Appendix A. Annotated Bibliography

This appendix describes key references that provide additional support for each of the Scoring Guide areas.

Scoring Guide Area 1: Implementing Systemic Change


This study of a small set of schools that were reconstituted in an urban area (pg. 555) found that students in reconstituted schools experience sizable and significant gains in ELA during the first two years of reconstitution, but insignificant effects for math. Changes in the state-wide assessment prevented these schools from being studied in subsequent years (pg. 556); however, case study data reflected that while reconstitution initially improves the student achievement at the school, the effects diminish over time (pg. 570). The authors suggest that it may be helpful for districts to maintain support in the form of funding and providing other resources for several years (pg. 571).


The authors note that there are limitations on the overall quantity and quality of the research base; however, the effects of the comprehensive school reform model appear promising. Schools that implemented the model for five years or more showed particularly strong effects (pg. 125).


The authors present the results of an 11-year longitudinal study of the America's Choice comprehensive school reform design focused on student learning gains. The study was conducted in Rochester, New York and compared test scores of students attending America's Choice schools with the scores of students who attended other schools and students who attended the same schools before America's Choice was implemented. There were significant annual effects, which accumulated over time in the elementary and middle grades (pg. 231). This study also found that over time, particularly after the fifth year of implementation, the effects dropped off and that although the effects were significant, students who were working below grade level did not catch up with grade-level peers (pg. 253). The America's Choice model emphasizes ongoing assessment and differentiation of instruction (pg. 252).


The authors reflect that only a few districts or schools have chosen to restart schools as charters. Case studies indicate several benefits of restarting a school as a charter including the freedom to hire, place, and remove staff; provide professional development and incentive; to use time as deemed best for students; adopt curriculum and implement other academic services; allocate dollars to priority areas and to own and maintain facilities (pg. 20). Case studies reflect improvements in student performance in some schools (pg. 12).

The author reflects that evidence regarding the effects of charter schools and education management organizations focuses on primarily on charter schools in general. Student achievement results are mixed when comparing student performance in charter schools to that of students in other schools ((pg. 27). It is unclear if true flexibility is afforded to charter schools that are low-performing or if that flexibility matters when it comes to student achievement (pg. 28).


The authors examine using magnet programs to improve the quality of education in urban areas and also to facilitate integration of schools. They note that “While desegregation does not ‘predict’ quality, within magnets a racial balance does predict academic gains. Integration and quality are highly associated; each is a correlative facet of effectiveness,” (pg. 134). A variety of factors in success are noted for schools that were studied. These include leadership of the principal, parental support, coordinated instructional program, and use of community resources (pg. 403, 412).


Results of a study conducted in Connecticut’s central cities indicate that “attendance at an interdistrict magnet high school has positive effects on the math and reading achievement of central city students and that interdistrict magnet middle schools have positive effects on reading achievement,” (pg. 323).


The author reflects that results of a study in American cities indicating that magnet schools were more effective than public comprehensive high schools in raising proficiency in science, reading and social studies (pg 1). In addition, principals of magnet schools reported “slightly more positive academic climates, on average, than principals in comprehensive schools,” (pg. 8).


The author states that there are “a number of studies over the past quarter-century that have found that magnet schools have higher levels of achievement than do other schools, and produce faster achievement gains in most subjects” (pg. 8). In addition, the magnet model is one where “schools seek to improve the performance of low-income students by drawing into a high-poverty school a contingent of middle class students” (pg. 8).


A study of magnet schools in Duval County Public Schools in Florida found that academic achievement of students attending magnet schools exceeded that of students who attended nonmagnet schools. The schools were established as part of a plan to desegregate the district (pg. 1).
Scoring Guide Area 2: Establishing Strong Leadership


This practice guide addressing turnaround of chronically low-performing schools recommends that strong leadership signal the need for dramatic change. It is important that principals “demonstrate commitment to developing a learning community for students and staff with the primary focus of the school on learning with staff and students working together toward that goal” (pg. 10). School leaders also signal change through clear communication, creating high expectations, sharing leadership and authority, demonstrating a willingness to make the same types of changes asked of their staff, identifying advocates with the staff, building a consensus that permeates the staff, ensuring that the maximum amount of classroom time is focused on instruction and establishing a cohesive culture (pg. 10-11). The current principal may be able to signal change; however, there may need to be a change in leadership to communicate the need for a dramatic change in the school (pg. 11).


The authors “describe the literature on principal behaviors linked to improved student achievement” (pg. 9). The behaviors are organized into five domains which include instructional management, internal relations, organizational management, administrative duties, and external relations. Under instructional management, behaviors such as “monitoring and providing feedback to teachers and student,” “having a vision for learning,” “providing support and professional development to teachers,” and “using data to drive decision-making,” were found to have positive relationships with student achievement. One study found that “promoting high standards for student learning ($r = .55-.61$) and having a rigorous curriculum ($r = .42-.47$) were most highly correlated with English language arts achievement in grades 3-5 and that performance accountability was significantly correlated in grade 3 ($r = .37$; Reardon, 2011)” (pg. 9-10). Eight of nine studies examined found a positive relationship between internal relations and student achievement while three of five studies reflected positive relationships between the time that principals spent on organizational management and student achievement. No studies found any relationship between principals’ time spent on administrative duties and student achievement. There were mixed results when it came to time spent devoted to external relationships and student achievement with school-community links in high-poverty and rural schools positively related to student achievement.


The authors of this study examined leadership at the school, district, and state level with the purpose to “identify the nature of successful educational leadership and to better understand how such leadership can improve educational practices and student learning” (pg. 7). At the school level, the authors reflected that among other findings that “collective leadership has a
stronger influence on student achievement than individual leadership" (pg. 19). Data suggests that “collective leadership has modest but significant indirect effects on student achievement” (pg. 28) as it positively effects teacher variables such as work setting and motivation which, in turn, impact student achievement.


The authors of this study examined the relationship between distributed leadership and academic capacity when observed over time and how distributed leadership impacts school improvement and subsequent growth in math (pg. 677). They “found support for the hypothesis that school leadership and capacity building are mutually reinforcing in their effects on each other over time,” and that “changes in these mutually reinforcing constructs were also positively associated with school growth rates in math. The effect size for change in academic capacity was almost 0.2” (pg. 679-680).


The authors examined a study investigating distributed or collaborative leadership. The study found that although there was no evidence of direct effect of collaborative or distributed leadership on student achievement, there was consistent indirect effects (pg. 12). The study found significant effect on changes in school academic capacity “which in turn had a significant effect on growth in student achievement in English language arts” (pg. 12).


The authors of this study examined “a sample of 20 Ohio schools that participated in a school turnaround program and found that participating schools experienced meaningful improvements in student achievement after completing the two-year program” (pg. 675). These schools investigated the implementation of a School Turnaround Specialist Program (STSP) where it was required that the principal and at least half of the school’s prior staff would be replaced. That said, the principal was replaced in only six of the 20 schools (pg. 691). Professional development to the principal and other leaders of the school the summer before the program was implemented and considerable support was provided to the principals through mentoring (pg. 679). “The schools examined as a part of this study demonstrated statistically and practically significant growth in student achievement within 2 years of participating in STSP” (pg. 694).
The practice guide states that a “comprehensive curriculum review can ensure that the curriculum aligns with state and local standards and meets the needs of all students (pg. 19). In addition, the What Works Clearinghouse establishes levels of evidence for assessing the quality of evidence supporting educational programs and practices (pg. 3).

The practice guide also indicates that schools need to “examine student achievement data to identify gaps and weaknesses in student learning….they can examine student learning through standards-based assessments and classroom assessments” (pg. 17). In addition, “school personnel can also look at data on factors that contribute to or impeded student learning, such as attendance, discipline, and fiscal expenditures” (pg. 17).


“This study examined data use and conditions influencing data use by typical principals and teachers, as well as the relationship between data use and student performance” (pg. 292). The authors note that data should be accessible, timely, and valid. In addition, the staff should have the expertise to analyze the data correctly (pgs. 296-297). “It is not data use per se that affects the quality of teaching and learning; rather it is the appropriateness of actions actually taken based on data-informed decisions about the nature of the problem and how it might be solved (pg. 321).


This study investigated a school-wide data-based decision-making (DBDM) intervention in primary schools in The Netherlands. The intervention involved a two-year training course in DBDM for primary school teams (pg. 366). It was hypothesized that “implementing DBDM will lead to changes in teacher’s classroom practices, which in turn will lead to student achievement growth in mathematics” (pg. 370). Results indicated that the intervention “can lead to a considerable improvement in the correct interpretation of student achievement data” (pg. 387) and there were positive effects on student achievement. In addition, the intervention “significantly improved the performances of students in low socioeconomic schools’ (pgs. 360-361).


This practice guide recommends that a variety of data is collected about student learning. Multiple data sources are important because, “no single assessment provides all the information teachers need to make informed instructional decisions” (pg. 11). Data collected may include “curriculum-based unit tests; class projects; classwork and homework; records from parent meetings and phone calls; classroom behavior charts; individualized education plans;
and prior data from students’ cumulative folders” (pg. 13).


The practice guide reflects that schools in need of improvement should “monitor progress and make adjustments” (pg. 17). Once schools have identified areas that needed improvement and develop a plan to improve instruction, they should continually monitor progress. In the schools cited in the practice guide, all of them used benchmark assessments or in some way systematically monitored student achievement and progress toward instructional goals (pg. 17). This was done so instruction could be modified as needed.


This practice guide recommends that teachers interpret data, develop a hypothesis about how to improve student learning (pg. 14), modify instruction to test the hypothesis, and continue the cycle to increase student learning (pg. 15). Modifying instruction may mean allocating more time, reordering the curriculum, identifying particular students in need of assistance with specific skills, attempting to teach complex skills in new ways, improving alignment between performance expectations among grade levels, or better aligning curricular alignment in the school (pg. 15).


The authors reflect that in order to improve teaching and learning, systematic, usable information regarding individual student performance and progress at the classroom level must be available (pg. 247). The study examined the “use of a computerized curriculum-based instructional management system in addition to ongoing math instruction” (pg. 248). The system allowed teachers to differentiate instruction based on data. Results reflect a positive effect with students in classrooms implementing the system demonstrating more growth than students in classrooms that did not implement the system (pg. 259).


This study examined the association between student performance on the 2012 Florida Comprehensive Assessment Test (FCAT) and their scores on the Florida Assessment for Instruction in Reading (FAIR) during three assessment periods throughout the year. In addition, the authors looked at the effects of adding FAIR as a means of preventing errors while identifying students in need of intervention (pg. i). The study showed a strong correlation between FAIR FCAT Success Probability (FSP) scores and performance on the 2012 FCAT at all grade levels. In addition, while FCAT could be used to identify students at risk/not at risk of meeting grade
level standards the following school year, implementing FAIR as a progress monitoring tool throughout the school year decreased the percentage of students that were misidentified. For example, “using FAIR FSP scores (which combine the FAIR Reading Comprehension Assessment with the 2011 FCAT 2.0 score) reduced underidentification from 21 percent in grade 4 to 4-6 percent” (pg. 9).


This study, conducted in a large urban district in Florida, examined how teachers and school staff administered computer-adaptive assessments of literacy to English learner students in grades 3-5 and how they used the assessments to monitor students’ growth in literacy skills. (pgs. 1-2). “Reliably measuring the literacy skills of English learner students can be challenging. Assessments typically address only grade-level proficiency, do not provide instructionally relevant information, and are not developmentally scaled to measure change over time” (pg. 2). The Florida Assessments for Instruction in Reading (FAIR) K-2 system was used because of the low level of English proficiency. The study found that teachers partnered with each other so that the assessment could be delivered within the required timeframe. Students’ literacy skills improved during the course of the year, but most students remained at the same grade level in the FAIR K-2 system at the end of the school year. Teachers found the data helpful as they could use it to plan and adjust instruction as needed.

**Scoring Guide Area 4: Developing and Retaining a High-Quality Staff**


The authors reflect that “the school leader needs to build a staff that is committed to the school’s improvement goals and qualified to meet them” (pg. 27). In addition, while not a focus of the specific recommendation in the practice guide, the author’s state that “professional development to help staff reach the school’s goals is an essential element of all school reform efforts and should be a part of turnaround schools,” (pg. 27).


The authors examined nine studies that addressed the effect of teacher professional development on student achievement in mathematics, science, and reading or English language arts. Five of the studies were randomized controlled trials and met the What Works Clearinghouse evidence standards without reservation. Four studies met the evidence standards with reservations (pg. iii). In all studies the professional development provided was directly to teachers and not through a “train the trainer” approach. It was delivered by those who created the professional development. It was also found that studies that had “more than 14 hours of professional development showed a positive and significant effect on student achievement from professional development” (pg.73). Further, the authors state that “First, professional development enhances teacher knowledge and skills. Second, better knowledge and skills improve
classroom teaching. Third, improved teaching raises student achievement….If a teacher fails to apply new ideas from professional development to classroom instruction, students will not benefit from the teacher’s professional development” (pg. 4).


Professional development was a key component of the set of instructional improvement interventions that were examined by this study. The study was conducted in high schools and included professional development for both mathematics and English teachers (pg. 3). The authors explain that professional development should be content focused, “meaning that it extends and intensifies teacher knowledge of a subject area and how children learn subject specific content” (pg. 5-6). Students attending treatment schools had higher math scores than those who attended schools not in the treatment group (pg. 19). Although the professional development component alone was not studied, it was a major component of the intervention set.


This study investigated a dynamic integrated approach to professional development as opposed to a holistic approach. The dynamic approach focused on factors that describe the teachers’ instructional role and are associated with student outcomes such as questioning, classroom assessment, and teacher-modeling while the holistic approach focused on teachers’ beliefs, experiences, and reflection on teaching practices (pgs. 291-292). The study found that teachers that had participated in the dynamic approach to professional development were more effective than those participating in the holistic approach model (pg. 303).


The authors conducted a quasi-experimental investigation focused on the effects of establishing grade-level teams focused on student learning on student achievement. Professional development was provided to the principal and the teachers on establishing the teams and professional development occurred during team meetings. Student achievement at schools in the treatment group improved at a faster rate than student achievement at comparable schools who did not implement grade-level teams (pg. 1).


The authors of this study investigated whether student reading comprehension could be improved through a professional development program emphasizing goals, data use, and instruction (pg. 1). Second and third grade teachers received 40 hours of professional development over the course of the school year. They attended meetings after school and completed homework assignments. Participation was voluntary and free of charge; however, no additional compensation was provided to teachers (pg. 140). The study found a positive effect on student achievement and at the end of the program “students in the experimental condition were more than half a year ahead of students in the control condition” (pg. 150).

The authors conducted an evaluation of a statewide reading coach program in Florida middle schools. "Using achievement data from nearly 1,000 Florida middle schools from the 1997-1998 through 2005-2006 school years, we find that receiving a state-funded coach was associated with statistically significant improvements in average annual reading achievement gains for two of the four cohorts of schools analyzed" (pg. 1). It is possible that the lack of effects for one of the cohorts (2006) may have been due to the fact that implementation had taken place for only one year. The other cohort (2004) was small and it is possible that idiosyncrasies of the schools came into play (pg. 383). Overall, "our results might be more supportive of positive coaching effects than the simple count of statistically significant findings would imply" (pg. 383).


The authors examined how coaches support data-driven decision-making and "the extent to which these efforts are associated with improvements in teaching and student achievement" (pg. 873). Data support was one of many activities to which coaches devoted their time. Coaches spent time administering and coordinating assessments, working with individual teachers, managing resources and materials, as well as working with groups of teachers. They also, in some cases, devoted time to non-coaching tasks such as substitute teaching or performing "duties" such as lunch duty or bus duty. More experienced coaches spent more time in supporting data-driven decision-making. A positive relationship was found between data analysis and student achievement (pg. 898).


The authors conducted a group-randomized trial in which schools within one district received a content-focused coach (CFC) and other schools continued with literacy coaching that was standard practice in the district (pg. 38). The CFC coaches helped teachers become more proficient at planning, teaching, and reflecting on their lessons and emphasized the Questioning the Author (QtA) approach which is a discussion-based approach to reading comprehension (pg. 37). Coaches met with teachers in weekly grade level teams and monthly in their classrooms. The study found that the CFC program had a positive effect on the quality of classroom discussions and "by the end of that academic year, students in the CFC schools demonstrated significantly higher reading achievement than their comparison group peers" (pg. 44). In addition, the CFC program helped to close the gap between ELL and non-ELL students in the study (pg. 44).


This study investigated whether schools implementing the Teacher Advancement Program (TAP) outperformed comparable schools on an annual basis, outperformed its controls, whether fidelity to implementation influenced student achievement and teacher satisfaction with the program (pg. 334). "By aggressively recruiting new teachers, providing a career continuum, introducing teacher-led professional development, implementing rigorous teacher accountability, and paying teachers based on their position, teaching skills and how much their students achieve, TAP schools change their organizational structure to support and reward
high-quality instruction” (pg. 327). The student achievement in TAP schools grew significantly when compared to the controls although the magnitude of the gains varied by school and fidelity of implementation (pg. 327).

**Scoring Guide Area 5: Creating a Positive School Climate and Culture**


This study examined a model designed to support school improvement efforts by emphasizing youth development, parent and family engagement and support, health and social services and community partnerships (pg. 192). The authors looked at the types of capacity-related innovations developed to support the model, whether school-level perceptions improve throughout implementation, and whether or not school-level indicators of academic achievement improve over the course of implementation. The study found that roles and responsibilities of staff changed to focus on the model and that innovations occurred that resulted in the use of data for planning. There was an improvement in the perception of the school climate and in academic motivation and implementation resulted in increased student achievement (pg. 198).


The authors of this study examined the caring practices in two higher performing and two lower performing urban high schools. It was found that “higher performing schools demonstrated caring communities, where interpersonal relationships and high academic expectations were prevalent throughout the school” (pg. 406). Factors such as “strong leadership support, caring as a core school value, and abundant curricular and extracurricular structures” (pg. 406) were less prevalent in lower performing schools that had only isolated instances of care.


The authors examined whether a supportive professional environment is associated with teacher improvement over time (pg. 476). The professional environment included factors such as the extent to which the school was a safe environment and order prevailed, the opportunity for peer collaboration, the support of the principal, the opportunity for teachers to participate in professional development, the respect, openness, and commitment to student achievement and a teacher evaluation process that provided teachers with meaningful feedback which could be used to improve instruction (pg. 480). The study concluded that teachers “working in more supportive professional environments improve their effectiveness more over time than teachers working in less supportive environments” (pg. 476).


The practice guide recommends providing “visible improvements early in the turnaround process” (pg. 22). These can include making improvements to the physical environment such as painting, ensuring the school building and grounds are clean, and fixing anything that is broken (pg. 25). In addition, establishing a safe and orderly environment by implementing an
approach to discipline that demonstrates the presence of administrators and safety officers, involves parents, and provides a means of dispensing discipline swiftly and fairly can also impact student learning and be implemented fairly quickly.