Educational Policy

How Instructional Coaches Support Data-Driven Decision Making: Policy Implementation and Effects in Florida Middle Schools

Julie A. Marsh, Jennifer Sloan McCombs and Francisco Martorell

How Instructional Coaches Support Data-Driven Decision Making

Policy Implementation and Effects in Florida Middle Schools

Julie A. Marsh
Jennifer Sloan McCombs
Francisco Martorell
RAND Corporation

This article examines the convergence of two popular school improvement policies: instructional coaching and data-driven decision making (DDDM). Drawing on a mixed methods study of a statewide reading coach program in Florida middle schools, the article examines how coaches support DDDM and how this support relates to student and teacher outcomes. Authors find that although the majority of coaches spent time helping teachers analyze student data to guide instruction, data support was one among many coach activities. Estimates from models indicate that data analysis support, nevertheless, has a significant association with both perceived improvements in teaching and higher student achievement.

Keywords: *Data-driven decision making; instructional coach*

In recent years, the education community has witnessed increased interest in data-driven decision making (DDDM)—making it a mantra of educators from the central office, to the school, to the classroom. Federal and state accountability policies; improved access to information technology; and a growing pool of commercial products, such as benchmark tests aligned with states' accountability tests, have facilitated the growth of available data for decision making. Some even suggest that educators are "drowning" in too much data (Celio & Harvey, 2005; Ingram, Louis, & Schroeder, 2004). Recent research suggests that although educators appreciate having access to various types of data (Hamilton et al., 2007; Kerr, Marsh, Ikemoto, Darilek, & Barney, 2006), they do not always know how

872

to use the information effectively (Choppin, 2002; Feldman & Tung, 2001; Marsh, Pane, & Hamilton, 2006; Mason, 2002; Supovitz & Klein, 2003).

One promising approach to providing school educators better guidance on using data to inform practice is the use of coaches—master teachers who offer on-site and ongoing instructional support for teachers. In fact, like DDDM, federal policies such as the Reading First, Striving Readers, and the No Child Left Behind Act (NCLB) have encouraged the expansion of coaching across the country. Numerous schools, districts, states, and school reform models (e.g., Accelerated Schools and America's Choice) currently employ coaching as a primary part of their improvement programs (Foltos, 2007; Galm & Perry, 2004; Russo, 2004). Moreover, coaching increasingly has become a centerpiece of literacy reform policies in many schools and districts and a few states.

Yet despite the widespread use of coaches and DDDM, there is little research examining how coaches support DDDM in schools and the extent to which these efforts are associated with improvements in teaching and student achievement. This article begins to fill this research gap by presenting empirical results from a mixed-methods study of a statewide reading coach program in Florida middle schools. In particular, this article investigates three broad research questions:

Research question 1: To what extent are reading coaches focusing their work on data analysis and support?

Research question 2: What kinds of support do coaches receive that promote their data-support activities?

Research question 3: To what extent is variation in coaches' time spent on data analysis related to teacher and student outcomes?

Answers to these questions will contribute to policy and practice and several important ways. First, given the significant federal, state, and local resources allocated to coaching programs and to generating data (e.g., interim assessment systems, state testing programs), it behooves policy makers to better understand if and how coaches support the effective use of data and whether these investments result in better outcomes. Second, should coaching prove to be an effective means of facilitating data use and improved outcomes, administrators, coaches, and teachers would benefit from information about what constitutes and enables effective coach practice in this area.

In the remainder of this article we first provide context for the study, including a review of literature and a description of the Florida coaching

program. Next we describe the study's conceptual framework and methods. We then present answers to the research questions, including findings on how coaches performed their data support roles and interacted with teachers, how districts helped develop coaches' DDDM knowledge and skills, and how frequency of data support relates to teacher and student outcomes. We conclude with a set of implications for policy, practice, and research.

Background

Two sets of literature inform our research and are important for understanding the convergence of coaching and DDDM policies.

DDDM

DDDM in education refers to teachers, principals, and administrators systematically collecting and analyzing various types of data, including input, process, outcome, and satisfaction data, to guide a range of decisions to help improve the success of students and schools (Marsh et al., 2006). Notions of DDDM in education are modeled on successful practices from industry and manufacturing, such as total quality management (TQM), organizational learning, and continuous improvement. These practices emphasize that organizational improvement is enhanced by responsiveness to various types of data, including input data (such as material costs), process data (such as production rates), outcome data (such as defect rates), and satisfaction data (including employee and customer opinions; see, for example, Deming, 1986; Juran, 1988; Senge, 1990).

The broad implementation of standards-based accountability under the federal NCLB act has presented new opportunities and incentives for data use in education by providing schools and districts with additional data for analysis, as well as increasing the pressure on them to improve student test scores (Massell, 2001). NCLB required states to adopt test-based accountability systems that meet certain criteria with respect to grades and subjects tested, the reporting of test results in aggregated and disaggregated forms, and school and district accountability for the improvement of student performance. Implicit in NCLB and other state accountability policies is a belief that data—particularly student test results—are important sources of information to guide instructional decisions. New state and local test results are adding to the data on student performance that teachers regularly collect via classroom assessments, observations, and assignments.

To date, most of the research on this topic has examined the implementation of DDDM, but only a limited amount of research has tried to assess the effects of DDDM (Marsh et al., 2006). For example, although they do not systematically track outcomes, a few case studies offer anecdotal evidence of increased test scores and improved student learning in schools that effectively implemented DDDM (Feldman & Tung, 2001; Schmoker & Wilson, 1995). In addition, the effective schools literature includes several studies that identify "planful" and extensive use of data as a common characteristic among schools and districts that are high performing (Casserly, 2002; Council of Great City Schools, 2002; Edmonds, 1979; Ragland, Clubine, Constable, & Smith, 2002; Snipes, Doolittle, & Herlihy, 2002; Supovitz & Klein, 2003; Symonds, 2003; Williams, Kirst, & Haertel, 2005). Finally, numerous studies link DDDM to changes in school culture and teacher practice that past research has linked to improved student performance. Common findings include teacher reports of greater differentiation of instruction, greater collaboration among school faculties, and improved identification of students' learning needs as a result of increased data use (Chen, Heritage, & Lee, 2005; Copland, 2003; Feldman & Tung, 2001; Wayman & Stringfield, 2006).

Despite the dearth of research on effects, many studies have identified a set of factors that are associated with more effective use of data by educators (for a review, see Marsh et al., 2006). Notably, several studies identify the importance of providing training to educators on how to use data and connect them to practice (Black & William, 1998; Datnow, Park, & Wohlstetter, 2007; Mason, 2002; Supovitz & Klein, 2003). Such training tends to address skills such as formulating research questions, interpreting results, and effectively developing and using classroom assessments, and often provide educators opportunities to discuss data and use their own real-life data issues and school challenges rather than hypothetical cases (Chen et al., 2005; Copland, 2003; Herman & Gribbons, 2001; Love, 2004; Mason, 2002; Murnane, Sharkey, & Boudett, 2005).

Coaching

Instructional coaching is one potential avenue for providing teachers with professional development on DDDM. Current coaching programs generally involve on-site specialists who work with classroom teachers to improve instruction in a particular content area, most often literacy (Knight, 2006). Some programs utilize coaches to support the implementation of particular instructional models or curricula, whereas others work to improve

general instructional practices. In nearly all models, instructional coaching is school based, collaborative, and conducted one on one or in small groups. Unlike other staff who support reading (e.g., reading resource teachers), coaches generally serve in a nonevaluative, support role for teachers and do not directly instruct or tutor students unless used as a means to model instruction for teachers.

Advocates and researchers often point to learning theory and research on professional development as the rationale for coaching. Learning theory suggests that individuals learn best when provided with opportunities to discuss and reflect with others, to practice application of new ideas and receive feedback from an expert, and to observe modeling (Brown, Collins, & Dugrid, 1989; Collins, Brown, & Holum, 1991; Lave, 1988; Lave & Wenger, 1991; Rogoff, 1990; Tharp & Gallimore, 1988; Vaughan, 1996). Empirical research further suggest that the transfer of ideas from the traditional professional development model of one-shot workshops into actual instructional change and increases in student learning is extremely limited (e.g., Garet et al., 1999; Garet, Porter, Desimone, Birman, & Yoon, 2001; Hawley & Valli, 1999; Showers & Joyce, 1996).

Though reading coaches are prevalent in many schools across the nation, there is little empirical evidence regarding the nature of coaching and its effectiveness in changing teacher practice and improving student achievement. Much of the current research on coaching focuses on Reading First coaches at the elementary level (e.g., Deussen, Coskie, Robinson, & Autio, 2007; Wong & Nicotera, 2006), with a few emerging studies at the high school level (e.g., Brown et al., 2007, 2006). Research on the effects of coaching on practice suggests that the policy shows some promise. For example, coaching pioneers Joyce and Showers (1996) found in several studies that teachers in coaching relationships practiced new skills more frequently and applied them more appropriately in their classrooms than other teachers. Kohler et al. (1999) reviewed the coaching literature from the 1980s and 1990s and reported several positive outcomes, including improvements in teachers' ability to plan and organize, provide instruction for students with disabilities, use classroom behavior management strategies, and address instructional objectives. More recently, authors have reported improvements in school culture and teacher collegiality and collaboration related to coaching programs (Guinney, 2001; Neufeld & Roper, 2003; Richards, 2003); positive changes regarding teachers' awareness, comprehension, alignment, and implementation of state standards (Wong & Nicotera, 2006); and knowledge of how to apply professional development learning in their classrooms (Brown et al., 2006). Two small-scale observational

studies by Kohler and colleagues (1997, 1995) found that teachers were more likely to implement changes in instruction while being coached than while working independently and that the changes made during the coaching phase were sustained after coaching ended. However, some studies have found coaching to have little to no impact on teacher behavior (Gutierrez et al., 2001; Veenman, Denessen, Gerrits, & Kenter, 2001).

Although some literature is suggestive of promising outcomes related to coaching models, it is limited in several ways. Much of the work done in the 1980s and 1990s focused on peer coaching rather than on content coaching, and so may not be completely relevant for current models of coaching that use designated on-site staff developers rather than peers in the coaching role (e.g., Kohler et al., 1999; Munro & Elliot, 1987; Showers & Joyce, 1996). Many studies from this era also focus on coaching for preservice rather than in-service teachers (e.g., Hasbrouk, 1997; Morgan, Menlove, Salzberg, & Hudson, 1994). Moreover, many studies have relied on anecdotal evidence or self-reported data and provided minimal explanation of methodology. Perhaps the largest gap in the existing research on coaching programs is the lack of evidence of coaching programs' effects on student achievement.

Florida's Coaching Program

With this literature in mind we designed a study to examine coaching at the secondary level and to provide empirical evidence on the effects of coaching for teachers and students. Florida provided a unique opportunity to study reading coaching situated within a broader state-led literacy policy, the "Just Read, Florida!" initiative. Established in September 2001 by then Governor Jeb Bush, the initiative's goal is that all students read at or above grade level by 2012. One key component of this effort has been the allocation of funds to districts to hire full-time, site-based reading coaches. Florida has been scaling up its reading coach initiative since 2002. In a span of 5 years, the number of participating schools increased from 300 in 30 districts to more than 2,200 in 72 districts in 2006-2007.[1] In 2006-2007, the year of this study, the state estimates that 2,360 coaches were funded through local, state, and federal funds: 1,413 served in elementary schools, 526 in middle schools, and 421 in high schools. Of these coaches, 1,977 served full-time at one school; 270 served part-time in one school; and 113 served full-time, splitting their time between two schools.

Florida's reading coach program does not provide a specific model per se[2] but instead provides an array of conceptual, policy, and practical supports

that are intended to guide the work of a coach. The overarching goal of Florida's coaching program is to improve students' reading ability by helping teachers implement effective, research-based instruction in reading and in content areas. State definitions of the reading coach emphasize that a middle school reading coach is an on-site person who

> will serve as a stable resource for professional development, progress monitoring, and student data analysis throughout a school to generate improvement in reading instruction and reading achievement. The middle school reading coach will both support and provide initial and ongoing professional development to teachers in each of the major reading components, administration and interpretation of instructional assessments, and differentiated instruction (Florida Department of Education, 2004).

Aside from the requirement that coaches be full-time employees, the state does not mandate any other aspects of a coach's job but instead provides districts with a basic job description suggesting basic coach qualifications (e.g., experience teaching, knowledge of reading research and of how to work with adult learners—in this case, teachers) and ways in which the coach should operate at the school level. Although the state job description explains that coaches will "train teachers in data analysis and using data to differentiate instruction," it is one of 10 suggested responsibilities. Furthermore, the state does not provide details on the percentage or amount of time coaches should spend on these data-related tasks. The state also encourages coaches to work with all teachers across content areas, with a focus on new teachers, new reading teachers, and those teaching struggling students; to prioritize their time on in-class coaching (e.g., modeling, mentoring, observing, providing feedback); and to avoid formally evaluating teachers and participating in activities that detract from work with teachers (e.g., administrative tasks, too much time administering assessments, tutoring students, substitute teaching). To encourage fidelity to the state's vision for coaching, the state provides annual training to coaches and principals. It also requires coaches to submit biweekly coach logs accounting for time spent and districts to submit reading plans that detail how coaches will be supported and utilized, both of which are monitored by the state.

Conceptual Framework

Our study design, data collection, and analysis were guided by a conceptual framework grounded in the empirical and theoretical research on

Figure 1
Conceptual Framework for the Study
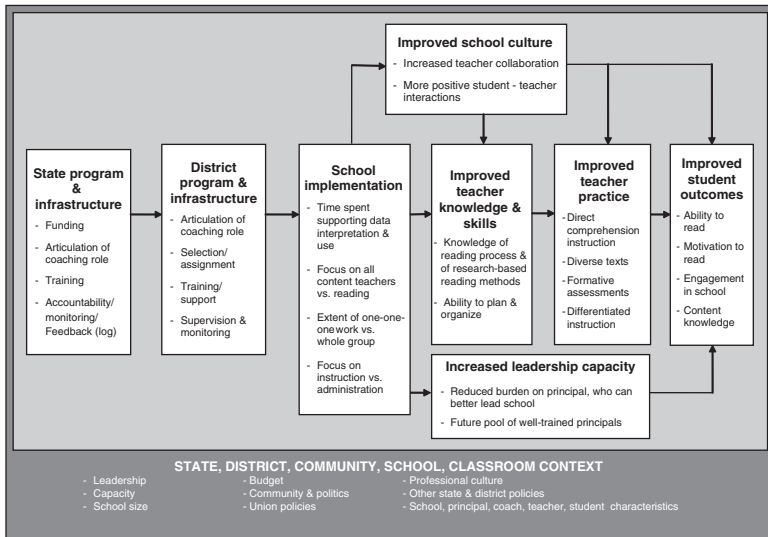
coaching, learning, and DDDM, as well as the state's implicit "theory of action" we deduced from our interviews and review of documents (Figure 1). The basic hypothesis is that increasing the expertise and availability of reading coaches to work with teachers at a school site will allow teachers to gain new knowledge and skills or enhance existing knowledge and skills, which in turn will improve their reading instruction and ultimately improve student achievement and other outcomes. The model recognizes that the state and district shape this process by articulating the roles and responsibilities of the coach, setting hiring qualifications, providing ongoing training and support to reading coaches, and monitoring their efforts. Schools also influence the coaching process by directing coaches' attention to certain priorities. Other aspects of a coach's actual work at the school level may also influence his or her effects on teachers, such as the amount of time spent working with teachers to support data interpretation and use. Theoretical notions of DDDM and organizational improvement (cited earlier) indicate that when properly examined, interpreted and acted upon, certain types of data can assist in improving individual practice and organizational

outcomes. Learning theory also suggests that the quality of coach–teacher interactions (e.g., how information is introduced, new practices are modeled, and teachers are provided opportunities for application and reflection) are likely to influence instructional responses.

The framework also posits that coaching can affect student learning through various other intermediate outcomes, such as building school leadership capacity and enhancing school climate, which in turn might either directly affect student achievement or indirectly affect achievement through changes in teacher practice. Finally, the framework for the study recognizes that Florida's coaching program, like all coaching programs, is embedded in a broader state, district, and local context that can influence coaching practice and its impact, and includes such factors as principal leadership, school size, and other state and district policies. Characteristics of participants may also affect the implementation and effects, most notably characteristics of each coach, such as their knowledge, skills, and confidence, as well as their experience teaching reading and coaching.

Data and Methods

As noted, the study from which this article draws used a combination of quantitative and qualitative methods (see Marsh et al., 2008 for further details).

Sample

To examine coaching implementation, we selected a purposive sample of large districts in Florida that represent a range of approaches to and experience with middle school coaching.[4] We drew an initial sample of 9 districts from among the largest 12 districts in the state (with approximately 10-45 middle schools each). One district declined to participate, giving us a final sample of eight districts.

In each of the eight study districts, we randomly sampled schools from all regular and charter middle schools (defined as schools serving Grades 6 through 8) that employed a part-time or full-time reading coach in 2006-2007, representing a total population of 226 schools.[5] In each district, we drew a random sample of schools proportional to district size, for a total sample of 180 schools. Ultimately, we recruited 113 schools to participate, representing an overall cooperation rate of 63%.

In each school, we surveyed the principal, all reading coaches, and 10 teachers. As a general rule, from each school roster we randomly sampled five reading teachers and five social studies teachers, stratified by grade, to obtain a representative sample that would be adequate for our analyses. (Florida requires students performing below proficiency on the state reading assessment to take a reading course in the middle grades from reading teachers.) We selected reading teachers because state interviews and a review of documents indicated that coaches were likely focusing much of their attention on these teachers. We selected social studies teachers to capture the perspectives of core-content area teachers who we were told were likely to interact with the coach.

From the eight participating districts, we selected two districts from which to collect more in-depth qualitative data and in which we were able to pretest our survey instruments. The districts selected differed both in size (the larger district oversaw approximately twice as many middle schools as the smaller one) and in their support for coaching (e.g., one provided small and large group professional development to coaches several times a month for extended full and half days, whereas the other provided training less frequently). Within each district we selected three schools to follow over the course of the year and within each school we selected the coach and three teachers with whom the coach had been working closely or planned to work with over the course of the year to follow.

Data Collection

Surveys. In Spring 2007, we administered Web-based surveys to principals, reading coaches, and reading and social studies teachers in our sample of schools from the eight participating districts. The four survey instruments drew on our conceptual framework, existing literature on coaching, data collected in the first round of case study visits to schools and districts, measures validated from other studies, and careful review by experts. We also pilot tested the draft surveys with teachers, coaches, and principals, who provided us feedback about the clarity of the items. Table 1 shows the response rates for each respondent group.[6] To adjust for potential differences due to differential sampling and nonresponse, we created weights that reflected both the known sampling probabilities and estimated response probabilities at the school and teacher level so that our responding sample would be representative of the entire population of middle schools in the eight study districts. We used these weighted data in our cross-tabulations. Achievement models were run with unweighted data.

Table 1
Survey Response

	Number Sampled	Number Ineligible[a]	Number Responding	Response Rate (%)
Principals	113	0	96	85
Reading coaches	124[b]	0	109	88
Reading teachers	554	1	386	70
Social studies teachers	563	3	348	62

a. Ineligible individuals were teachers originally identified on rosters as teaching reading or social studies who, after receiving the survey, told us they were either no longer teaching at the school or not teaching that particular subject.
b. Because some schools had two full-time coaches, the number of coaches is greater than the number of schools and principals.

Case study visits. Researchers visited each case study school three times during the 2006-2007 school year. During the visits, researchers conducted interviews with the reading coach, principal, and three case study teachers at each school; observed one period of instruction of each case study teachers; shadowed the reading coach for at least half a day; and conducted focus groups with core-content-area teachers. Researchers also collected documents pertaining to reading coaches and reading improvement efforts. We conducted a total of 64 interviews, 13 focus groups (with 43 teachers in total), and 28 observations over the course of the academic year.

District and state interviews, observations, and documents. Interviews with state-level staff, attendance at the state's annual leadership conference, and documents provided us with information on Florida's coaching program and supports. Telephone interviews with the supervisors of middle school reading coaches in the six non–case study districts also provided information on district support for coaching.

Student achievement and demographic data. We obtained from the state department of education's K-12 Data Warehouse FCAT (Florida's Comprehensive Assessment Test) score information (criterion-referenced portions for reading and math) for individual students in all schools in the state that included any of Grades 6 to 8, from 2001-2002,[7] the school year prior to the first year of implementing the state's middle school reading coach initiative, through 2006-2007. We also obtained background information for individual students, including gender, ethnicity, socioeconomic status, limited English proficiency (LEP) status, participation in special

education or gifted programs, attendance, mobility, age, and grade retention history.

Descriptive Analysis

The answers to the first two descriptive research questions are based on weighted survey responses, as well as case study data. We integrated findings from the different data sources to identify cross-district findings and themes regarding the nature, quality, perceived impact, and potential barriers and enablers of coaching. We also examined how coaches' work varied by such factors as coaches' experience and school characteristics, using simple cross-tabulations of data and Pearson's chi-squared test to determine whether these relationships were statistically significant. Throughout the article we use the term *significantly* (e.g., "significantly more likely") to indicate relationships that were found to be statistically significant at $p < .05$.

Analysis of Effects: Modeling Approach

We address the third research question by estimating a set of regression models.

Perceived effects. To more precisely examine the relationship between coach activities and perceived effects on instruction when other factors are held constant, we employed least squares regression analyses to model various predictors of perceived influence on teacher practice. The variables used in the models are defined in Table 2 and are derived from the conceptual framework for the study. The outcome variable modeled is teachers' perception of coach influence on their practice. As we could use only the 86 schools that provided survey responses from both the principal and the coach, we selected a fairly parsimonious set of program features identified as important in prior research and in our own survey findings. We focus on indicators of coach skill, knowledge, and ability, including their reading credential status, experience teaching reading, ability with adult learners, whether or not they are a more experienced coach, their confidence or self-efficacy, and teachers' perceptions of coaches' overall quality; school contextual factors that may enable or hinder coaches' work, including coach caseload; the percentage of new teachers in the school; the number of years the school had a coach; and principal leadership. We also include measures of coaches' activities from teacher surveys and reports on the focus of coaches' work from coach surveys.

Table 2
Definition of Variables Used in Models

Constructs	Definitions
Outcome	
Perceived influence on teacher practice *M* = 2.3 *SD* = 0.6 Source: Teacher surveys	To what extent did the coach influence any changes made to your instruction over the course of the year? Measured on a 4-point scale (1 = *not at all*, 2 = *to a small extent*, 3 = *to a moderate extent*, 4 = *to a great extent*)
Predictors	
Coach expertise and experience	
Reading credential *M* = 0.79 *SD* = 0.39 Source: Coach survey	Defined as coach having a master's degree in reading, a reading certification, state reading endorsement, or combined state reading/English to speakers of other languages (ESOL) endorsement.
Years teaching reading *M* = 11.2 *SD* = 9.67 Source: Coach survey	Defined as total years experience teaching reading and serving as a reading specialist or reading resource teacher.
Perceived coach quality scale (alpha = .91) *M* = 3.03 *SD* = 0.41 Source: Teacher surveys	The reading/literacy coach(es) at my school • has strong knowledge of best practices in reading instruction • *has a limited understanding of the particular needs of students that I teach* • has a strong understanding of my needs as a teacher • helps me adapt my teaching practices according to analysis of student achievement data (e.g., test results) • maintains confidentiality of what we discuss or work on together • understands the middle school culture and student • *has little time to regularly support teachers* • is someone I trust to help me and provide support • provides feedback in a nonevaluative way • explains the research, theory, or reasons underpinning the strategies (s)/he suggests or the feedback (s)/he provides • (social studies teacher only) *does not have sufficient understanding of my content area to help me with my teaching*. Measured on a 4-point scale with an additional "don't know/NA" option (1 = *strongly disagree*, 2 = *disagree*, 3 = *agree*, 4 = *strongly agree*). Statements in italics were reverse coded.

(continued)

Table 2 (continued)

Constructs	Definitions
Ability to support adult learners *M* = 2.50 *SD* = 0.61 Source: Principal survey	How would you rate your reading/literacy coach's knowledge and skills in the following area? If your school has more than one reading coach, answer the question for the reading coaches as a team. Understanding of how to support adult learners Measured on a 3-point scale (1 = *weak*, 2 = *medium*, 3 = *strong*)
Coach confidence scale (alpha = .59) *M* = 3.54 *SD* = 0.45 Source: Coach survey	To what extent do you agree or disagree with the following statements about your work as a reading/literacy coach? • I feel confident in my ability to support teachers with reading instruction. • *I do not feel prepared to help content-area teachers incorporate reading strategies into their classrooms* (reverse coded). Measured on a 4-point scale (1 = *strongly disagree*, 2 = *disagree*, 3 = *agree*, 4 = *strongly agree*)
More experienced coach *M* = 0.49 *SD* = 0.49 Source: Coach survey	Defined as having been a coach for three or more years (yes/no)
Coach activities	
Focus on integrating instruction across content areas *M* = 3.15 *SD* = 0.75 Source: Coach survey	Considering all of the work you have done with teachers this school year, how much emphasis did you place on supporting the following area of instruction? • Integrating reading instruction across the content areas Measured on a 4-point scale (1 = *no emphasis*, 2 = *minor emphasis*, 3 = *moderate emphasis*, 4 = *major emphasis*)
Reviewed assessment data with coach Reading Teachers (RT), *M* = 2.14 RT, *SD* = 0.59	How often has your school's reading/literacy coach(es) performed the following actions? Since the beginning of the school year, my school's reading/literacy coach(es) has reviewed student assessment data with me (individually or in a group).
Social Studies Teachers (SST) test, *M* = 1.65 SST test, *SD* = 0.50 Source: Teacher surveys	Measured on a 4-point scale (1 = *never*, 2 = *a few times this year*, 3 = *once or twice a month*, 4 = *once or twice a week or more*)
Received individual coaching scale (alpha = .88)	How often has your school's reading/literacy coach(es) performed the following actions?

(continued)

<div align="center">

Table 2 (continued)

</div>

Constructs	Definitions
RT, M = 1.88 RT, SD = 0.51 SST, M = 1.49 SST, SD = 0.42 Source: Teacher surveys	Since the beginning of the school year, my school's reading/literacy coach(es) has • come to my classroom to coteach or model a lesson or reading strategy, • assisted me with planning a lesson or curricular unit, • visited my classroom to observe my instruction, • given me feedback on my teaching or facilitated reflection on my practice. Measured on a 4-point scale (1 = *never*, 2 = *a few times this year*, 3 = *once or twice a month*, 4 = *once or twice a week or more*)
Context for coaching	
Number of years the school had a coach M = 4.01 SD = 2.18 Source: Principal survey	For how many years (including this year as one) has your school had a reading/literacy coach?
Coach caseload M = 6.97 SD = 0.50 Sources: Principal survey and common core of data	Log (number students per coach) Ideally we would have used a measure of teacher-to-coach ratio; however, we did not have reliable data to construct such a variable. Given that the ratio of students to teachers generally does not vary considerably across schools, the student-to-coach ratio is a useful proxy.
Percentage of new teachers in the school M = 27.36 SD = 16.02 Source: Principal surveys	A new teacher is defined as someone teaching less than 3 years.
Principal leadership scale (alpha = .94) M = 3.14 SD = 0.36 Source: Teacher surveys	The head principal at my school • communicates a clear academic vision for my school, • sets high standards for teaching, • encourages teachers to review the Sunshine State Standards and incorporate them into our teaching, • helps teachers adapt our curriculum based on an analysis of FCAT (Florida's Comprehensive Assessment Test) results, • expects all staff to work with the reading coach to reflect on and improve their teaching, • ensures that teachers have sufficient time for professional development, • enforces school rules for student conduct and backs me up when needed, • makes the school run smoothly, • is someone I trust at his/her word. Measured on a 4-point scale (1 = *strongly disagree*, 2 = *disagree*, 3 = *agree*, 4 = *strongly agree*)

Achievement. To understand associations between coaching implementation and achievement, we estimated school random-effects models using our survey data and student-level achievement data obtained through the Florida Department of Education. In these models, student achievement in reading and mathematics on the 2007 FCAT Sunshine State Standards (SSS) test were modeled as a function of coaching program features during the 2006-2007 school year (the same measures described in Table 2) and other school and student characteristics. Specifically, we estimated models of the form Y_{is} = $X_{is}b + a_s + e_{is}$, where Y_{is} is the dependent variable (e.g., test scores) of student i in school s, X_{is} is a vector of covariates (including the reading coach program measures), a_s is a school random effect, and e_{is} is an individual-level random term. The parameter "b" measures the influence of each of the covariates in X_{is}. We assume that a_s and e_{is} are normally distributed i.i.d. (independent and identically distributed) random variables that are statistically independent from each other and also that a_s and e_{is} are both uncorrelated with the covariates in X_{is}. Under these assumptions random-effects regression will provide consistent estimates of "b."

We model mathematics outcomes because improved reading skills may also improve students' scores on the mathematics FCAT, which is a text-heavy assessment that includes many word problems as well as performance tasks in which students must solve a problem and explain their methods used. Achievement scores were modeled as a *z* score with a mean of 0 and a standard deviation of 1. For this analysis, all coaching implementation measures were aggregated to the school level because we do not have any way of linking individual students to teachers who worked with the coach.

For consistent estimates of the effects of various aspects of coaching implementation to be obtained, omitted influences on student achievement must be unrelated to coaching implementation variables. To better understand the relationship between the coach activities and student outcomes, we control for school and student characteristics that might be associated with both the coaching program and student achievement, including the percentage of new teachers, the number of coaches a school has, the years of experience the coaches have, and teachers' perception of the principal's leadership. Our models also control for student characteristics including gender, ethnicity (Hispanic, African American, Other), limited English proficiency, special education, percentage of school days attended, free-lunch eligibility, reduced-price lunch eligibility, grade retention, and grade level. We also control for school-level covariates, including the number of students enrolled in the school, the percentage of students eligible for free or reduced lunch, and the percentage of minority students in the school.

In addition, we control for student prior achievement by including fourth-grade test scores in the model. Thus, the estimates we obtained reflect the association between aspects of the coaching program and achievement relative to students' baseline performance observed in fourth grade. Note that the fourth-grade test scores provide a good proxy for "pre-coaching" baseline achievement because in fourth grade most students were in a different school that either had no reading coach or a totally different coaching program. Thus, even if the nature of the coaching program differs by whether a school has higher or lower achieving students, our estimates will still be unbiased so long as the coaching program is unrelated to the potential gains between fourth grade and middle school.

The number of schools in our analysis is 86 schools. The number of student observations in our achievement models is 71,234.[8]

Study Limitations

Given resource constraints, we were able to examine coaching implementation in only eight moderate-to-large districts. Although this approach does not allow us to generalize to all districts in Florida, particularly smaller districts, it does allow us to describe how different models of district implementation are translated into school- and coach-level practices and classroom practice. Second, due to limited resources, we could survey only 10 teachers across 2 content areas in each participating school. Clearly a sample of all teachers would have provided more reliable estimates of coach interactions with and perceived influences on teachers throughout a school. Thus, it is possible that our responding teachers do not accurately represent the experiences of all teachers in a school, particularly content-area teachers. Third, our measures of teacher practice were limited by the reliance on self-reported practices from surveys and interviews (despite attempts, we were unable to conduct enough meaningful classroom observations to use in an analysis of changes in teacher instruction). Although prior research suggests that well-designed surveys can measure some aspects of instructional practice with a reasonable degree of accuracy (Mayer, 1999; Mullens & Gayler, 1999; Smithson & Porter, 1994), these measures tend not to be as rich or nuanced as those collected through first-hand observations. Nevertheless, our in-depth case studies that included observations helped mitigate this limitation. Finally, this is a cross-sectional analysis that examines the effects of coaching on student achievement in 1 year only. The full impact of coaching may be better measured over time, particularly if the full effects of coaching are lagged or grow over time.

Descriptive Findings

In this section we answer the first two research questions, describing the extent to which reading coaches focused their work on data analysis and the ways in which districts supported coaches to perform this data-support role.

The Majority of Coaches Reported a Major Focus on Analyzing Data to Guide Practice

Congruent with the state's vision, the majority of coaches focused considerable attention on DDDM. When asked to consider all of the work they did with teachers during the 2006-2007 school year, 62% of coaches reported placing a major emphasis on supporting the analysis of data to guide instructional practice. The results also indicate coaches focused on other key components of reading instruction endorsed by the state and others as relevant to middle school students: More than half placed a major emphasis on supporting comprehension, vocabulary, fluency, and differentiating instruction to meet student needs. They were considerably less likely to focus support on other instructional areas, such as classroom management, phonemic awareness, and writing.

In some case study schools, coaches' DDDM work included presenting FCAT results to teachers in schoolwide or department meetings at the start of the year to identify student weaknesses. In other schools, the coach may have worked individually with teachers to help them understand diagnostic assessment results for their students and how to use them to identify on-level reading material or appropriate instructional strategies. The coach described in the following vignette represents someone who strongly embraced the role of coach as data analyst:

Elaine, a seasoned reading coach in her second year at the school, used data to drive much of her work throughout the year. At the start of the year, Elaine examined schoolwide FCAT test results with members of the reading leadership team to develop a "reading treatment plan." The plan identified several problem areas and strategies to address them, such as working on vocabulary development, establishing a reading workshop model in several classrooms, and setting up model "lab" classrooms. Periodically, Elaine also analyzed schoolwide data to evaluate various programs. For example, she examined assessment results from a reading software program that several teachers were using and reported back to school administrators on students' strengths and weaknesses and some potential problems with the program. This analysis

indicated that most students were not at the "standard productivity level" and may not have been taking the program seriously and that teachers may not have understood how to effectively use the program. One administrator noted how valuable this analysis has been for school leaders:

> [S]he did a report for [this] reading program . . . [which] helped me out because she pinpointed what the problems were and what the students' strengths were. She really understands those reports. She is really able to look at the data and really assess where we should be. . . . She diagnosed a problem; she took a big write-up with some solutions of things we could do or things we should do to assist with those problems and she shared that information with myself [and other administrators] and from there we met as a reading leadership development committee team and we talked about some of those things and . . . how we can integrate it into doing some other things . . . and try to make some improvements.

Throughout the year, Elaine also helped the reading teachers review FCAT and other assessments results and understand how to use them to identify areas that needed more instructional attention. For example, a sixth-grade reading teacher reported that the coach not only modeled how to administer the oral reading fluency and comprehension tests but also discussed the results with teachers:

> She is good about making the spreadsheet and showing us what they [students] have done and what they haven't done. And she breaks it down into different levels like comprehension and whatever else . . . that deals with the strategies on the FCAT so that we will know what we need to work on as a group and what we need to pull for small groups.

Another teacher noted that "before [Elaine] came you were basically on your own" to understand assessment results and how to adjust your teaching to address them. She explained, "We knew there was a problem but we did not know what to do. . . . She helped us to understand that we were presenting a sixth-grade-level assignment to a sixth-grade student who was reading on a third-grade level."

Data Support Was One of Many
Activities to Which Coaches Devoted Time

Coaches in all districts reported dividing their time among a wide range of activities (Table 3). Half of all coaches spent 6 or more hour every 2 weeks analyzing and training teachers on how to analyze and use student data to inform instruction (including FCAT, MAZE, Fluency checks, student

Table 3
Time Coaches Spend on Activities During a Typical 2-Week Period

	5 Hr or Less	6 to 16 Hr	17 to 24 Hr	More Than 24 Hr
Working with individual teachers one on one on their instruction (including classroom observations)	19	42	23	15
Providing a "listening ear" for teachers' concerns	25	39	25	11
Administering or coordinating student assessments (including managing assessment materials)	37	35	16	12
Analyzing and training teachers on how to analyze and use student data to inform instruction (including FCAT (Florida's Comprehensive Assessment Test), MAZE, Fluency checks, student work)	50	28	14	8
Managing reading resources and materials (including ordering, budgeting, doing inventory, locating written materials as well as overseeing computer software and reading labs)	53	26	17	4
Working with groups of teachers on their instruction (including large group professional development sessions)	60	27	9	3
Attending meetings or professional development sessions (not ones that you lead) in the school, district, or region	52	38	8	2
Performing noncoaching administrative duties (including lunch duty, bus duty)	75	18	6	2
Teaching or tutoring students in class or in computer labs	75	16	4	5
Substitute teaching	92	5	2	2

Note: Response options were "I generally do not do this every 2 weeks," "a small amount (1-5 hr)," "a moderate amount (6-16 hr)," "a large amount (17-24 hr)," "a very large amount (more than 24 hr)." Several of these categories were collapsed above to obtain the columns/categories in the table.

work).[9] A subset of coaches (22%) devoted a significant amount of time to data analysis—spending at least a quarter of their time over a 2-week period on this activity. Nevertheless, the average reported times were greater for several other categories of activities. First, individual instructional work with teachers topped the list of coaches' activities, which would include observing instruction, providing feedback on instruction, and modeling.[10] A little more than one third of coaches reported spending 17 hour or more in the 2-week window working one on one with teachers. The second most frequent activity reported by coaches could be categorized as informal coaching or lending a "listening ear" for teachers' concerns.

Third, coaches were more likely to spend time administering and coordinating assessments than actually analyzing or helping teachers analyze and use the results to guide practice. More than two thirds of coaches reported spending 6 hour or more every 2-week period administering or coordinating student assessments, which could include state or local tests. In most case study schools, coaches spent significant amounts of time administering state- and district-required individual reading assessments, such as fluency tests, to some students and inputting results from all teachers into the state's database. In some case study schools, coaches also organized and managed the administration of the state FCAT exam. Teachers in one school identified the FCAT administrative duties as ones that greatly competed with the intended coaching duties. "She gets pulled out of her reading role into an FCAT administrator role, which someone else could be doing," explained one teacher. Interestingly, more than half of coaches cited the large amount of time it takes to coordinate and administer assessments as a moderate or great hindrance to their work.

More Experienced Coaches and Coaches in Low-Performing Schools Were More Likely Than Counterparts to Spend Time Supporting DDDM

More experienced coaches (3 or more years of coaching experience) were significantly more likely than less experienced coaches (1-2 years) to spend a large amount of time (17 or more hour every 2 weeks) on data analysis (32% compared with 12%, respectively). This is perhaps understandable, given the skills and knowledge required to successfully interpret, communicate, and assist teachers with data. Newer coaches may have less time initially to devote to building data analysis capacity, given the more immediate concerns of understanding the roles of the coach, building relationships

with teachers, and gaining access to classrooms. Coaches in low-performing schools (receiving state school grades of C, D, or F) were also much more likely than their counterparts in high-performing schools (receiving A or B grades) to spend a large amount of time analyzing and helping teachers use the results (51% compared with 16%, respectively). This is consistent with patterns of time spent on administering assessments. Given that low-performing schools likely have more students with severe reading problems and more students enrolled in the more "intensive" reading courses requiring assessments, it follows that coaches in these schools may have a lot more achievement data with which to work as well as greater pressures to use these data to improve instruction.

Reading Teachers Were More Likely Than Social Studies Teachers to Receive Coach Data Support

Overall, reading teachers were far more likely than social studies teachers to interact with the reading coach in general, be it attending a meeting in which the coach presented information or working with the coach to plan instruction or locate reading materials. This distinction holds true for data support as well: 29% of reading teachers reported receiving this support from their coach once or twice a month or more compared with 12% of social studies teachers. In fact, half of the social studies teachers reported that their school's reading coach *never* reviewed assessment data with them (individually or in a group) in the past year, compared with only 23% of reading teachers.

Districts Often Helped Develop Coaches' DDDM Knowledge and Skills

Coaches generally characterized central office staff as supportive of their work and valued the professional development opportunities they offered. Furthermore, almost all agreed or strongly agreed that their districts gave them necessary guidance on how to improve reading instruction and performance in their schools (92%). As the state envisioned, all the study districts provided at least monthly professional development opportunities for coaches. Described as either mandatory or strongly encouraged, the sessions were generally well attended by coaches. The focus of these meetings varied across and within districts over time, although most district coordinators described a strong instructional focus and several emphasized training for DDDM. More than half of coaches reported that district professional

development for coaches placed a major emphasis on analyzing and using student data to improve instruction. In fact, most coaches were satisfied with the amount of support received: Only 21% wanted "much more" support in this area and 49% wanted a "little more." The following vignette describes a district meeting that emphasized training for DDDM:

> During the monthly coach meeting in one case study district, district coordinators focused a majority of the day on interpreting and using data to guide instruction and modeling how to assist teachers in this process. First, one of the coordinators provided a PowerPoint presentation on differentiated instruction. Using an example of a "real kid," she modeled how to interpret results from multiple assessments to identify this student's reading needs and possible lessons and interventions that might address those needs, as well as how a coach might work with a teacher to develop and implement these strategies. Throughout the activity, the coordinator was explicit about her intentions; for example, she explained, "My purpose is to show you the curriculum that you could put together with a teacher." As she described how to move "from data to a plan," coaches regularly interjected with questions and suggestions of their own. One coach asked, "How does the first part of the lesson look?"; later, another coach explained the value of doing preassessments with students; still another suggested a particular series of books that might be helpful for teachers. At the end of the discussion, the administrator returned to overhead slides describing differentiated instruction and the criteria for "tiered lessons." She then asked the coaches to use those criteria and "reflection" sheets to evaluate lessons they had prepared as "homework" for today's meeting—lessons that they developed to use or have used with teachers at their school.
>
> For the next 45 minutes, coaches worked in groups, discussing the lessons they created, how they fit the criteria for tiered lessons that differentiate instruction for different learners, and how they would coach the teacher to take the next steps. The conversations were spirited, and coaches exchanged many ideas and questions.
>
> Later in the meeting staff addressed the topic of "progress monitoring." They presented a chart and worksheet with "dummy" assessment data for six students—some with clear deficiencies in fluency, others with low scores in comprehension—and asked coaches what reading-level placement would be appropriate for each individual student and what instructional support they would suggest teachers offer to build each student's reading skills. After small-group discussion, coaches returned to the large group to review what they decided for each student. "I would do another assessment to see if that student really has a problem or they just blew off the test," suggested one coach. "Maybe they have had no practice or exposure to the MAZE [a group-administered test that measures fluency of silent reading and low-level

comprehension of passages that are like those students will encounter on the FCAT]. . . . Try balanced reading, spelling books," another coach offered. To close the activity, the administrator once again explained her intent, "So my attempt here has been to model for you what to do with teachers. Go through and look for the gray area kids. Who stands out? Use it to do lesson planning and tiered lessons. Look at the data and patterns that emerge." After this discussion, the facilitator moved on to some administrative matters.

Analysis of Effects

In this section we answer the final research question, examining the extent to which variation in coaches' time spent on data analysis related to teacher and student outcomes.

Frequency of Data Support Associated With Perceived Effects on Teaching

When asked directly to what extent the coach influenced any changes the teachers made to their instruction over the course of the year, 47% of all reading teachers and 40% of all social studies teachers reported that the reading coach had influenced them to make changes in their instruction to a moderate or great extent.[11] A minority of reading teachers (24%) and social studies teachers (34%) noted that the coach did not influence their instructional change at all.[12] Notably, teachers who received more frequent data support from the coach were significantly more likely than teachers with less frequent to no data support to attribute changes in their instruction to working with the coach. Specifically, 75% of reading teachers and 72% of social studies teachers who had received data support from the coach once a month or more reported that the coach had influenced changes in their instruction to a moderate or great extent, compared with 36% of reading teachers and 35% of social studies teachers who had received this support a few times a year or never.

Some of the teachers in our case study schools described how working with the coach to review assessment data enhanced their teaching methods. One reading teacher explained:

> The coach creates these charts where she has all of her classes on them. And then you get to look and see the differences [in student test scores]. She even puts the assessment in the scanner and she will show us how many of our kids got number one wrong—so that we can make instructional decisions.

> She is teaching us how to make instructional decisions based on assessment.
> It's not just okay that they took the test and these people failed it. She teaches
> us how to group kids for small groups for DI [direct instruction] lessons. . . .
> I found out a lot of my kids were not getting the main idea. So I had four kids
> that were not getting it. So she told me that I needed to put those kids
> together. I did a small group lesson up here [at her desk].

To understand coaches' influence on specific reading practices, we assessed any changes teachers reported in their instruction in general over the course of the year—changes that may or may not be attributable to the coach per se. These data provided an important overall picture of how teachers were adjusting their reading instruction over time. Not surprisingly, regardless of whether they worked with a coach or not, the vast majority of reading and social studies teachers reported making a number of specific changes to their instruction aimed at improving students' reading abilities over the course of the year (e.g., techniques to assist with reading comprehension, vocabulary, fluency).

Once again, these reported instructional changes were associated with the frequency with which the coach reviewed data with them. For example, reading teachers who received more frequent data support from the coach (meeting with the coach to review assessment data once or twice a month or more) were significantly more likely than their peers with less frequent data support (a few times a year or never) to report various changes in their instruction to a moderate or great extent, including the following:

- Introducing texts more thoroughly, providing students background knowledge about the text we will read (63% vs. 41%);
- Working to connect the readings to students' existing knowledge and lives more often (45% vs. 30%);
- Taking into account students' reading abilities/levels more often when designing tasks and assigning work (64% vs. 50%);
- Allowing students to select more of what they read (42% vs. 30%);
- Asking students to read texts out loud in class more frequently (48% vs. 38%).

The results of our Model 1 (Table 4) also indicate that the frequency with which the coach helped social studies teachers review assessment data had a strong positive association with perceptions of coaches' influence on instructional change. This suggests that, on average, social studies teachers who met with their coach more often to review assessment data reported more positive perceptions of coach influence (controlling for other variables in

Table 4
Model Results, Reading and Social Studies Teachers' Reports of Work With Coach Reported Separately

	Model 1: Teachers' Perception of Influence	Model 2: Students' Reading Achievement	Model 3: Students' Mathematics Achievement
Reading credential	−0.037 (0.093)	0.007 (0.018)	0.062* (0.025)
Years teaching reading	−0.009* (0.005)	−0.002* (0.001)	−0.003* (0.001)
Perceived coach quality	0.391** (0.120)	−0.032 (0.024)	−0.060 (0.034)
Ability to support adult learners	0.132* (0.060)	0.001 (0.011)	−0.017 (0.015)
Coach confidence	−0.123 (0.078)	0.021 (0.016)	−0.005 (0.022)
More experienced coach	0.161* (0.079)	−0.012 (0.017)	−0.024 (0.024)
Focus on integrating instruction across content areas	0.130* (0.054)	0.004 (0.011)	0.018 (0.016)
Reviewed assessment data with coach (reading teacher report)	−0.035 (0.089)	0.082*** (0.019)	0.073** (0.028)
Reviewed assessment data with coach (social studies teacher report)	0.355*** (0.098)	0.023 (0.019)	−0.031 (0.027)
Received individual coaching (reading teacher report)	0.510*** (0.099)	−0.061*** (0.018)	−0.029 (0.026)
Received individual coaching (social studies teacher report)	0.053 (0.133)	−0.020 (0.027)	0.002 (0.039)
Number of years the school had a coach	−0.013 (0.017)	0.010** (0.004)	0.007 (0.005)
Coach caseload	−0.084 (0.073)	0.043 (0.024)	0.081* (0.035)
Percentage of new teachers in the school	0.003 (0.002)	0.000 (0.000)	−0.001 (0.001)
Principal leadership	0.117 (0.111)	0.029 (0.020)	0.045 (0.028)

Note: All models include controls for student and school characteristics (not shown). Variable values are not standardized. Standard errors are given in parentheses.
*$p < .05$. **$p < .01$. ***$p < .001$.

the model). Similarly, the frequency with which the coach worked one on one with reading teachers had a positive association with perceptions of their influence.

As Table 4 indicates, a number of other variables were also significantly and positively related to perceptions of coach influence, including teachers'

overall views about coaches' quality, principals' ratings about coaches' understanding of how to support adult learners, and more experience coaching.

Frequency of Data Support
Associated With Reading Achievement

Table 4 shows the results of the model of reading achievement (Model 2). One of the only program features that we found significantly and positively related to better reading scores was the frequency with which reading teachers reported that the coach reviewed assessment data with them (either individually or in a group). The point estimate indicates that a one-unit increase in the scale used to measure how often the coach reviewed assessment data with the reading teachers increases reading achievement by 0.082 standard deviations. The standard deviation of this scale is 0.613, implying that an increase of one standard deviation leads to an improvement of reading achievement of 0.05 standard deviations. The magnitude of this association is small, especially relative to what would be expected from a program or intervention aimed at improving student achievement. Nonetheless, it should be borne in mind that data analysis support is at most only one aspect of having a reading coach and certainly not the entire coaching program. Thus, it remains noteworthy that we find a statistically significant association between data analysis support and achievement, even if the magnitude of this link is fairly small. Social studies teachers' reports of reviewing assessment data with the coach were also positively related to reading scores, but the estimated relationship was smaller than it was for reading teachers and not statistically significant.

How often reading teachers report receiving one-on-one coaching was negatively associated with reading achievement. This result is surprising because the program's theory of action posits that coaching is effective when coaches are able to work with individual teachers on issues related to classroom instruction. Furthermore, our models of proximal outcomes found that teachers who worked one on one with the coach were more positive about the impact of the coach than were teachers who did not have this experience, holding other factors constant.

How do we make sense of these puzzling results? First, when we estimated models where only one of these two program features was included, the results indicate that the coefficient on reviewed data remains positive and statistically significant (although it is smaller in magnitude), whereas the coefficient on individual coaching is very small and not statistically significant, which implies that this is not a robust finding for individual

coaching. Second, teacher reports of individual coaching and reviewing assessment data are highly correlated with each other (the correlation coefficient is .7), so there is limited variation with which to separately identify the effects of these two program features. Thus, the separate effects were estimated from only a small number of schools in which individual coaching is relatively high but reviewed assessment data is low. It is true that reading achievement in these schools is lower than it is in observationally similar schools. However, schools with relatively high levels of individual coaching generally are schools where coaches review assessment data with teachers relatively frequently, and in these schools, achievement also tends to be high. This finding may indicate that the content of individual coaching is important. In other words, if individual coaching is done without including a review of student data, it may be counterproductive.[13]

Frequency of Data Support Associated With Mathematics Achievement

As with reading achievement, we again found a positive relationship between reading teacher reports of coach reviewing assessment data and mathematics achievement (Model 3). The magnitude of this association is similar to that for reading achievement, although it is slightly smaller.

Implications

These results provide important insights into the role of coaches in supporting DDDM. Most important, it provides empirical evidence linking coach data support activities with both perceived teacher influence and student achievement. These findings provide the first step in filling a gap in the research base on coaching and DDDM. Although our data do not support causal inferences, they nonetheless suggest that more frequent data support from a coach is associated with higher student achievement (albeit small in magnitude) and more positive perceptions of coaches' influence on teacher practice.

These findings have several important implications for theory, policy, practice, and future research. Given the increased popularity of using data to guide school improvement and using coaches to assist teachers in this process, it is critical to understand how coaches perform this data support role and how to do it effectively. Our results and evidence from other research suggest that what makes this practice effective is not just helping teachers interpret the data (which may be particularly important in the

context of literacy for content-area teachers, many of whom lack a deep understanding of fluency, comprehension, etc.) but also helping them identify instructional strategies in response to these data. Analyzing data and taking action based on data are two different tasks. Taking action is often more challenging and requires more creativity than does analysis. Yet to date, taking action generally receives less attention, particularly in the professional development provided to educators (Marsh et al., 2006). Other research confirms the importance of providing training on how to use data and to connect them to practice (see earlier literature review). Thus, coaches may be bridging this important divide for teachers, helping them identify students' strengths and weaknesses and providing them with specific instructional strategies aligned with their needs. Furthermore, as learning theory predicts, teachers may be more likely to apply new skills and practices when they have a solid understanding of the reasons behind their use and why they are important.

Returning to one of the reading teachers mentioned earlier in this article, who discovered through working with the coach that a group of students were not grasping main ideas and needed additional small-group instruction, it is quite possible that examining the assessment results alone without coaching support or receiving coaching without the assessment results would not have resulted in the same instructional response. If the coach had simply told the teacher to work with those few students in a small group around main idea, she may not have done so effectively because she may not have understood the reason for doing so. Conversely, had the teacher simply discovered on her own through the data the students' weakness in this area, she may not have known how to address their needs or even considered grouping them together to provide tailored instruction in this way. In other words, there is nothing inherent in the way these assessment results are reported that clearly indicate an effective "solution" or response. And whereas organizational theory and notions of DDDM indicate that improvement is enhanced by responsiveness to data, one's responsiveness is clearly contingent on having the skills and knowledge to match the identified areas of need with effective instructional strategies—skills and knowledge that not all teachers posses. Thus, coaches appear to be situated in a critical nexus of data and action. Our understanding of DDDM and theories of organizational improvement may be incomplete without a more nuanced understanding of how individuals identify effective responses to data. Similarly, our understanding of how adults learn and how teachers improve their practice is greatly enhanced by an understanding of the role data play in facilitating the learning process.

Ultimately, our research suggests several lessons for administrators. First, to encourage this data analyst and support role, administrators should

continue providing professional development for coaches in this area, with a particular focus on taking action in response to these results. The case study district highlighted previously provides a good example of the in-depth training and support districts can provide to coaches around data use. Important questions to answer before designing such professional development include the following: What types of data are important (e.g., state tests scores, diagnostic assessment results, observational data on quality of instruction)? What is the most effective way to engage with teachers in this activity (e.g., individually, in groups)? What tools would assist coaches in their data work with teachers (e.g., user-friendly displays of student data, templates to help analyze individual student data)? What specific reading strategies are recommended to align with students' specific needs?

Second, district and school leaders should attend to several factors that may be constraining coaches' ability and opportunity to work with teachers and provide data-driven instructional support to teachers. Policy makers and administrators should consider ways to free up more time for coaches to spend in classrooms. For example, there may be easy steps to take to minimize administrative, assessment-related demands on coaches (which our analysis also indicates is negatively associated with teacher perceptions of the coach's influence on instruction, see Note 3). For instance, can other school staff do more of this administrative work? Could volunteers or temporary staff be trained to input the results into the state database? It also behooves school, district, and state leaders to discourage coaches from participating in excessive, non–reading-related assessment tasks (e.g., cases where coaches served as the FCAT coordinators for all subjects across the school). A lack of planning time built into the school day also may be minimizing opportunities for coaches to work individually with teachers. Obviously, addressing this barrier would require structural policy changes at the school or district level to make teachers more available to participate in one-on-one work.

Given that teacher resistance or lack of rapport between coach and teachers may constrain coaches' ability to provide support to all teachers (a third of coaches felt teacher reluctance to work with them was a hindrance to their work), administrators also may want to ensure that principals know how to hire high-quality coaches, to provide coaches with professional development focused on how to develop relationships with teachers and build trust, and to link new coaches with mentors who have faced similar situations. Finally, in some cases the coach caseload may be too large to allow coaches enough time to work with all teachers needing support if a school has a lot of new teachers or teachers who are new to reading (the median number of teachers per middle school was 65). Though higher

coach caseload had a negative association with perceptions of influence over teacher practice, it was not associated with reading achievement (and was positively related to mathematics achievement). Nevertheless, the issue remains important if in fact state policy makers want coaches to prioritize their time working directly with teachers and to improve instruction across the school. Obviously actions pertaining to coach caseload involve difficult resource decisions. Leaders should consider the needs of each school (i.e., student performance, numbers of inexperienced teachers) when assigning coaches and consider allocating more than one coach in large, high-needs schools when possible.

Finally, although these results provide useful information for policy makers and practitioners, the limitations of our data suggest several fruitful avenues for future researchers. First, researchers should consider assessing coaching implementation and achievement over a longer period of time than a year. This type of longitudinal coaching study could allow for a more careful discernment of the relationship between coaches' activities and teacher and student outcomes. In particular, if the effect of coaching on student achievement grows over time, a longitudinal analysis would be more sensitive to determining the relationship between coaches' activities and student outcomes. Such longitudinal studies could focus at the coach level (examining how an individual coach's effectiveness changes as he or she gains experience), at the student level (examining the cumulative effects of students' exposure to teachers who have benefited from coaching), and at the teacher level (examining how teachers' effectiveness changes as they work with a coach). Research with more direct observational measures of teacher practice would also add depth to our understanding of how coaches influence instruction, a critical intermediate outcome of coaching. Furthermore, researchers may consider using measures of achievement beyond state-standardized tests, such as specific reading assessments that provide more detailed information about specific reading skills and abilities. Finally, to further understand and enhance the roles coaches play in supporting DDDM, more research is needed to identify the specific skills and knowledge needed to effectively bridge the divide between data and practice for teachers, and how to build this capacity on a large scale.

Notes

1. There are 67 county districts in Florida and 8 nontraditional districts (e.g., Florida School for Deaf and Blind in Dozier/Okeechobee). In 2006-2007, virtually all of these districts participated in the program.

2. State documents and administrators interviewed at the state and local level frequently mentioned the "state's reading coach model," but this often referred to a job description and not to a delineation of the process, content, supports, and expected outcomes of reading coaching.

3. As a condition of participation, all districts, schools, and individuals were promised anonymity. Thus, we do not provide specific data or details on any organizations or individuals that could inadvertently disclose their identity. We use pseudonyms for some individuals and schools throughout the report.

4. One district denied access to its lowest-performing schools, which removed eight schools from the eligible population.

5. We achieved these response rates by conducting extensive online and paper follow-up reminders and by providing honoraria to responders ($25 gift card to coaches and teachers: Principals' schools received $100 for participating and providing us with lists of teachers and email addresses).

6. The academic year 2001-2002 was the first year when the FCAT (Florida's Comprehensive Assessment Test) was administered to all students in Grades 3 to 8.

7. For additional technical details on the modeling and the full set of model results see Note 3.

8. We asked for hours within a 2-week period because the state logs ask coaches to report their time for this window of time and coaches were thus familiar with thinking about their work in 2-week blocks.

9. Note that not all these categories are mutually exclusive.

10. This survey question followed a pervious multiitem question asking teachers to think about the ways in which their teaching in general was different at the end of the year compared with the beginning of the year and the extent to which they made a series of changes to their teaching over the course of the year. Thus, their reports of coach influence were anchored in an understanding of this list of changes.

11. Only 2% of reading teachers and 5% of social studies teachers reported not making any of the changes in instruction asked about on our survey over the course of the year—these teachers are included in the percentage of teachers reporting that the coach did not influence changes in their instruction at all.

12. Another possibility is that teachers with the most challenging students are more in need of coaching; thus, time teachers get might be associated with lower student achievement. Because our models control for prior academic achievement, we view this interpretation as unlikely. However, there may be unobserved student ability not captured by our controls for prior achievement, thus we cannot rule out this interpretation entirely.

References

Black, P., & William, D. (1998). Assessment and classroom living. *Assessment in Education: Principles, Policy, and Practice, 5*(1), 7-74.

Brown, J. S., Collins, A., & Dugrid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher, 18*(1), 32-42.

Brown, D., Reumann-Moore, R., Hugh, R., Christman, J. B., Riffer, M., du Plessis, P., et al. (2007). *Making a difference: Year two report of the Pennsylvania high school coaching initiative*. Philadelphia: Research for Action.

Brown, D., Reumann-Moore, R., Hugh, R., du Plessis, P., & Christman, J. B. (2006). *Promising inroads: Year one of the Pennsylvania high school coaching initiative*. Philadelphia: Research for Action.

Casserly, M. D. (2002). *Beating the odds: A city-by-city analysis of student performance and achievement gaps on state assessments*. Washington, DC: Council of the Great City Schools, 2002.

Celio, M. B., & Harvey, J. (2005). *Buried treasure: Developing an effective management guide from mountains of educational data*. Seattle, WA: Center on Reinventing Public Education.

Chen, E., Heritage, M., & Lee, J. (2005). Identifying and monitoring students' learning needs with technology. *Journal of Education for Students Placed at Risk, 10*, 309-332.

Choppin, J. (April, 2005). *Data use in practice: Examples from the school level*. Paper presented at the Annual Conference of the American Educational Research Association, New Orleans, LA.

Collins, A. M., Brown, J. S., & Holum, A. (1991). Cognitive apprenticeship: Making thinking visible. *American Educator, 15*(3), 6-11, 38-46.

Copland, M. (2003). Leadership of inquiry: Building and sustaining capacity for school improvement. *Educational Evaluation and Policy Analysis, 25*, 375-395.

Council of the Great City Schools. (2002). *Beating the odds II*. Washington, DC: Author.

Datnow, A., Park, V., & Wohlstetter, P. (2007). *Achieving with data: How high-performing school systems use data to improve instruction for elementary students*. Los Angeles, CA: Center on Educational Governance, University of Southern California.

Deming, W. E. (1986). *Out of crisis*. Cambridge: MIT Center for Advanced Engineering Study.

Deussen, T., Coskie, T., Robinson, L., & Autio, E. (2007). *"Coach" can mean many things: Five categories of literacy coaches in Reading First* (Issues and Answers Report, REL 2007-No. 005). Washington, DC: U.S. Department of Education.

Edmonds, R. (1979). Effective schools for the urban poor. *Educational Leadership, 37*(1), 15-24.

Feldman, J., & Tung, R. (April, 2001). *Whole school reform: How schools use the data-based inquiry and decision making process*. Paper presented at the 82nd Annual Meeting of the American Educational Research Association, Seattle, WA.

Florida Department of Education. (2004). *Just read, Florida! Middle school reading coach grant: Request for application*. Retrieved June 19, 2008, from http://www.justreadflorida.org/grants/MRC-RFP.asp

Foltos, L. (2007). Peer coaching: Changing classroom practice and enhancing student achievement. *Utah Special Educator, 27*(5), 34-37.

Galm, R., & Perry, G. (2004). Coaching moves beyond the gym. *Journal of Staff Development, 25*(2), 1-4.

Garet, M. S., Birman, B. F., Porter, A. C., Desimone, L., Herman, R., & Yoon, K. S. (1999). *Designing effective professional development: Lessons from the Eisenhower Program*. Washington, DC: U.S. Department of Education.

Garet, M. S., Porter, A. C., Desimone, L., Birman, B. F., & Yoon, K. S. (2001). What makes professional development effective? Results from a national sample of teachers. *American Educational Research Journal, 38*, 915-945.

Guinney, E. (2001). Coaching isn't just for athletes: The role of teacher leaders. *Phi Delta Kappan, 82*, 740-743.

Gutierrez, K., Crosland, K., & Berlin, D. (April, 2001). *Reconsidering coaching: Assisting teachers' literary practices in the zone of proximal development*. Paper presented at the annual meeting of the American Educational Research Association, Seattle, WA.

Hamilton, L. S., Stecher, B. M., Marsh, J., McCombs, J. S., Robyn, A., Russell, J., et al. (2007). *Implementing standards-based accountability under No Child Left Behind: Responses of superintendents, principals, and teachers in three states*. Santa Monica, CA: RAND Corporation.

Hasbrouk, J. E. (1997). Mediated peer coaching for training preservice teachers. *The Journal of Special Education, 31*, 251-271.

Hawley, W., & Valli, L. (1999). The essentials of effective professional development: A new consensus. In L. Darling-Hammond & G. Sykes (Eds.), *Teaching as the learning profession: Handbook of policy and practice* (pp. 151-180). San Francisco: Jossey Bass.

Herman, J., & Gribbons, B. (2001). *Lessons learned in using data to support school inquiry and continuous improvement: Final report to the Stuart Foundation*. Los Angeles: National Center for Research on Evaluation, Standards, and Student Testing.

Ingram, D., Louis, K. S., & Schroeder, R. G. (2004). Accountability policies and teacher decision making: Barriers to the use of data to improve practice. *Teachers College Record, 106*, 1258-1287.

Joyce, B., & Showers, B. (1996). Staff development as a comprehensive service organization. *Journal of Staff Development, 17*(1), 2-6.

Juran, J. M. (1988). *Juran on planning for quality*. New York: Free Press.

Kerr, K. A., Marsh, J. A., Ikemoto, G. S., Darilek, H., & Barney, H. (2006). Districtwide strategies to promote data use for instructional improvement. *American Journal of Education, 112*, 496-520.

Knight, J. (2006). Instructional coaching. *School Administrator, 63*(4), 36.

Kohler, F. W., Crilley, K. M., Shearer, D. D., & Good, G. (1997). Effects of peer coaching on teacher and student outcomes. *The Journal of Educational Research, 90*, 240-260.

Kohler, F. W., Ezell, H. K., & Paluselli, M. (1999). Promoting changes in teachers' conduct of student pair activities: An examination of reciprocal peer coaching. *Journal of Special Education, 33*, 154-165.

Kohler, F., McCullough, K., & Buchan, K. (1995). Using peer coaching to enhance preschool teachers' development and refinement of classroom activities. *Early Education and Development, 6*, 215-239.

Lave, J. (1988). *Cognition in practice*. New York: Cambridge University Press.

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. New York: Cambridge University Press.

Love, N. (2004). Taking data to new depths. *Journal of Staff Development, 25*(4), 22-26.

Marsh, J. A., Pane, J. F., & Hamilton, L. S. (2006). *Making sense of data-driven decision making in education: Evidence from Recent RAND research*. Santa Monica, CA: RAND Corporation. Retrieved June 19, 2008, from http://www.rand.org/pubs/occasional_papers/OP170/

Marsh, J. A., McCombs, J. S., Lockwood, J. R., Martorell, F., Gershwin, D., Naftel, S., et al. (2008). *Supporting Literacy Across the Sunshine State: A Study of Florida Middle School Reading Coaches*. Santa Monica, CA: RAND Corporation.

Mason, S. (2002). *Turning data into knowledge: Lessons from six Milwaukee public schools*. Madison: Wisconsin Center for Education Research.

Massell, D. (2001). The theory and practice of using data to build capacity: State and local strategies and their effects. In S. H. Fuhrman (Ed.), *From the capitol to the classroom: Standards-based reform in the states*. Chicago: University of Chicago Press, 148-169.

Mayer, D. (1999). Measuring instructional practice: Can policymakers trust survey data? *Educational Evaluation and Policy Analysis, 21*, 29-45.

Morgan, R. L., Menlove, R., Salzberg, C. L., & Hudson, P. (1994). Effects of peer coaching on the acquisition of direct instruction skills by low-performing pre-service teachers. *Journal of Special Education, 28*, 59-76.

Mullens, J. E., & Gayler, K. (1999). *Measuring classroom instructional processes: Using survey and case study field test results to improve item construction*. Washington, DC: National Center for Education Statistics.

Munro, P., & Elliot, J. (1987). Instructional growth through peer coaching. *Journal of Staff Development, 8*, 25-28.

Murnane, R. J., Sharkey, N. S., & Boudett, K. P. (2005). Using student-assessment results to improve instruction: Lessons from a workshop. *Journal of Education for Students Placed at Risk, 10*, 269-280.

Neufeld, B., & Roper, D. (2003). *Coaching: A strategy for developing instructional capacity*. New York: The Annenberg Institute for School Reform and the Annenberg Institute Program on Education.

Ragland, M., Clubine, B., Constable, D., & Smith, P. A. (2002). *Expecting success: A study of five high performing, high poverty schools*. Washington, DC: Council of Chief State School Officers and the Charles A. Dana Center at the University of Texas at Austin.

Richards, A. (2003). *"Making our own road": The emergence of school-based staff developers in America's public schools*. New York: Edna McConnell Clark Foundation.

Rogoff, B. (1990). *Apprenticeship in thinking: Cognitive development in social context*. New York: Oxford University Press.

Russo, A. (July/August, 2004). School-based coaching. *Harvard Educational Letter*. Retrieved June 30, 2008, from http://www.edletter.org/past/issues/2004-ja/coaching.shtml

Senge, P. (1990). *The fifth discipline: The art and practice of the learning organization*. Garden City, NY: Doubleday.

Showers, B., & Joyce, B. (1996). The evolution of peer coaching. *Educational Leadership, 53*(6), 12-16.

Smithson, J. L., & Porter, A. C. (1994). *Measuring classroom practice: Lessons learned from efforts to describe the enacted curriculum—the reform up close study* (CPRE Research Report Series #31). Madison: University of Wisconsin.

Snipes, J., Doolittle, F., & Herlihy, C. (2002). *Foundations for success: Case studies of how urban school systems improve student achievement*. Washington, DC: MDRC and the Council of Great City Schools.

Supovitz, J. A., & Klein, V. (2003). *Mapping a course for improved student learning: How innovative schools systematically use student performance data to guide improvement*. Philadelphia: Consortium for Policy Research in Education, University of Pennsylvania Graduate School of Education.

Symonds, K. W. (2003). *After the test: How schools are using data to close the achievement gap*. San Francisco: Bay Area School Reform Collaborative.

Tharp, R. G., & Gallimore, R. (1988). *Rousing minds to life: Teaching, learning, and schooling in social context*. Cambridge, UK: Cambridge University Press.

Vaughan, D. (1996). *The Challenger launch decision: Risky technology, culture, and deviance at NASA*. Chicago: University of Chicago Press.

Veenman, S., Denessen, E., Gerrits, J., & Kenter, J. (2001). Evaluation of a coaching program for cooperating teachers. *Educational Studies, 27*, 317-340.

Wayman, J. C., & Stringfield, S. (2006). Technology-supported involvement of entire faculties in examination of student data for instructional improvement. *American Journal of Education, 112*, 549-571.

Williams, T., Kirst, M., & Haertel, E. (2005). *Similar students, different results: Why do some schools do better? A large-scale survey of California elementary schools serving low-income students*. Mountain View, CA: EdSource.

Wong, K., & Nicotera, A. (2006). Peer coaching as a strategy to build instructional capacity in low performing schools. In K. Wong & S. Rutledge (Eds.), *System-wide efforts to improve student achievement*. Greenwich, CT: Information Age Publishing, 271-302.

Julie A. Marsh is a Policy Researcher at RAND, specializing in research on policy implementation, district-level educational reform, accountability, and school-community collaboration. She has recently conducted studies on standards-based accountability, district-intermediary partnerships, and promotion and retention policies.

Jennifer Sloan McCombs is a Policy Researcher at RAND. Her education research has focused on policies aimed at improving outcomes for at-risk students including improving adolescent literacy, out of school time programs, teacher education and professional development, high-poverty schools, and instructional practices.

Paco Martorell is an Associate Economist at RAND, specializing in applied microeconomics. He has conducted research on high school exit exams, college remediation, principal training programs, and the No Child Left Behind Act.