's major initiatives is to increase student achievement, the organization will benefit from continuing to align its initiatives around that goal in a clear and purposeful manner.

Recommendations include:

- *Bring the objective setting to scale with instructional support.* Crafting objectives is a key initial step in planning and delivering instruction. It is not merely an exercise in writing. It will be important to align instructional support to assist teachers to meet the specific targets in their objectives.

- *Strengthen the linkage between classroom objectives, school improvement plans and district standards and goals.* To the extent that objectives, school plans, and standards and goals are mutually reinforcing, the implementation of all three will be strengthened and there will be greater clarity of purpose.
- *Increase the connection between student information systems and human resources systems.* Building on the district's progress in supporting the pilot schools, this recommendation focuses on establishing greater district-wide linkages among the data systems for student assessment, planning, and human resources. These linkages are critical for any effort that seeks to examine the contribution of a teacher to student achievement.
- *Project the costs of changing internal practices and requirements.* There are direct financial costs to implementing Pay for Performance systemwide. To maintain teacher commitment, the momentum of the pilot needs to continue under a range of financial circumstances.

Issue Two: Assessment

A portfolio and appropriate usage of high quality assessments are the marks of a district that is aligned and accountable for its students. Using student assessments for a new purpose—compensation—requires greater assessment and data capacity, as well as a skillfully supervised and consistently administered effort at the school sites so that all students have regular assessments.

Recommendations include:

- *Expand the district's assessment strategy.* The existing assessment plan should become a component of a more comprehensive assessment strategy that includes aligning fair and valid assessments with the district's standards and curricula, ensuring practitioner input into the selection and use of assessments, and establishing a clear direction about who is to be assessed.
- *Define which assessments can be used for objective setting and compensation purposes.* There continues to be a need for a rational level of prescriptive direction regarding which assessments can be used as part of any new system that involves pay for performance. A pay for performance system or companion educational initiative that has too many allowable assessments will be unmanageable, will cause discord and will fail to promote valid increases in student achievement.
- *Make the use of multiple measures a developmental priority.* For four years, Denver teachers and site administrators have been raising questions about the fairness and accuracy of single measures. The charge is for the district to develop a means to link several assessments together to more meaningfully identify student progress.
- *Increase the district capacity to disaggregate and analyze student achievement data.* Regular analyses of these data strengthen decisions about delivering classroom instruction, developing school improvement plans, and managing strategically at the district level and, in the era of No Child Left Behind, it will provide communities with high quality information about its schools.
- *Convene select urban districts to analyze and take action on problems in assessments.* As a result of the Pay for Performance pilot, Denver is positioned as a national leader in the area of tying teacher compensation, in part, to student achievement. The ensuing challenges that Denver faces are shared in common by other districts. Denver should convene a small number of urban districts and assessment specialists to guide test developers to link their efforts more directly to the growing needs of urban districts.

Issue Three: Professional Development

Virtuosity in teaching is the goal of professional development for teachers. Initiatives often assume that teachers will embrace the concept of the reform and change their practices when, in fact, they may not know about new practices that would be appropriate. Both educational research and the pilot outcomes indicate that there is a profound connection between objectives based on learning content, a teacher's subject matter knowledge, specific teaching practices, and student achievement.

Recommendations include:

- *Establish district standards for professional development.* Establishing quality standards for professional development is a natural and necessary complement to instructional priorities. They should be tied to the Colorado Teacher Standards, research about best teaching practices, the district's curriculum standards, and the assessment strategy described above. This work will result in a roadmap for providing professional development services and ensuring quality control.
- *Predicate professional development on student achievement.* The priorities for professional development need to be based on continuous reviews of student achievement results by school staffs. Such a review identifies schoolwide, classroom and individual student instructional needs and instructional areas which need to be updated or improved. This, in turn, may reveal areas in which school staff or the community may need assistance in meeting these needs.
- *Create opportunities for teachers and principals to shape professional development.* One of the key findings from the pilot is that the ability of site practitioners to influence implementation contributes to improvements in the overall effort. This kind of involvement increases the prospects of professional development to effectively target teacher needs, school priorities and district goals.

Issue Four: Leadership

Many reforms fail for lack of sustained leadership. The Board of Education and the Association demonstrated leadership as they joined to create the pilot. The Design Team has provided creative leadership in advancing the pilot through uncharted pathways. As the effort moves forward to institutionalize the critical elements of the pilot into district practice, quality leadership will be essential to shape and guide the reform through its next steps.

Recommendations include:

- *Broaden the collaboration on behalf of student achievement.* The pilot is the result of an unprecedented collaboration between the Board of Education and the Association. This collaboration has been substantive and effective. It should be extended to other parts of district educational operations, regardless of the outcome of the Association and Board votes on a new compensation plan.
- *Continue to place problems on center stage.* A central factor contributing to the accomplishments of the pilot has been the ability to place critical issues that affect the district on center stage. The district will benefit by continuing and extending this function.
- *Create a Principals Leadership and Achievement Institute.* All principals need to understand deeply how learning occurs and how it is nourished, measured and supported. They need ongoing, sustained opportunities to identify salient site issues, analyze trends in student achievement data, reflect on emerging issues, develop their skill in observing classrooms and providing support to teachers, and build the knowledge to work effectively with diverse students and families. Building these capacities will complement the current district plans to prepare principals to carry out targeted educational initiatives.

- *Prepare for the post-pilot and post-vote transition.* The pilot benefited greatly from having a special internal implementation team with the commitment and sense of urgency that is essential to create change. As the learnings and practices from the pilot are implemented district-wide, it will be essential to institutionalize the qualities that the Design Team brought to the implementation of the pilot.

C. Summary

A major initiative that focuses on student achievement—while concurrently exploring changes in the teacher compensation system—goes to the heart of the district mission and structure. As such, the Pay for Performance pilot and study were significant undertakings.

Denver introduced Pay for Performance as a new element in a large urban district. The pilot has been a catalyst for changing the district so that it could become focused on student achievement in a more coordinated and consolidated way. A key part of Denver's story is how a pilot, with key internal and external supporters, engendered positive change in a larger institution. Many of the changes have been systemic—changing how the system thinks and behaves. They are, though, works in progress. Challenges of organizational alignment remain distinct.

As in many large urban districts, Denver experienced leadership transitions over the four years of the pilot. However, the Board of Education and the Association stayed the course. As a result, the pilot achieved a substantial degree of reach into the system. By so doing, Denver has contributed to its own systemic improvement efforts as well as to those of other districts who may want to go down this path.

The issue of aligning a district in support of a pay for performance system cuts to the very essence of how—and to what extent—a school district is functioning in support of student learning. The changes required to identify, strengthen and reward individual student growth and individual teacher contributions under pay for performance have the added effect of stimulating other parts of the school system to improve the quality of support and service. The result is a *catalyst for change* that benefits all students and teachers.