

District and School Practices and Assessments to Support a Learning-Centered Vision

Elena Diaz-Bilello, University of Colorado Boulder

Jared Anthony, Carnegie Foundation for the Advancement of Teaching

CONTENTS

INTRODUCTION	168
THE ROLE OF THE DISTRICT.....	170
A Culture of Testing	171
ASSESSMENTS TO SUPPORT AMBITIOUS TEACHING AND LEARNING	172
An Example of Classroom Assessment Supporting Ambitious Teaching	174
Distinguishing the Role of Distal Assessments.....	176
Establishing School Partnerships	177
ESTABLISHING AN INSTRUCTIONAL INFRASTRUCTURE FOR AMBITIOUS TEACHING	178
Access to High-Quality Curriculum	179
Professional Learning.....	184
Grading	188
TAKING A <i>LEARNING SYSTEMS</i> APPROACH TO IMPLEMENTING A LEARNING-CENTERED VISION	192
CONCLUSION	193
REFERENCES.....	194

INTRODUCTION¹

In 2017, *Education Week* reported that in a nationally representative sample of more than 500 U.S. K–12 teachers, approximately 85 percent indicated that they had experienced new changes or reforms in the past two years, and more than 58 percent indicated that they were experiencing “reform fatigue” (Loewus, 2017). Most teacher respondents (85 percent) further shared that “as soon as they get a handle on a new reform, it changes” (Loewus, 2017). In Chapter 8 of this volume, “Developing, Implementing, and Institutionalizing Complex Educational Innovations: Considerations for Balanced Assessment Systems,” Peurach and Linn frame this persistent state of churning reforms as part of the conventional narrative that “policy-level fragmentation, incoherence, and turbulence” foster the same within local districts, as their central offices and schools attended to changing policy ambitions and priorities.

Based on the perspectives from educators and educational researchers outlined in *Education Week*, it should come as little surprise that incoherence also characterizes the design of assessment systems in many districts and schools. For example, the website of one large urban school district in a Western state shows the district’s “balanced assessment system framework” as a large menu of assessments that fall under either Assessments for Learning (formative) or Assessments of Learning (summative). Under the formative category, there is a list of more than 12 assessments, including universal screeners and district-mandated interim assessments. Under the summative category, there is a large list of district-required and state-required year-end tests. The website notes that multiple types of assessments and data from multiple occasions are needed to guide instruction and improve student performance. In other words: it seems that the simple act of selecting and administering assessments under both categories defines balance in this assessment system without consideration for whether data from all of these assessments are communicating a coherent picture of student performance to effectively inform instructional steps.

Unfortunately, the misconception that a balanced assessment system means using several different types of assessments is also perpetuated by some vendors that sell interim assessments. As Marion (2021) notes, several test vendors claim that a balanced assessment system should consist of a selection of assessments—formative, interim, and summative—that teachers and administrators can combine to form a comprehensive picture of student learning. This conception of a balanced assessment system will likely lead to a patchwork of assessments that do not advance a particular vision or model of learning and often inspires a refrain heard in the education field about districts and schools being “data rich but information poor.”

In this chapter, we discuss the practices and structures employed by districts and schools functioning as *learning systems* (see Chapter 8 of this volume, “Developing, Implementing, and Institutionalizing Complex Educational Innovations: Consider-

¹ We are so grateful to our reviewers (Amy Berman, Debbie Durrence, Peter Leonard, Jonathan Supovitz, and Scott Marion) for providing us with thoughtful feedback that challenged and extended our thinking. We also want to thank the school districts that we work with for the incredible partnership opportunities. We could not have written this chapter without learning from our work with you. And finally, we cannot thank Lorrie Shepard enough, as she always and generously made the time to share her insights and provide us with feedback throughout this project.

ations for Balanced Assessment Systems,” for a fuller discussion). We argue that these practices and strategies can be used to support and sustain assessments that focus on ambitious teaching and learning. Districts and schools characterized as being in the *learning systems* stage of the systems continuum that includes *school systems*, *education systems*, and *learning systems* are “distinguished by capabilities to engage diverse stakeholders ... in collaborating to develop the shared understandings, knowledge, and values needed to identify and address local educational ambitions, needs, and problems” (Chapter 8 in this volume, p. 259). Thus, districts and schools in the *learning systems* stage are best positioned to implement a learning-centered vision as addressed throughout this volume. *Learning systems* stage institutions demonstrate the importance of building collaborative networks between districts, schools, and other key partners to tackle important problems and issues, such as addressing inequities in the education system.

This chapter opens with a brief account of how school districts have recently influenced the instructional work taking place in U.S. schools. This accounting shows that districts do not have a long legacy of engaging with schools as *learning systems* and demonstrates that several districts remain in what Peurach and Russell (in Chapter 8 of this volume, “Developing, Implementing, and Institutionalizing Complex Educational Innovations: Considerations for Balanced Assessment Systems”) refer to as the *education systems* and *school systems* stages. Districts with *school systems* characteristics focus on supporting the business and administrative functions of operating schools that serve their communities and tend to more weakly support improving schools’ educational work. Districts with *education systems* characteristics, on the other hand, focus on improving teaching and learning, but their goals often involve achieving technical effectiveness and efficiency in response to federal and state policy goals and interventions. Balanced assessment systems operating in districts located in either of these two stages would not resemble the type of instructional and assessment work that supports ambitious teaching and learning practices, described in great detail by Ruiz-Primo and Furtak in Chapter 4 of this volume, “Classroom Activity Systems to Support Ambitious Teaching and Assessment.”

We then consider what it would look like for a *learning system* district to use assessments supporting ambitious teaching and learning. Under this scenario, districts would prioritize the use of classroom assessments and use federal- and state-required test results in ways that do not detract and instead support effective classroom assessment practices. We walk through an example of classroom assessments that support this learning-centered vision, and in so doing clarify ideal features and qualities of these assessments.

We note, as Kang et al. (2016) documented in their study of teachers who use rich instructional tasks, that having access to rich tasks and assessments does not necessarily mean that they will be implemented effectively—the successful enactment of rich classroom tasks and assessments requires a strong supportive infrastructure. Thus, we discuss *instructional infrastructure* and provide examples of ambitious teaching and learning practices and structures that district and school personnel located in a *learn-*

ing system stage should support and invest in (Hopkins & Spillane, 2015; Hopkins & Woulfin, 2015). *Instructional infrastructure* components discussed in order of priority are:

- high-quality curricula,
- professional learning, and
- grading.

Although we discuss each component separately, the three should work together to establish an assessment system integrated with instruction. The *instructional infrastructure* literature commonly distinguishes assessment as a separate component of teaching that supports instruction, but we do not do so, because, as discussed by Ruiz-Primo and Furtak in Chapter 4 of this volume, “Classroom Activity Systems to Support Ambitious Teaching and Assessment,” the types of assessment that support ambitious teaching and balanced assessment are inseparable from instruction.

Lastly, we address how districts might begin working with schools to move toward this bold vision for teaching and learning while simultaneously engaging in necessary evaluations to monitor implementation. Because districts and schools can be located anywhere along the *school systems, education systems, or learning systems* continuum, we provide only general ideas for how districts and schools can begin this highly complex work aimed at changing organizational behaviors, cultures, policies, and structures.

THE ROLE OF THE DISTRICT

In most school districts in the United States today, central office leadership sets the vision and policies for improving teaching and learning (Honig & Coburn, 2008). Districts communicate budget priorities, provide instructional frameworks and curricular materials, set expectations for assessment strategies, and provide professional development opportunities for school-based personnel (Coburn et al., 2009; Honig & Venkateswaran, 2012; Penuel et al., 2017).

This has not always been the case. Historically, district personnel tended to focus on fulfilling business and compliance functions rather than implementing a vision of teaching and learning for their schools (Honig, 2013). Peurach and Russell in Chapter 8 of this volume, “Developing, Implementing, and Institutionalizing Complex Educational Innovations: Considerations for Balanced Assessment Systems,” classify districts that focus on business and compliance as being in a *school systems* stage, in which district management concern themselves with the structural and procedural activities needed to deliver educational services.

Districts first began to play a more active role in teaching and learning in the 1980s, when research highlighted the important role that they could and should play in fostering effective schools (Mac Iver & Farley, 2003). Over the subsequent two decades, such research generated momentum for researchers and policymakers to consider how districts could prominently steer instructional reforms (Honig & Coburn, 2008). The advent of the federal No Child Left Behind Act (NCLB) of 2001 further cemented districts’ direct involvement in steering the teaching and learning visions for their schools, particularly because NCLB provided monetary incentives to encourage districts to take a leading role in evidence-based school improvement work (Anderson & Young, 2018; Leithwood et al., 2019; No Child Left Behind Act, 2001).

The passage of NCLB not only encouraged many districts to take a prominent role in improving teaching and learning in their schools but was also the starting point for these districts to focus substantial resources and energy on raising test scores to improve school accountability ratings (Au, 2007; Blazar & Pollard, 2017). In shifting their resources and attention to these types of educational work, many districts entered what Peurach and Russell in Chapter 8 of this volume, “Developing, Implementing, and Institutionalizing Complex Educational Innovations: Considerations for Balanced Assessment Systems,” classify as the *education system* stage. By establishing new structures and organizational practices to directly influence education work, district personnel became directly involved with instruction to meet or respond to federal and state policy goals, including implementing academic content standards and accountability policies for student outcomes in mathematics and English language arts (e.g., Marsh, 2002; Massell & Goertz, 2002; Snipes et al., 2002). The implications of this type of organizational structure, with its intention to improve educational practices, opened the door to the testing-focused quandary discussed next.

A Culture of Testing

Despite NCLB’s well-intentioned efforts to focus on the performance of minoritized student groups, many authors point to how district and school practices and policies shaped by state testing and accountability largely exacerbated rather than mitigated inequalities (e.g., Au, 2007; Blazar & Pollard, 2017; Hamilton et al., 2008). These practices included restructuring or narrowing curricula to focus instructional activities on content for state tests, adjustments to programming and scheduling (e.g., removing art classes from the school’s program or reducing or removing recess time) to better prepare students for the tests, and using highly scripted curricula and strict pacing to help improve test scores (Au, 2007; Blazar & Pollard, 2017; Crocco & Costigan, 2007; Dresser, 2012; Duncan-Owens, 2009; Hamilton et al., 2008; Heiser et al., 2015).

NCLB also motivated districts to make large investments in commercially developed interim assessments designed to efficiently monitor student learning and collect predictive information about student performance on high-stakes summative state tests (Shepard, 2017). Despite persistent calls from multiple stakeholders for reduced state testing (Olson & Jerald, 2020), at present, many school districts continue to administer a large array of tests to students, as highlighted in Figure 6-1.

The areas in Figure 6-1—showing a Grade 11 calendar during the 2020–2021 school year—highlighted in yellow mark time in which different groups of students are scheduled to be taken out of their classrooms to participate in formal district interim testing or summative state testing. Most days of most months are earmarked for some sort of testing, underscoring how testing continues to shape scheduling and behaviors at many U.S. districts and schools.

During the NCLB period (2001–2015), many educational researchers documented how learning experiences for students of color and other minoritized groups were negatively impacted by testing and accountability (Blazar & Pollard, 2017; Heiser et al., 2015; Teoh et al., 2014). Today, researchers still note the persistence of such practices within a testing culture and have documented how they continue to harm minoritized groups (Gitomer & Iwatani, 2022; Randall et al., 2022). The call Marion and colleagues



FIGURE 6-1 Testing calendar from an urban high school.
 NOTE: This photo was shared with the authors of this chapter on the condition of not revealing the school or district location.

make in Chapter 1 of this volume, “Reimagining Balanced Assessment Systems: An Introduction,” to re-center the focus of assessments on classroom assessment, not only represents a demand to reprioritize the types of assessments used by districts and schools, but is a call to action for states, districts, and schools to ensure that this shift affords greater equity and fairness in the educational opportunities offered to students.

ASSESSMENTS TO SUPPORT AMBITIOUS TEACHING AND LEARNING

Because the learning-centered vision articulated in this volume is informed by sociocultural learning theory (see Chapter 3 of this volume, “Human Learning and Development: Theoretical Perspectives to Inform Assessment Systems”) and plays an integral part in ongoing teaching and learning activities (see Chapter 1 of this volume, “Reimagining Balanced Assessment Systems: An Introduction”), district and school leaders will want to prioritize classroom assessments that are situated in *classroom activity systems*. Such systems are “largely determined by a teacher’s pedagogical actions [and] provide affordances for participation in a community of practice” (Kang & Furtak, 2021, p. 75). According to Kang and Furtak (2021), a *classroom activity system* can be identified by the set of interactions—or relational work—inspired by activities involving the materials and structures deployed in classrooms (e.g., lesson plans, assessments, instructional tasks, instructional routines, etc.). Kang and Furtak (2021) further note that the degree to which learners participate in these activities depends on “who they are and their historical relationship with the discipline, and actors [i.e., other students, teachers] in classrooms” (p. 75). This implies that when the materials and structures used in a *classroom activity system* are intentionally designed to improve the quality

of relational interactions between participants, the participation of all students in the classroom can be optimized. Since proximal and near classroom assessments represent the major focus of learning-centered assessment systems, forms of assessment that are distal from a *classroom activity system*, such as state tests and other locally required assessments, have more utility for district and school administrators than for educators and students (Ruiz-Primo et al., 2012).

Forms of classroom assessment that support this dynamic and relational *classroom activity system* enable districts to play a highly influential role in supporting ambitious teaching and learning practices. That is, a district can serve as a learning hub for schools by providing school leaders and educators with professional learning and resources to support the selection, development, and use of classroom assessment as an integral part of high-quality curriculum and instruction. Taken together, the formative and summative classroom assessments discussed in Chapter 4 of this volume, “Classroom Activity Systems to Support Ambitious Teaching and Assessment,” form a coherent and relatively comprehensive picture of student learning. Importantly, the specifications for the design of such assessments lean heavily on processes that not only reveal student thinking and reasoning but also help foster a trusting and inclusive learning environment for all.

Such classroom assessments described in Chapter 4 of this volume, “Classroom Activity Systems to Support Ambitious Teaching and Assessment,” are connected to the learning environment because assessments that support rich and culturally responsive instructional and learning opportunities can generate observable classroom practices such as teachers and students co-constructing knowledge as they explore ideas in depth; respectful dialogue that values the ideas of every learner; and teachers avoiding “front-loading” vocabulary in classroom interactions to signal that there is not one “correct” way to use language in a given discipline (Darling-Hammond & Cook-Harvey, 2018; Suárez et al., 2020; Thompson et al., 2021; Windschitl et al., 2012, 2018). These practices indicate that an important feature of a classroom with a culture of ambitious teaching is that it positions students to collaborate with teachers in the learning and assessment experiences enacted in the classroom.

Performance assessments that can be used in multiple ways (e.g., end-of-unit assessments, capstone performance demonstrations, or common district-wide assessments) are a natural fit for this ambitious teaching vision. Such assessments engage students in complex tasks and activities—synthesizing information, evaluating evidence, problem-solving—and also provide relatively accurate markers of the higher-order skills and knowledge students have acquired (Conley, 2015; Darling-Hammond & Adamson, 2014; Faxon-Mills et al., 2013; Hofman et al. 2015; Linn & Burton, 1994). In fact, a wide range of informal and formal formative assessment strategies and processes can be used to elicit student thinking in line with this vision. These include discourse-based strategies employing extended discussions to discover how students are thinking about the topic at hand and then adjusting daily instruction accordingly. Such assignments are instrumental in enacting ambitious teaching. We next walk through an example of the types of formative and summative assessments that can be used in a *classroom activity system* and how the assessments work together to support ambitious teaching and learning.

An Example of Classroom Assessment Supporting Ambitious Teaching

To illustrate how formative and summative classroom assessments can cohere and produce a rich body of information to help inform teaching and learning, we point, as an example, to the curriculum development and assessment work of the Storylines Project, based at Northwestern University. This project seeks to advance the implementation of Next Generation Science Standards (NGSS) in districts and schools by providing high-quality, open-resource materials. The Storylines Project exemplifies a coherent approach for using science assessments in service of curricular and instructional goals since the curriculum materials or units of study developed for different grades has embedded informal and formal assessments that support the learning targets connected to big disciplinary ideas in science, and ultimately support the enactment of NGSS standards.

According to the Storylines project team, each open-source unit with lessons and embedded assessments presents “a coherent sequence of lessons in which each step is driven by students’ questions that arise from their interactions with phenomena” (Next Generation Science Storylines, n.d.). In other words, students serve as key collaborators in the learning process, helping to move the classroom forward by explaining scientific phenomena or solving problems. By positioning students as active participants in their learning, Storylines units attend to key features of sociocultural learning theory and can help educators implement ambitious teaching in their classrooms.

Each Storylines unit developed is intended to elicit the intentional enactment of formative instructional strategies and tasks from teachers and help surface student reasoning and inquiry. Teachers can then use rich summative performance assessments to evaluate student learning at the end of the instructional period or unit of study. Importantly, teachers can use these assessment experiences as part of their lessons. For example, one portion of the set of curricular resources provided for a Grade 4 unit was titled, “*Why do some things wash up on the beach and others don’t?*” Informal checks of collaborative group work produced for one of the seven lessons include having teachers:

Look at student responses in the “Make Predictions” and “Make Plans” sections on page 1 of Student Handout 4.1 to see students’ ideas about how to create the type of waves they need, and how to consistently carry out their plans. Also check page 2 of the handout to see if students are able to accurately record their data from multiple trials. (Aycock et al., 2019)

If group work products indicated that students were struggling with the tasks assigned for this lesson, the unit encourages teachers to draw on formative strategies as follows:

If students are not able to summarize this thinking about how waves move floating objects, have them look back at the “Finding Patterns” section on page 3 of Student Handout 4.1. Ask them what was similar among most or all of the group’s data. If students are struggling to comprehend that the results were not what they predicted, remind them that science is often surprising, and unexpected results help us ask more questions and design better investigations next time. It may also be helpful to remind students that the “How we represent our thinking” section of their Progress Tracker

can be done with drawings in addition to words - some students may be more able to explain their thinking with a labeled model than complete sentences, and that's okay! (Aycock et al., 2019)

The intention of these informal instructional and assessment strategies is to encourage students to build on prior knowledge to advance their thinking. The strategies also encourage students to consider different modalities for demonstrating knowledge and skills and to take risks with their investigations.

Once teachers reach the end of a Storylines unit, they may opt to administer a summative unit test that allows students to formally demonstrate their learning. Figure 6-2 presents a student's response to a single item on the summative test for the Grade 4 unit described above. This rich item attends to the performance expectations evaluated throughout the unit but also asks students to demonstrate the depth of their learning by transferring the acquired knowledge and skills to an entirely new set of tasks.

1A. What will you do to make the most, tallest waves you can? Use the words disturbance, amplitude, and wavelength to describe how you will push the water to make the most, tallest waves.

To make taller waves, you need to make a ~~big~~ disturbance. Making a faster, bigger disturbance will make high ~~amplitude~~, and short wave length, that will make a lot of high waves to hit your friend.

1B. Draw a model of what your waves would look like in the water. Be sure to label your model including the words amplitude, and wavelength.

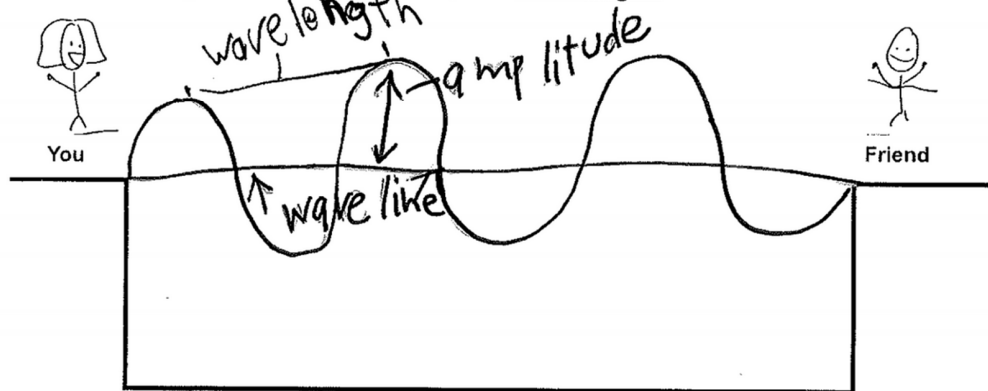


FIGURE 6-2 Example of a rich curriculum-embedded item on a Grade 4 end-of-unit summative test. SOURCE: Aycock et al. (2019).

Storylines units developed for science educators provide an example of how a district and school can achieve *horizontal coherence*² in assessments supporting one disciplinary area (science) using a variety of informal and formal curriculum-embedded assessments that accompany each lesson. These lessons achieve *horizontal coherence* because the set of assessment experiences and opportunities embedded in each unit address important learning targets linked to the disciplinary “big ideas” established in the curriculum (Shepard et al., 2018).

Taken together, the set of rich formative and summative assessment experiences provided in the curriculum units developed by the Storylines Project give teachers a clear picture of what students know and can do relative to key learning goals deliberately aligned with the performance expectations set by academic standards. The Storylines Project also works with a conception of coherence that requires teachers to factor in student perspectives and agency as key design principles for developing instructional and assessment experiences. The project’s developers believe that learning and assessment can only gain “coherence” when students participate in both activities as co-constructors of knowledge (Reiser et al., 2021). Thus, in the Storylines Project, student perspectives and agency operate as important defining features for establishing coherence in the set of learning and assessment experiences offered under the banner of ambitious teaching and learning.

Distinguishing the Role of Distal Assessments

Distal assessments—assessments that are external to classroom learning activities—can play an important role in helping school and district leaders evaluate broader school-level performance. However, these assessments serve a distinct purpose from classroom assessments. Because federal and state education agencies recognize that districts exercise direct oversight over schools, districts are charged with ensuring that required state and federal assessments are administered to students. This oversight function requires districts to review student results from standardized assessments because such assessments generate comparable information about schools, allowing district and school leaders to identify important student performance trends relative to key school and student initiatives, reforms, and other interventions.

Other locally required assessments, such as interim tests classified as formative classroom assessments by many district and school leaders and educators, would actually be considered distal assessments. Although some disagree with classifying these district-developed or commercially developed interim tests as distal assessments (e.g., Dyer, 2017), these assessments are not designed to support the culturally responsive practices and relevant learning experiences embedded in a *classroom activity system* as described in Chapter 4 of this volume, “Classroom Activity Systems to Support Ambi-

² To clarify this term, we refer back to Chapter 1 of this volume, “Reimagining Balanced Assessment Systems: An Introduction”: “At the classroom level, coherence generally means ensuring that assessments are consistent with high-quality curricula and instructional materials that reflect contemporary understandings of disciplinary learning and knowledge development. *Horizontal coherence* is alignment among curriculum, instruction, and assessment to help students develop proficiency in a content domain” (National Research Council, 2006). . . . *Horizontal coherence* is most critical at the classroom level, especially because formative and other classroom assessments must cohere with ambitious instruction and an equity-centered curriculum. School districts generally have the authority to support *horizontally coherent* systems of assessment since curriculum and other related decisions are generally made at the district level” (p. 5, italic in the original).

tious Instruction and Assessment.” Given that interim tests tend to rely on a selected response format, they provide limited opportunities for making student thinking and reasoning visible. Thus, these tests are not ideal for a classroom assessment system that seeks to deepen teaching and learning (Perie et al., 2009; Shepard, 2019).

Districts will want to clearly communicate to school leaders that distal assessments are not part of a thriving *classroom activity system*. While the results of distal assessments are essential for broader program evaluation and district-wide monitoring of schools, they should not drive classroom instruction. Ensuring that districts message these priorities to schools is consistent with state practices that take up two of the high-leverage state actions highlighted in Chapter 7 of this volume, “State Practices and Balanced Assessment Systems,” clearly communicating the role of state summative assessments and mitigating their misuse and the misuse of other locally required tests. In the next section, we address the infrastructure practices districts and schools should adapt to support and sustain balanced assessment systems.

Establishing School Partnerships

An important first step for districts seeking to encourage ambitious teaching is partnering with schools to strengthen investment in this vision of teaching and learning. As mentioned in the previous section, this type of collaboration includes clearly communicating which assessments should be prioritized and why. Recognizing that some schools (e.g., charter and innovation-zone schools) are autonomous from their districts and can shape their own educational goals and vision and that some districts have adopted decentralized structures (i.e., more authority and resources shift to schools), this chapter addresses a common scenario found in many states where districts guide and motivate the operations and performance of schools—district leadership setting the vision for teaching and learning and central offices expected to coordinate with schools to implement this vision.

Even under this scenario, schools do not necessarily follow the district’s strategic direction and vision, and some schools face difficulties enacting desired reforms. Factors that can engender difficulties include the size of the school district; school-based leaders and personnel who misinterpret the vision and accompanying policies; internal conflicts within the district; and the extent to which the district’s existing organizational structures, policies, and norms obstruct school-based initiatives (Massell & Goertz, 2002; Snipes et al., 2002; Togneri & Anderson, 2003). Literature on school turnaround is rife with case studies that document the failure of reforms to take root in schools, particularly when districts enforce top-down implementation (Meyers, 2020).

The idea that schools can effectively improve teaching and learning on their own, without resources and guidance from their districts, their state, or both, is not supported by the available evidence (Honig & Rainey, 2015; McLaughlin & Talbert, 2002; Polikoff, 2021). Even so, most researchers recognize that districts need to provide schools with the flexibility to take up proposed reforms in ways that best fit their needs and contexts (Elmore & Burney, 1997; Marsh, 2002; Massell & Goertz, 2002; McLaughlin & Talbert, 2002; Meyers, 2020; Togneri & Anderson, 2003). That is, to have school personnel buy into a district’s vision, school leaders must have the flexibility to make decisions about how to engage in proposed changes since schools have different levels of readiness to

take up complex reforms. At the district level, flexibility entails paying attention to the localized context of each school and the communities it serves to determine what level of support and resources are appropriate on a site-by-site basis (Massell & Goertz, 2002; Meyers, 2020).

In addition to exercising flexibility, districts can establish reciprocal relationships that benefit both districts and schools. To do this, they can look to strong partnership models such as the research–practitioner partnership model (Coburn et al., 2013b). Coburn et al. (2013b) define the research–practitioner partnership model as a long-term collaborative relationship established between researchers and practitioners to attend to persistent issues or problems. By encouraging schools to take up productive adaptations of the support and resources provided by central offices, districts can enable schools to identify and implement actions or steps that would best facilitate their adoption of the district’s teaching and learning vision. This type of partnership work will likely promote schools’ sustained support of the district’s vision. We now turn to a discussion of *instructional infrastructure* in considering how ambitious teaching can be supported through partnerships between districts and schools.

ESTABLISHING AN INSTRUCTIONAL INFRASTRUCTURE FOR AMBITIOUS TEACHING

In this section, we outline examples of district and school practices and policies to support three critical components of *instructional infrastructure*—high-quality curricula, professional development, and grading—that can catalyze both district and school efforts to recenter their focus on classroom assessments. Ideally, the state would partner with school districts to support classroom assessment, in a similar way that we would envision districts to partner with their schools (see Chapter 7 of this volume, “State Practices and Balanced Assessment Systems”). That is, districts can sustain their work to provide high-quality curricula and support high-quality professional learning opportunities at schools through the funding, resources, and other support they receive from the state. This partnership model would also be taken up by schools, so that school personnel could partner with parents, students, and other community members to improve the school’s *instructional infrastructure*.

Districts and schools seeking to implement this learning-centered vision are likely functioning as *learning systems* (see Chapter 8 of this volume, “Developing, Implementing, and Institutionalizing Complex Educational Innovations: Considerations for Balanced Assessment Systems”). As indicated in the introduction to this chapter, districts located in the *learning system* phase actively seek broad stakeholder input when embarking on and learning from this transformational work. Leaders at the district and school level whose institutions are at the *learning system* stage typically engage in interactive and relational practices with stakeholders to facilitate trust and buy-in from their stakeholders to deepen reforms by learning from stakeholder experiences and feedback. By learning from and through leaders’ multiple layered interactions with their broader network of stakeholders, these leaders can target their reform efforts. A key part of this broader relational work is to codesign the infrastructure—policies, structures, and practices—with the stakeholders who will support this complex instructional and assessment work in schools.

Instructional infrastructure, or what Penuel (2019) refers to as “infrastructuring,” includes components that contribute to the successful adoption of educational reforms, including how assessments are used. Specifically, *instructional infrastructure* includes the components aimed at shifting instructional practices and the set of interactions that occur within this infrastructure (Cohen et al., 2013; Mehta & Fine, 2015; Spillane et al., 2011). According to Hopkins and Spillane (2015), *instructional infrastructure* “forms a system intended, by design or default, to guide and monitor instruction and its improvement” (p. 422). Components within this system include but are not limited to, professional learning, assessments, instructional materials, instructional frameworks, school and district-level policies, roles, and positions focused on instructional support, programming, and oversight (Cohen et al., 2013; Hopkins & Spillane, 2015; Spillane et al., 2011). In theory, these components should work together to foster rich interactions in the *classroom activity system*.

If the components of an *instructional infrastructure* cohere to support a *classroom activity system* with ambitious teaching and learning as the organizing design, there is promise that the infrastructure will lead to desired outcomes like establishing an equity- and learning-centered environment for all students (Bryk et al., 2009). However, building such infrastructure will only gain traction in schools if district personnel engage with school leaders, educators, and other important stakeholders (e.g., community members) in a meaningful partnership.

Our discussion of *instructional infrastructure* begins with practices and policies that support the development of a high-quality curriculum since the presence of such a curriculum is a critical lever for “establishing coherent, consistent high-quality instruction in ... schools” (Polikoff, 2021, p. 103). In other words, ambitious teaching and learning—including classroom assessment practices—cannot happen unless schools have access to a high-quality curriculum.

Next, we shift the discussion to professional learning, as practices included in this component of *instructional infrastructure* enable schools and teachers to enact the curriculum and the assessments that support this infrastructure. We then address grading, as this component provides teachers and schools with an additional avenue for providing feedback that has the potential to improve learning. We do not define instruction as a separate component of *instructional infrastructure* because the above components combine to directly support instructional routines, including assessment. We acknowledge that additional infrastructural components could be examined—such as talent, career development, or teacher evaluations—but have limited the discussion to these three areas because they are critical levers for creating assessments that support a *classroom activity system* focused on ambitious teaching and learning.

Access to High-Quality Curriculum

We open this section with an illustrative vignette sourced from personal conversations with Peter Leonard about the current work underway at Chicago Public Schools (CPS) to advance high-quality curriculum. This vignette (see Box 6-1) is intended to illustrate how a large school district has made significant investments in an *instructional infrastructure* to help spread this work in schools.

BOX 6-1
Chicago Public Schools: Advancing High-Quality Curriculum

In 2020, the Chicago Public Schools (CPS) embarked on a reform to provide a high-quality curriculum (Skyline curricula) to all schools. This curriculum was a key component of their vision to advance student equity. The district made the decision to prioritize the provision of this curriculum and accompanying high-quality instructional materials and resources to all schools in order to ensure that all students, regardless of the school they attended, would have access to engaging and rigorous lessons and instructional materials. CPS also moved in this direction to signal a sweeping change to their vast network of schools: the district was turning away from using tests as the primary means for organizing instructional priorities and evaluating students, and would instead focus on improving the quality of teaching and learning provided to them. This meant that the district's assessment focus would prioritize curriculum-embedded classroom assessment that supported teachers in enacting the district's high-quality curriculum.

To engage in this work, the district established a broad group of curriculum experts—both within and external to CPS. Their task was to clearly define curricula in all content areas and grades, to select high-quality materials, and to work with curriculum partners to ensure that the selected materials were culturally responsive and relevant to Chicago students. In addition to defining the curriculum and materials, the district hired curriculum specialists in every disciplinary area and at every grade level to serve as professional learning partners for all schools throughout the district. Knowing that this initiative would not gain traction if schools were mandated to adopt the curriculum and materials for each disciplinary area, CPS showcased this work as a model that schools could choose to follow or to directly adopt (or not). The district's hope was that schools would see the quality of the investments it had made in this instructional infrastructure, and would be motivated to shift toward adopting the curriculum and the extensive resources the district had provided. Although the district is still in the early stages of elevating this instructional infrastructure, it has expended vast resources to implement it. Presently, 470 schools in the district use Skyline curricula in at least one grade band and content area. The support that central administrators and staff at CPS have given to this large endeavor communicates a unified front: that every stakeholder in the system prioritizes giving students opportunities to learn and benefit from a high-quality curriculum.

A high-quality curriculum—one that includes instructional frameworks, curriculum maps, instructional materials (inclusive of classroom assessment), and programming decisions—is the most critical component in a district or school's *instructional infrastructure*. This is because curriculum determines both the materials and resources that will directly support instructional activities, as well as the progression or sequencing of disciplinary content and skills teachers will use. The curriculum also determines when and which assessments should be used and how to organize the school schedule to maximize instructional time. In the following subsections, we focus on two specific curricular components needed to support an ambitious teaching vision: curriculum coherence and curriculum materials.

Curriculum Coherence

“In the absence of a learning plan with clear goals, how likely is it that students will develop shared understandings on which future lessons might build?” (Wiggins

& McTighe, 2005, p. 21). This question gets to the heart of why instructional frameworks and curricular maps are critical tools for communicating instructional guidance and identifying important areas of assessments. Because these frameworks and maps provide teachers with guidance and content for instruction, they can also facilitate coherence in the curriculum across and within grades by articulating expectations for each disciplinary area of the *instructional infrastructure* (Cohen et al., 2013; Hopkins & Spillane, 2015). As the authors of Chapter 4 of this volume, “Classroom Activity Systems to Support Ambitious Teaching and Assessment,” discuss, districts and schools can gain a clear understanding of the evidence needed to evaluate learning relative to established learning goals if they have a clear roadmap that outlines expectations for what should be taught in each disciplinary area. This understanding, in turn, should lead districts and schools to intentionally design or select assessments that match the learning goals, such as those used in the Storylines Project units.

Instructional frameworks and curriculum maps can help teachers specify assessment tasks and clarify which instructional moves should follow those tasks. When anchored to ambitious teaching as an organizational design, these frameworks and maps should also encourage teachers to attend to the developmental needs of students while simultaneously providing lessons that feature “well-scaffolded instruction and ongoing formative assessment(s) that support conceptual understanding, take students’ prior knowledge and experiences into account, and provide the right amount of challenge and support on relevant and engaging learning tasks” (Darling-Hammond et al., 2020, p. 98). As is noted in Chapter 4 of this volume, “Classroom Activity Systems to Support Ambitious Teaching and Assessment,” curricula designed for ambitious teaching not only respond to but also sustain the knowledge, practices, cultures, and languages of learners.

Teachers can also reference instructional frameworks and curriculum maps as guides in the development of instructional materials—such as units of study—and the accompanying set of curriculum-embedded assessments used to evaluate student performance. When teachers closely align their instructional and assessment practices to an instructional framework or curriculum map, this can be promising to establish more equitable access to high-quality learning opportunities for all students, regardless of the school they attend.

An example of an instructional framework designed to connect with ambitious teaching and learning is Schoenfeld’s Teaching for Robust Understanding (TRU) framework for math (Schoenfeld, 2013, 2017). The TRU framework consists of five dimensions, each of which focuses on what students are expected to do in math within the context of the learning activities enacted by teachers. These five dimensions are described in Figure 6-3.

Burkhardt and Schoenfeld (2019) explain how the five dimensions of the TRU framework can work together as principles for designing instruction connected to a well-specified sequence of learning activities and tasks that utilize formative assessment strategies. For example, Burkhardt and Schoenfeld (2019) note that, by using the TRU framework, teachers can design lessons that “uncover students’ existing ways of thinking, then create cognitive conflicts or disturbances that lead students to realize and confront inconsistencies ... through student-student and student-teacher discussion, in pairs or small groups, and then across the class as a whole” (p. 51). Teachers

The Five Dimensions of Powerful Classrooms				
The Content	Cognitive Demand	Equitable Access to Content	Agency, Ownership, and Identity	Formative Assessment
<i>The extent to which classroom activity structures provide opportunities for students to become knowledgeable, flexible, and resourceful disciplinary thinkers. Discussions are focused and coherent, providing opportunities to learn disciplinary ideas, techniques, and perspectives, make connections, and develop productive disciplinary habits of mind.</i>	<i>The extent to which students have opportunities to grapple with and make sense of important disciplinary ideas and their use. Students learn best when they are challenged in ways that provide room and support for growth, with task difficulty ranging from moderate to demanding. The level of challenge should be conducive to what has been called "productive struggle."</i>	<i>The extent to which classroom activity structures invite and support the active engagement of all of the students in the classroom with the core disciplinary content being addressed by the class. Classrooms in which a small number of students get most of the "air time" are not equitable, no matter how rich the content: all students need to be involved in meaningful ways.</i>	<i>The extent to which students are provided opportunities to "walk the walk and talk the talk" – to contribute to conversations about disciplinary ideas, to build on others' ideas and have others build on theirs – in ways that contribute to their development of agency (the willingness to engage), their ownership over the content, and the development of positive identities as thinkers and learners.</i>	<i>The extent to which classroom activities elicit student thinking and subsequent interactions respond to those ideas, building on productive beginnings and addressing emerging misunderstandings. Powerful instruction "meets students where they are" and gives them opportunities to deepen their understandings.</i>

FIGURE 6-3 The Teaching for Robust Understanding framework dimensions.
SOURCE: Teaching for Robust Understanding Framework (n.d.).

using the TRU framework would then be asked to reflect on their formative practices and check whether they support rich mathematical content, achieve high levels of cognitive demand by maintaining productive struggles with content, ensure meaningful engagement for all students, and strengthen opportunities for student sense-making that fosters agency and identity (Burkhardt & Schoenfeld, 2019). Thus, implementing an instructional framework that supports ambitious teaching practices can also directly impact student academic performance. According to the authors, "Classrooms that did well on the rubric [connected to the TRU framework] did well on mathematics [classroom-based] measures ... [whereas] classrooms that scored poorly did not" (Burkhardt & Schoenfeld, 2019, p. 41).

Building on this example, we argue that if districts provided schools with resources to adopt this type of instructional framework for mathematics, it would set the groundwork for curriculum specialists to design and develop instructional and assessment strategies and routines. At the district level, district-based curriculum specialists—who are typically separated from district assessment staff—would be encouraged by leadership to work together to create educative resources and models that schools could directly adopt or reference in their selection of frameworks and materials. These curriculum specialists would be tasked with building capacity to support professional development work provided by the district and backed by school leadership.

District leaders who support this vision should recognize that offices supporting different functions will have to collaborate closely to establish coherence across the work of personnel who are charged with influencing teaching and learning. For example, in CPS, the curriculum, instruction, and assessment functions all reside within the Office of Teaching and Learning, which creates the structural conditions for closer collaboration between offices at the district level. This type of collaboration is rarely encountered in school districts (Latham, 2018). Nevertheless, the example is an important proof that departmental shifts can happen if they are prioritized by district leadership, and staff are provided the authority and resources to implement organizational change.

Curriculum Materials

High-quality curriculum materials, inclusive of classroom assessments, should ideally reflect the content and activities specified in related instructional frameworks and curricular maps. We define “high-quality materials” as those that embody the learning-centered features described at length in previous chapters of this volume and the broader literature (e.g., Armstrong, 2021; Ladson-Billings, 2014; Reiser et al., 2021; Wang et al., 2021). Primary among these features is that the materials connect to the diverse experiences and interests of students. This feature is important because it has direct implications for designing learning and assessment experiences. Kaufman et al. (2020) note that, despite the importance of high-quality curriculum materials and resources, they are often lacking even though a few states have tried to help districts identify and invest in such materials. The disconnect between the high-quality curriculum materials promoted by these states and lower-quality materials can be attributed to the common policy of “local control,” under which curricular decisions are made by the district rather than by the state.

As addressed earlier in this chapter, each district should have a vision for its curriculum, but this vision is not always shared by schools. This is particularly the case when the district is large and oversees several networks of schools, as seen in the CPS vignette. Districts operating at the same scale as CPS understand that their vision for a high-quality curriculum is more likely to gain traction if it is promoted using strategies that build buy-in with schools. One aspect of the CPS strategy was to ensure that school leaders and educators were involved in the process of piloting and refining these high-quality curriculum models and resources so that they would be motivated to adopt these materials and resources. According to a CPS leader, this included working with school leaders and educators to formalize a clear definition of high-quality curriculum, evaluating current curricula along that quality definition, performing a non-evaluative curriculum audit in collaboration with schools and networks, and setting a multi-year goal toward high-quality curricula that empowers school leaders to lead that process in their buildings (P. Leonard, personal communication, 2023).

Another important strategy used by CPS to facilitate the adoption of high-quality materials in schools was to ensure that all educators had access to curriculum-specific professional learning from district-based content experts in order to effectively use these resources. These content experts, located in the district’s curriculum office, provided school-based staff with resources, learning opportunities, and additional support for implementing new curricular resources. Deploying strategies like those used by

CPS can lower some of the anticipated barriers and anxieties that can surface in schools when reforms, including new curricular and assessment directions, are implemented.

Polikoff (2021) recommends that states recruit teachers to participate in curriculum reviews to help and endorse high quality curriculum materials, as this is likely to improve buy-in for teachers to use those materials in the classroom. This strategy can be used by districts and schools that are in the process of adopting new curriculum materials and can also include the broader community, which would help ensure that the materials are responsive to the diverse backgrounds and interests of the local community. Another strategy is to have teachers participate in a curriculum adaptation of materials to better align existing materials with ambitious teaching and learning goals (Cook-Endres et al., 2014). Allowing teachers to engage in this type of adaptation work by reworking existing materials may be a more viable strategy in some schools and districts, particularly if a site does not have sufficient resources to supplant existing materials. This strategy would enable districts and schools to work with educators and modify selected units, tasks, and assessments to better align with learning-centered approaches and instructional models.

As discussed earlier in this chapter, districts can offer flexibility to schools in selecting instructional materials, especially when schools favor distinct programmatic and focal areas like expeditionary learning or a STEAM model (in which science, technology, engineering, and math are the focus, with the arts infused). However, flexibility will need to be balanced with the assurance that selected materials align with the high-quality criteria endorsed by the district in consultation with schools and other stakeholders (e.g., community-based groups). These criteria are critical for setting clear expectations about the features the instructional materials should embody—including embedded assessments—when individual schools are allowed to consider the materials they wish to select or adopt.

Professional Learning

We begin this section with a vignette as an illustrative example of one school and district collaboration to support professional learning intended to elevate ambitious teaching practices. This vignette (see Box 6-2) summarizes key highlights from an existing report (Diaz-Bilello et al., 2022).

In the prior section, we touched on aspects of curriculum—coherence and materials—that provide the necessary support for important classroom-based instructional and assessment activities. Here, we note that these resources require instructional coaches, leaders, and teachers who have the knowledge and capacity to enact and use them in skillful ways. In other words, ensuring that all schools can access a high-quality curriculum is not enough—this work must be accompanied by personnel capacity building. To advance equity and create learner-centered instruction and assessment routines, both district and school leaders must provide teachers with professional development that enables them to build their knowledge and skills in the cognitive, social-emotional, and cultural dimensions of learning (see Chapter 5 of this volume, “Assessment Literacy and Professional Learning”).

Unless teachers build a repertoire of pedagogical content, knowledge, and skills to successfully enact a high-quality curriculum—which must include instructional and

BOX 6-2
Prairie Heights Middle School: Collaborative Opportunities in Professional Learning

The Prairie Heights Middle School in Greeley, Colorado, participated in a network of turnaround schools established by the state and supported by their school district. Because the school had already invested in high-quality instructional materials approved by district and state partners, the school's leadership team turned its attention to revamping its professional learning structures and practices to improve their instruction and assessment practices. An important aspect of this restructuring work entailed ensuring that professional learning offered collaborative opportunities for teachers to try out and learn from the important formative instructional strategies used to engage students in their learning.

At Prairie Heights, all teachers review student work together. Every week in both grade-level and disciplinary-specific teams, they discuss instructional strategies and consider the home-life and personal circumstances of students. Coaching cycles were established so that school leaders and teacher leaders could mentor and provide feedback to novice teachers or teachers new to the school. The cycles were an opportunity for less experienced teachers to learn how to provide meaningful and actionable feedback to students. When teachers within these professional learning communities (PLCs) discuss the personal experiences of students, they use the understandings they have gained from having engaged with students and their families. This knowledge enables them to consider the issues and contexts that influence the learning experiences of students. School administrators serve as facilitators of these PLCs, along with other mentor teacher leaders. Together, the PLC and coaching cycles enable teachers to develop and test strategies—including building classroom structures intended to foster student engagement and collaboration—and receive feedback from mentors on how to refine the enacted routines. In the words of one teacher who had joined the school during the difficult turnaround period, “If it wasn't for the administrators and mentor teachers doing feedback loops, sticking with me, showing me what I'm missing, [and] pairing me with other teachers who could model for me, I probably would not have stayed.... I did not leave the school because of all of these supports to make me significantly better.”

assessment materials and resources—take-up of ambitious teaching vision will likely be less effective (Darling-Hammond, 2004; Darling-Hammond et al., 2017; Kang et al., 2016). For this reason, districts should work with school leaders to find creative ways of establishing and planning professional learning opportunities as a regular, ongoing, and frequent part of the school week (Darling-Hammond et al., 2017; Penuel et al., 2017, 2020a). For example, the Boulder Valley School District (BVSD) in Colorado recently enacted a policy of allowing their schools to set a late-start day once per week to allow the morning to be used for professional development. In an interview, a BVSD principal indicated that school leaders were “so grateful for this policy because [prior to this policy], we had so little time during the school year to spare on professional development” (S. Minnich, personal communication, April 25, 2023).

Districts that support this important infrastructure work can be highly influential in providing a policy framework and establishing clear expectations for school leaders in cultivating strong professional learning communities. In the opening vignette for this section, the Greeley district did just this in supporting Prairie Heights Middle School's decision to implement weekly horizontal and vertical team meetings for all teachers.

Even though each school ultimately defines the structure and culture of the professional learning it enacts for and with teachers, Kraft et al. (2021) note that the sense of success teachers have as professionals can be bolstered when they feel that they can depend on their district and school leaders to clearly communicate instructional priorities and provide them with targeted and relevant training.

Learning Communities and Distributed Teams

Effective professional learning occurs in the context of a learning community—or culture—in which teachers engage in collective sense-making and knowledge construction (Coburn et al., 2013a; Fullan & Quinn, 2015; Hargreaves, 2000; Watson, 2014). Under the vision of ambitious teaching and learning outlined throughout this volume, such a learning community would ideally regularly convene to examine and discuss student work produced through rich classroom assessments and tasks and identify the instructional and assessment work needed to challenge, scaffold, and improve student learning. Districts can play an influential role in creating such learning communities, primarily by encouraging schools to build distributed leadership teams that empower teacher leaders to take on important mentorship roles. Central office staff can also provide schools with guidance, tools, and resources to support these professional learning communities. For example, furnish school leaders with protocols and tools to guide the analysis of student work produced in response to instructional tasks and assessments, and capture outcomes from these discussions. The district can also help model assessment focused work in professional learning communities by organizing the analysis of student work across educators located in different schools. One model to consider for facilitating the assessment focused work of professional learning communities is the twice-per-year student work analysis event that takes place across participating schools in the New York Performance Standards Consortium (Willis et al., 2022). The Consortium, which focuses on supporting curriculum-embedded, performance-based assessments, uses student work analysis to calibrate expectations and refine both the tasks and scoring rubrics used across participating schools.

However, it is important to note that the reach of the district is limited within schools and the efficacy of these structures depends on how school leaders define and organize their distributed leadership teams. When teacher leaders are provided with “resources, support, structure, and the [formal] authority” by school leaders, they can help drive school improvement efforts (Supovitz, 2018). Distributed teams can inform the selections of school curricula and/or assessments in their districts and provide effective coaching and expert support to challenge and transform common classroom practices that run counter to the vision of teaching, learning, and assessment their districts propose (Darling-Hammond et al., 2017; Wang et al., 2021). Even so, if the distributed leadership team structure lacks formal authority, then these teams are less likely to improve instruction for students (Supovitz, 2018).

Sustaining Professional Development Through Active Learning

Districts can deliver both the policy environment and the resources schools need to ensure that their teachers can analyze and discuss student work using rich assessments

within their learning communities. However, whether and how these professional learning opportunities are effectively adopted depends on the support and mentorship that school administration and leadership teams provide to teachers. Districts should know if school-based professional learning opportunities have been designed to allow teachers to use the instructional and assessment strategies and knowledge discussed in this section to engage in important sense-making and set goals for their learning (Coburn et al., 2013a). If these key features are missing, district administrators can partner with school leaders to strengthen their capacity to build such practices into professional learning. School leaders who shape professional learning within the context of the improvement sciences (e.g., Plan-Do-Study-Act or coaching cycles) are likely to emphasize the importance of active learning, or having teachers reflect on and refine their enactment efforts continuously (Darling-Hammond et al., 2017; Hanno, 2022; Penuel et al., 2020a). This active learning approach serves as another important characteristic of a *learning systems* district or school, where the professional learning work becomes highly adaptative, dynamic, and responsive to the varying and changing needs of students.

The vignette about Prairie Heights Middle School that opens this section describes active learning overseen by school leaders and supported by district administrators. At that school, educators collaborate in disciplinary teams to design strategies they can apply in their classrooms. Then, the teachers receive input from their mentors to help transform their teaching practice. In each subsequent meeting with their disciplinary teams and mentors, Prairie Heights Middle School teachers engage in reflection and inquiry as they consider how to refine their teaching strategies, based on their analysis of student work and their interactions with students. Another facet of active learning that is, as Darling-Hammond (2020) notes, crucial for teachers who aim to understand the development and learning of students, is to reflect on how students' personal lives and circumstances interact with and can directly influence the ways in which students learn. Paying close attention to how students' personal lives and circumstances interact with learning can provide critical information about other types of supports (e.g., social-emotional learning supports) that should be built into classroom activities to better attend to the learning needs of students (Darling-Hammond, 2020).

Active learning serves as an important characteristic for effective professional development because teachers can directly relate to, apply, and continuously refine what they learn to activities that take place in their classrooms (Darling-Hammond, 2020; Hanno, 2020; Penuel et al., 2020a). As they continually reflect upon and evaluate their enacted practices within an active learning professional context, an active learning approach provides teachers with the flexibility to embark on a "change sequence" or the ability to continuously adapt and rework planned instruction, which serves as an important marker to gauge whether they are engaging in equitable instruction (Clarke & Hollingsworth, 2002). Equitable instruction, in this case, means the application and modification of strategies that teachers continuously adapt and revise to optimize student learning.

This type of active learning professional development approach also encourages school leadership teams to consider what additional resources are needed to support the coaching cycles established by schools so that, following each cycle, teachers can see the direct benefits and applications of their new knowledge and skills. When active learning is used in coordination with coaching and mentoring cycles, the activities taken

up in these learning communities maintain coherence—teachers revisit topics and build on the knowledge they have gained in each subsequent coaching cycle. This, in turn, helps ensure that what teachers learn becomes integrated into their regular teaching routines, rather than fading over time, as has been documented in several professional development studies (Boston & Smith, 2009; Hanno, 2022).

Grading

In this section, we review the third critical area of *instructional infrastructure* that is needed to support instruction and assessment: grading. Although the signals associated with grading typically accompany ongoing instruction and assessment, grading practices will need to be examined to ensure that they support rather than work against ambitious teaching. We begin with another vignette to illustrate how one district engaged parents, students, and educators to transform a grading system that can better support student learning. This vignette (see Box 6-3) was assembled from conversations with Lori Cooper from Fountain Fort Carson School District who provided information about this ongoing work to shift grading practices.

We discuss grading as an important component of *instructional infrastructure* as it is tied to a set of district- and school-defined policies and practices that lie outside those governing the scoring of classroom-based and distal assessments. Feldman (2023) states that grades serve as a powerful means for communicating student progress and also represent one of the few areas of teaching that falls directly within an individual teacher's control. As a result, grading practices can vary widely among teachers, even within the same school, and often vary across schools. In some situations, grading practices can create environments that decrease student motivation to learn. Within the framework of an *instructional infrastructure*, pursuing equitable grading policies and guidance for implementation provides an opportunity for districts and schools to support and align grading practices with the values that undergird how to provide feedback to students in a learning-centered vision.

Changing the way grading takes place in the classroom requires districts to provide leaders with guidance and flexibility so that they can determine the best way to shift grading practices among their teachers. Additionally, given both the role of grades as a monitoring device for many parents and how school structures have been set up in support of this purpose, challenges and potential pushback from some parents may be unavoidable—which means that clear communication is necessary, as is inviting parents into the conversation to provide them with the reasoning behind shifts in grading practice.

Many scholars have pointed to the inherent challenges in shifting traditional grading practices at schools, especially as these challenges relate to teacher attitudes and perceptions about grading (Guskey, 2021; McMillan, 2001; McMillan et al., 2002). The FFC vignette that begins this section shows a district that is intentional about bringing teachers into conversations about grading practices. By inviting teachers to discuss their thoughts and criticisms about existing practices, the district provided them with an opportunity to reflect and begin correcting some of the more problematic practices deployed in their classrooms. This example highlights one way that districts and

BOX 6-3

Fountain Fort Carson School District: Shifting Grading Practices

Fountain Fort Carson School District (FFC), located in Colorado Springs, Colorado, recently adopted standards-based grading practices. As the district began its transition to a more equitable grading practice, it determined that a philosophical shift around grading would need to occur in order to provide a rationale for the practices that would follow. The district revisited existing and common grading practices such as: using averages to determine final grades, including student behavior as a part of grades, and using grades as punishment. One of the first steps it took was to generate a buy-in among high school teachers for changing grading practices. Thus, FFC encouraged teachers to submit lists that described existing problematic grading practices at the district's two high schools. After a review of the input provided, school leadership and staff identified the top ten practices teachers considered to be exacerbating the inequitable evaluations of students. These practices included providing extra credit, using grades as a punitive or disciplinary tool, assigning zero grades to missing or incomplete homework, and factoring attendance into grades.

The next step the district took, in partnership with the two high schools, was to bring parents into the conversation. To gather input directly from parents about the educational experiences offered to their children, the high schools established a structure they refer to as Learning Walks. Learning Walks are an opportunity for parents to tour the school with the goal of "lifting the curtain" to showcase instructional practices and to hear from students about their learning experiences. In a panel with students during one Learning Walk, parents engaged with and heard directly from students about how the newly instituted grading practices were impacting their learning. This, in turn, helped bolster parent support for these practices.

Before instituting the new grading practices, school leaders, with district support, determined that the practices would not be implemented as a top-down mandate. School leadership made this decision because they wanted to enact sustainable and incremental changes that would intrinsically motivate their teachers to adopt the new practices. To this end, teachers were encouraged to try out the new practices, to learn by doing, and to ask students for their feedback along the way. As a result, the high schools have increased equitable grading practices in their district.

Presently, an increased number of teachers across the district use a decaying average grading system that emphasizes a student's current performance (or the most recent evidence of this performance) to determine their final grade. Homework and quizzes are treated as (ungraded) practice that provides students with opportunities to receive feedback and correct their errors and misunderstandings. Additionally, teachers no longer assign zeros for missing work. The district implemented a 0–4 scale with clear proficiency criteria for each level, and teachers grade students exclusively on their mastery of content while providing students with separate opportunities to self-assess on essential skills such as collaboration, critical thinking, and communication. As the principal of one FFC high school stated: students have voiced their support for this system in that they now experience more hope in their learning compared to in the past. One district administrator recalled a conversation with a student who said, "This is the first time I've ever had hope in my math class." The administrator explained that the student now felt motivated because they had "never passed with more than a D."

schools can begin shifting teacher mindsets about grading. Educational systems that seek to build deeper learning opportunities for all students will need to consider how to best evaluate student knowledge and competencies to meet learning-centered goals. While there are many ways to evaluate student performance, each method carries its own set of tradeoffs—students will benefit from clearly defined success criteria that also help to maintain fairness and accuracy in the evaluation of all students (Berns, 2015; Nieto, 2013).

Shifting Grading Practices to Support Ambitious Teaching

Grading practices that support a vision of ambitious teaching and learning communicate a clear understanding of the skills and knowledge being assessed at any given time. One model for establishing clearer grading descriptors and criteria, which is employed by many districts and schools, is standards-based—or competency-based—grading. This approach evaluates student mastery of skills or knowledge relative to specific competencies or state standards (Brookhart, 2013a, 2013b). Researchers view this approach as a relatively fair and accurate means for grading student performance because the rubrics used are intended to provide an abundance of relevant and meaningful information that students can use to improve (Feldman, 2023; Lewis, 2020; Muñoz & Guskey, 2015). This focus on mastery is also intended to emphasize deep content learning, particularly when the learning targets and success criteria are clearly defined (Lewis, 2020).

Feldman (2023) argues that standards-based or competency-based grading approaches should provide all students with grades that are fair and meaningful, regardless of any one student’s personal context or learning needs. Feldman (2023) advises that “equitable grading is accurate, bias-resistant, and motivational” (p. 77). These pillars highlight an important guide for grading policies implemented in schools: grading practices should be an accurate reflection of what students know, and should not factor in criteria such as attendance, behavior, or completed homework. In practice, this would mean that teachers ensure that grades reflect criteria that are fair and transparent to all students and are weighted to reflect a student’s most recent performance as opposed to averaging their performance over time (Feldman, 2023). Within this model, teachers provide students with rubrics that contain clear descriptors, or they co-create rubric criteria with their students, allowing students to demonstrate their knowledge through various formats and modalities over time, and provide opportunities for revisions and improvement (Brookhart, 2013b; Feldman, 2023; Nieto, 2013).

In addition to shifting toward clear, qualitatively established, criteria-based grading that has been developed in consultation with teachers, district and school leaders may need to revisit established policies for entering grades into learning management or information systems. In the specific case of FFC, teachers recognized that they were grading students on assignments that were intended to encourage risk-taking and learning from mistakes. Grading those assignments undermined this intention since grading elevated the stakes associated with submitted work. Revisiting the types of assignments and products graded and entered into the learning management system is critical for ensuring that the appropriate form of assessment aligns with and supports the underlying learning-centered values of ambitious teaching.

Transparency for Grading Practices and Policy Shifts

Along with establishing guidance on the use of equitable grading practices within their schools, a *learning system* district or school should engage the school community in policy decisions that would shift grading practices. In the same way that teachers may be resistant to enacting new grading approaches, students and families need adequate time to both fully understand and adjust to these new practices (Hany et al., 2016; Townsley, 2019). Schools that provide students and parents with opportunities to engage with these changes can increase school and family collaboration, engagement, and trust in their implementation (Henderson & Mapp, 2002; Stosich & Bae, 2018). The FFC vignette that begins this section shows one way to do this: the Learning Walks provided families with an opportunity to gather information about changes to grading practices and to see them in action. FFC also provided families with an opportunity to hear directly from students about the positive impact of these changes. Attempts to change grading practices are often unsuccessful because district or school leaders tend to enact them before being transparent about the rationale for the changes being made (Guskey, 2021). In addition, Guskey (2021) notes, leaders may not allow for adequate discussion on how the new practices align with other parts of the educational process (i.e., curriculum, instruction, assessment). While such clarification may not avoid all potential opposition, Guskey (2021) argues that communicating and soliciting feedback from stakeholders can help leaders address the challenges that may lie ahead “with patience, purpose, and resolve” (p. 196).

In summary, fair and equitable grading policies and practices complement work in other parts of the *instructional infrastructure* that supports ambitious teaching. If grading practices are not considered when adjusting the *instructional infrastructure*, existing practices could potentially work against efforts to promote equitable assessment practices. Steps that can be taken to promote equitable grading practices include:

- Avoiding grading methods that demotivate students and work against ambitious teaching and learning practices (e.g., assigning zeros to unsubmitted work, factoring participation and attendance into grades, using 100-point grading scales, and grading non-academic or soft skills such as collaboration).
- Re-evaluating the body of evidence used for grading and ensuring that instructional tasks and assignments used for formative purposes are not factored into grades.
- Providing students with multiple opportunities to demonstrate their mastery of a skill.
- Developing clear criteria in grading rubrics and ensuring that they connect to expected learning targets and goals. Ideally, criteria should be co-constructed with students to ensure a shared understanding.

District and school leaders who provide policy guidance to teachers can help establish consistency in grading practices by creating transparency into these practices and by encouraging the creation of learning environments in which expectations of high-quality teaching and learning are present for both teachers and students (Brookhart, 2013a; Diefes-Dux, 2018; Feldman, 2023).

TAKING A *LEARNING SYSTEMS* APPROACH TO IMPLEMENTING A LEARNING-CENTERED VISION

Enacting the practices and policies described in this chapter requires significant district investment, time, and resources. The transition, for districts, into the *learning systems* stage, focused on advancing ambitious teaching and learning, is a journey. This journey can be daunting, especially when considering the amount of coordination and support work that is required across classrooms and schools. For schools taking up this vision for the first time, the journey can also feel overwhelming, as it requires attending to many areas, including revisiting the school schedule to make it possible for teachers to engage in meaningful professional learning and collaboration.

For districts and schools moving in the *learning systems* direction, the work might begin gradually, in one disciplinary area, one or two grades, just a few classrooms, or a few schools. Pursuing incremental changes will be more manageable for all involved, and will likely help avoid “shallow” results, which are more likely if a transition is implemented at scale and all at once (Cook-Endres et al., 2014). Districts leading the work to implement infrastructure changes may also choose to implement a manageable, small pilot in a single school, or for a single grade level or content area, to learn deeply from this work. Indeed, by using effective partnership models, districts and schools will also facilitate or improve their ability to learn from and improve on this work together at a manageable scale, before considering and potentially re-evaluating steps needed to spread the work gradually and incrementally into other grades, content areas, and schools (see Chapter 8 of this volume, “Developing, Implementing, and Institutionalizing Complex Educational Innovations: Considerations for Balanced Assessment Systems”). Additionally, including the voices of community members in these decisions should also broaden ownership and sustainability in these types of reform (Arriaza, 2004; Penuel et al., 2020b).

For districts and schools wanting to move into a *learning system* phase, it is critical to consider what to prioritize or refine in the existing infrastructure to support ambitious teaching and learning. Prioritization is important because districts and schools face resource constraints that limit what work can be taken up at a given time. Addressing just one area of the *instructional infrastructure* (e.g., providing high-quality curriculum materials) requires significant investment and work. For districts seeking to build on existing infrastructure, drawing on improvement science approaches to identify priorities in consultation with school- and community-based stakeholders, including students, will be helpful (Penuel et al., 2020b).

Improvement science approaches, such as design thinking or Plan-Do-Study-Act cycles, integrated into a community-based design are particularly useful for facilitating the adoption or the improvement of reforms enacted. First, these approaches can provide an effective process for districts and schools to engage stakeholders in appraising the existing infrastructure and identifying areas to prioritize. Many schools and districts already use these types of approaches as process evaluations for improving instructional strategies, including assessment. Second, these collaborative approaches provide districts with opportunities to ensure that school-based personnel, students, and the broader community can help codesign or provide substantive input into the areas of infrastructure that require rethinking and additional investment. Lastly, these

approaches provide natural opportunities for districts and schools to experiment and learn from infrastructure adjustments and changes.

There is no “right” or specific pathway to approach the work of becoming a *learning systems* district or school, centering culturally responsive and learning-centered classroom activity systems. However, districts and schools that seek to become *learning systems* can consider engaging in some of the activities below to begin the process of identifying and determining areas to strengthen or transform:

- Develop a theory of action or change as a starting point to help define the implementation work and supports needed to build or bolster one or more components of the *instructional infrastructure*. This development work could include conducting interviews with a broad range of stakeholders to understand the existing barriers, challenges, and opportunities encountered in each infrastructure area.
- Undertake a review and /or audit of curriculum materials, inclusive of assessments, used in schools to determine the extent to which materials can support teachers in enacting ambitious teaching and learning. Results from this review can inform which materials and tasks are good candidates for adaptation and identify those that may need to be replaced.
- Establish research partnerships with organizations, including higher education institutions, to build an evaluation plan for engaging in studies that will enable districts and schools to learn from their continuous improvement work.

In reference to the third bullet, whether a district is starting this work or building on existing infrastructure, it is necessary to establish an evaluation plan for all areas impacted by the change. The plan should focus on building ownership of this work with schools and the broader community, if the efforts are to be sustained and if the district is to maximize opportunities for all students to learn (Arriaza, 2004; Coburn, 2003; Penuel et al., 2020b). This continuous learning will need to be monitored to determine if policies and resources directed at a particular area of the infrastructure are ultimately leading to expected outcomes. Districts may decide that this learning work should focus on monitoring areas such as the maintenance of access to and provision of high-quality curriculum materials and other resources. Alternatively, planned evaluation work may focus on preventing the development of separate learning tracks (e.g., removing policies and existing barriers for students to access higher-level course opportunities). It could also attend to monitoring how and whether teachers are taking up professional learning in their classrooms or prioritize the evaluation of how implemented policies and practices are expanding deeper learning opportunities for all students.

CONCLUSION

Moving toward a learning-centered vision that supports ambitious teaching and learning will require stakeholders to take up the courageous and difficult work of disrupting and dismantling existing infrastructures that continue to perpetuate student inequities in many districts and schools across the United States. This work is especially salient now, after disruptions to schooling caused by the COVID-19 pandemic, which

exacerbated existing inequalities in educational opportunities and learning on a global scale (Barron Rodriguez et al., 2021; Nicola et al., 2020; United Nations Educational, Scientific and Cultural Organization et al., 2021). Although attempts to address these inequalities have sparked well-intentioned efforts to reignite learning at schools, some states, districts, and schools have renewed their efforts to reinvigorate infrastructures like those seen in the NCLB era. For example, there is a growing interest across states to develop through-year assessments that entail heavy investments in vendor-developed benchmark and interim tests (Marion, 2021). Marion (2021) notes that these through-year assessments are being developed to serve multiple purposes including providing predictive information about student performance on the state end of year test during the school year, as well as providing “instructionally useful” information to students and teachers. We have also observed a few districts beginning to reshape programming offered to students by increasing instructional time dedicated to tested subjects (English language arts and math) and refamiliarizing students with testing strategies to accelerate learning from learning losses attributed to the pandemic. Increases in learning time for these two tested subjects come at the expense of programming for students in the sciences, social studies, and other disciplinary areas that contribute to and enhance student well-being and learning.

State, district, and school leaders should heed the fact that scant evidence supports the idea that short-term gains achieved by enacting test-supportive teaching and learning practices will lead to deeper learning (McTighe & Gareis, 2021). Instead, evidence shows that collectively building and sustaining instructional infrastructure designed to advance ambitious teaching will, in the long run, provide students with the knowledge and skills to address urgent societal issues given the cross-cutting tensions impacting our social, racial, economic, environmental, and political climate today. This means cultivating thriving classroom activity systems that motivate students to build a repertoire of knowledge, skills, and confidence to productively engage with others as citizens, participants, and contributors to communities at the local, national, and global levels.

REFERENCES

- Anderson, E., & Young, M. D. (2018). If they knew then what we know now, why haven't things changed? An examination of district effectiveness research. *Frontiers in Education*, 3. <https://doi.org/10.3389/educ.2018.00087>.
- Armstrong, A. L. (2021, December 1). *The representation of social groups in U.S. educational materials and why it matters*. New America. <https://www.newamerica.org/education-policy/reports/the-representation-of-social-groups-in-u-s-educational-materials-and-why-it-matter>.
- Arriaza, G. (2004). Making changes that stay made: School reform and community involvement. *The High School Journal*, 87(4), 10–24. <https://doi.org/10.1353/hsj.2004.0007>.
- Au, W. (2007). High-stakes testing and curricular control: A qualitative metasynthesis. *Educational Researcher*, 36(5), 258–267. <https://doi.org/10.3102/0013189X07306523>.
- Aycock, B., Drew, S., Housman, G., Ivory, S., McGill, T. A. W., Rosvally, H., Servino, M., & Wampler, R. (2019). *Why do some things wash up on the beach and others don't?* Next Generation Science Storylines. <https://www.nextgenstorylines.org/why-do-some-things-wash-up-on-the-beach-and-others-dont>.
- Barron Rodriguez, M., Cobo, C., Munoz-Najar, A., & Sanchez Ciarrusta, I. (2021). *Remote learning during the global school lockdown: Multi-country lessons*. The World Bank. <http://hdl.handle.net/10986/36141>.
- Berns, R. (2015). *Child, family, school, community: Socialization and support* (10th ed.). Cengage Learning.
- Blazar, D., & Pollard, C. (2017). Does test preparation mean low-quality instruction? *Educational Researcher*, 46(8), 420–433. <https://doi.org/10.3102/0013189X17732753>.

- Boston, M. D., & Smith, M. S. (2009). Transforming secondary mathematics teaching: Increasing the cognitive demands of instructional tasks used in teachers' classrooms. *Journal for Research in Mathematics Education*, 40(2), 119–156.
- Brookhart, S. M. (2013a). Grading. In J. H. McMillan (Ed.), *SAGE handbook of research on classroom assessment* (pp. 257–271). Sage.
- Brookhart, S. M. (2013b). *How to create and use rubrics for formative assessment and grading*. Association for Supervision and Curriculum Development.
- Bryk, A. S., Sebring, P. B., Allensworth, E., Easton, J. Q., & Luppescu, S. (2009). *Organizing schools for improvement: Lessons from Chicago*. University of Chicago Press. <https://doi.org/10.7208/chicago/9780226078014.001.0001>.
- Burkhardt, H., & Schoenfeld, A. (2019). Formative assessment in mathematics. In H. L. Andrade, R. E. Bennett, & C. J. Cizek (Eds.), *Handbook of formative assessment in the disciplines*. Routledge.
- Clarke, D., & Hollingsworth, H. (2002). Elaborating a model of teacher professional growth. *Teaching and Teacher Education*, 18(8), 947–967. [https://doi.org/10.1016/S0742-051X\(02\)00053-7](https://doi.org/10.1016/S0742-051X(02)00053-7).
- Coburn, C. E. (2003). Rethinking scale: Moving beyond numbers to deep and lasting change. *Educational Researcher*, 32(6), 3–12. <https://doi.org/10.3102/0013189X032006003>.
- Coburn, C. E., Toure, J., & Yamashita, M. (2009). Evidence, interpretation, and persuasion: Instructional decision making at the district central office. *Teachers College Record*, 111(4), 1115–1161.
- Coburn, C. E., Mata, W. S., & Choi, L. (2013a). The embeddedness of teachers' social networks: Evidence from a study of mathematics reform. *Sociology of Education*, 86(4), 311–342. <https://doi.org/10.1177/0038040713501147>.
- Coburn, C. E., Penuel, W. R., & Geil, K. E. (2013b). *Research-practice partnerships: A strategy for leveraging research for educational improvement in school districts*. William T. Grant Foundation. <https://files.eric.ed.gov/fulltext/ED568396.pdf>.
- Cohen, D. K., Peurach, D. J., Glazer, J. L., Gates, K. E., & Goldin, S. (Eds.). (2013). *Improvement by design: The promise of better schools*. The University of Chicago Press.
- Conley, D. (2015). A new era for educational assessment. *Education Policy Analysis Archives*, 23, 8. <https://doi.org/10.14507/epaa.v23.1983>.
- Cook-Endres, T., Taylor, A., & Penuel, B. (2014). *Using curriculum adaptation as a strategy to help teachers learn about NGSS and developing aligned instructional materials*. Brief 5. STEM teaching tools. <https://stemteachingtools.org/brief/5>.
- Crocco, M. S., & Costigan, A. T. (2007). The narrowing of curriculum and pedagogy in the age of accountability urban educators speak out. *Urban Education*, 42(6), 512–535. <https://doi.org/10.1177/0042085907304964>.
- Darling-Hammond, L. (2004). Inequality and the right to learn: Access to qualified teachers in California's public schools. *Teachers College Record*, 106(10), 1936–1966. <https://doi.org/10.1111/j.1467-9620.2004.00422.x>.
- Darling-Hammond, L. (2020). Culture, learning, and policy. In N. S. Nasir, C. D. Lee, R. Pea, & M. McKinney de Royston (Eds.), *Handbook of the cultural foundations of learning* (pp. 404–426). Routledge. <https://doi.org/10.4324/9780203774977-28>.
- Darling-Hammond, L., & Adamson, F. (2014). *Beyond the bubble test: How performance assessments support 21st century learning*. Jossey-Bass & Pfeiffer Imprints, Wiley.
- Darling-Hammond, L., & Cook-Harvey, C. M. (2018). *Educating the whole child: Improving school climate to support student success*. Learning Policy Institute. <https://doi.org/10.54300/145.655>.
- Darling-Hammond, L., Hyler, M. E., & Gardner, M. (2017). *Effective teacher professional development*. Learning Policy Institute. <https://doi.org/10.54300/122.311>.
- Darling-Hammond, L., Flook, L., Cook-Harvey, C., Barron, B., & Osher, D. (2020). Implications for educational practice of the science of learning and development. *Applied Developmental Science*, 24(2), 97–140. <https://doi.org/10.1080/10888691.2018.1537791>.
- Diaz-Bilello, E., York, A., & Mork, K. (2022). *Learnings from a multi-site case study of former turnaround schools*. Center for Assessment, Design, Research and Evaluation, University of Colorado Boulder. <https://www.colorado.edu/cadre/2022/08/12/learnings-multi-site-case-study-former-turnaround-schools>.
- Diefes-Dux, H. A. (2018). Student self-reported use of standards-based grading resources and feedback. *European Journal of Engineering Education*, 44(6), 838–849. <https://doi.org/10.1080/03043797.2018.1483896>.

- Dresser, R. (2012). The impact of scripted literacy instruction on teachers and students. *Issues in Teacher Education*, 21(1), 71–87.
- Duncan-Owens, D. (2009). Scripted reading programs: Fishing for success. *Principal*. https://www.naesp.org/sites/default/files/resources/2/Principal/2009/J-F_p26.pdf.
- Dyer, K. (2017, July 13). *Understanding formative, interim, and summative assessments and their role in student learning*. NWEA. <https://www.nwea.org/blog/2017/understanding-formative-interim-summative-assessments-role-student-learning>.
- Elmore, R. F., & Burney, D. (1997). *Investing in teacher learning: Staff development and instructional improvement in Community School District #2, New York City*. National Commission on Teaching & America's Future. <https://files.eric.ed.gov/fulltext/ED416203.pdf>.
- Faxon-Mills, S., Hamilton, L. S., Rudnick, M., & Stecher, B. M. (2013). *New assessments, better instruction? Designing assessment systems to promote instructional improvement*. RAND Corporation. https://www.rand.org/pubs/research_reports/RR354.html.
- Feldman, J. (2023). *Grading for equity: What it is, why it matters, and how it can transform schools and classrooms* (2nd ed.). Corwin.
- Fullan, M., & Quinn, J. (2015). *Coherence: The right drivers in action for schools, districts, and systems*. Corwin: A Sage Company.
- Gitomer, D. H., & Iwatani, E. (2022). Two communities' views on test fairness. *Educational Assessment*, 27(2), 197–203. <https://doi.org/10.1080/10627197.2022.2087624>.
- Guskey, T. R. (2021). Learning from failures: Lessons from unsuccessful grading reform initiatives. *NASSP Bulletin*, 105(3), 192–199. <https://doi.org/10.1177/01926365211029375>.
- Hamilton, L. S., Stecher, B. M., & Yuan, K. (2008). *Standards-based reform in the United States: History, research, and future directions*. RAND Corporation. https://www.rand.org/content/dam/rand/pubs/reprints/2009/RAND_RP1384.pdf.
- Hanno, E. C. (2022). Immediate changes, trade-offs, and fade-out in high-quality teacher practices during coaching. *Educational Researcher*, 51(3), 173–185. <https://doi.org/10.3102/0013189X211062896>.
- Hany, K., Proctor, M., Wollenweber, J., & Al-Bataineh, A. (2016). Teacher perception of standards-based grading: Implication and effectiveness. *Journal of Teaching and Education*, 5(1), 17.
- Hargreaves, A. (2000). Mixed emotions: Teachers' perceptions of their interactions with students. *Teaching and Teacher Education*, 16(8), 811–826. [https://doi.org/10.1016/S0742-051X\(00\)00028-7](https://doi.org/10.1016/S0742-051X(00)00028-7).
- Heiser, P., Simidian, G., Albert, D., Garruto, J., Catucci, D., Faustino, P., McCarten, K., & Caci, K. (2015). *Anxious for success: High anxiety in New York's schools*. New York Association of School Psychologists and New York State School Boards Association. https://www.nyssba.org/clientuploads/nyssba_pdf/Test_Anxiety_Report.pdf.
- Henderson, A. T., & Mapp, K. L. (2002). *A new wave of evidence: The impact of school, family, and community connections on student achievement*. National Center for Family and Community Connections with Schools: SEDL. <https://sedl.org/connections/resources/evidence.pdf>.
- Hofman, P., Goodwin, B., & Kahl, S. (2015). *Re-balancing assessment: Placing formative and performance assessment at the heart of learning and accountability*. McREL International. <https://files.eric.ed.gov/fulltext/ED568906.pdf>.
- Honig, M. I. (2013). *From tinkering to transformation: Strengthening school district central office performance*. American Enterprise Institute. <https://www.aei.org/research-products/report/from-tinkering-to-transformation-strengthening-school-district-central-office-performance>.
- Honig, M. I., & Coburn, C. (2008). Evidence-based decision making in school district central offices: Toward a policy and research agenda. *Educational Policy*, 22(4), 578–608. <https://doi.org/10.1177/0895904807307067>.
- Honig, M. I., & Rainey, L. R. (2015). *How school districts can support deeper learning: The need for performance alignment*. Jobs for the Future. <https://files.eric.ed.gov/fulltext/ED560756.pdf>.
- Honig, M. I., & Venkateswaran, N. (2012). School–central office relationships in evidence use: Understanding evidence use as a systems problem. *American Journal of Education*, 118(2), 199–222. <https://doi.org/10.1086/663282>.

- Hopkins, M., & Spillane, J. P. (2015). Conceptualizing relations between instructional guidance infrastructure (IGI) and teachers' beliefs about mathematics instruction: Regulatory, normative, and cultural-cognitive considerations. *Journal of Educational Change*, 16(4), 421–450. <https://doi.org/10.1007/s10833-015-9257-1>.
- Hopkins, M., & Woulfin, S. L. (2015). School system (re)design: Developing educational infrastructures to support school leadership and teaching practice. *Journal of Educational Change*, 16(4), 371–377. <https://doi.org/10.1007/s10833-015-9260-6>.
- Kang, H., & Furtak, E. M. (2021). Learning theory, classroom assessment, and equity. *Educational Measurement: Issues and Practice*, 40, 73–82. <https://doi.org/10.1111/emip.12423>.
- Kang, H., Windschitl, M., Stroupe, D., & Thompson, J. (2016). Designing, launching, and implementing high quality learning opportunities for students that advance scientific thinking. *Journal of Research in Science Teaching*, 53(9), 1316–1340. <https://doi.org/10.1002/tea.21329>.
- Kaufman, J. H., Tosh, K., & Mattox, T. (2020). *Are U.S. teachers using high-quality instructional materials?* RAND Corporation. <https://doi.org/10.7249/RR2575.11-1>.
- Kraft, M. A., Simon, N. S., & Lyon, M. A. (2021). *Sustaining a sense of success: The protective role of teacher working conditions during the COVID-19 pandemic*. Annenberg Institute, Brown University. <https://doi.org/10.26300/35nj-v890>.
- Ladson-Billings, G. (2014). Culturally relevant pedagogy 2.0: a.k.a. the remix. *Harvard Educational Review*, 84(1), 74–84. <https://doi.org/10.17763/haer.84.1.p2rj131485484751>.
- Latham, B. (2018, March 22). Looking for lasting change? Start talking to other departments. EdSurge. <https://www.edsurge.com/news/2018-03-22-looking-for-lasting-change-start-talking-to-other-departments>.
- Leithwood, K., Sun, J., & McCullough, C. (2019). How school districts influence student achievement. *Journal of Educational Administration*, 57(5), 519–539. <https://doi.org/10.1108/JEA-09-2018-0175>.
- Lewis, D. (2020). Gender effects on re-assessment attempts in a standards-based grading implementation. *PRIMUS*, 30(5), 539–551. <https://doi.org/10.1080/10511970.2019.1616636>.
- Linn, R. L., & Burton, E. (1994). Performance-based assessment: Implications of task specificity. *Educational Measurement: Issues and Practice*, 13(1), 5–8. <https://doi.org/10.1111/j.1745-3992.1994.tb00778.x>.
- Loewus, L. (2017, December 19). *Majority of teachers say reforms have been “too much.”* EducationWeek. <https://www.edweek.org/teaching-learning/majority-of-teachers-say-reforms-have-been-too-much/2017/12>.
- Mac Iver, M. A., & Farley, E. (2003). *Bringing the district back in: The role of the central office in improving instruction and student achievement*. Johns Hopkins University and Howard University. <https://files.eric.ed.gov/fulltext/ED481804.pdf>.
- Marion, S. (2021, April 29). *It might just be a pile of bricks! The challenges of creating balanced assessment systems*. Nation Center for the Improvement of Educational Assessment. <https://www.nciea.org/blog/it-might-just-be-a-pile-of-bricks>.
- Marion, S. (2021, September 29). *Trying to serve multiple uses with through year assessments. Assumptions, claims, and the evidence needed to support them*. National Center for the Improvement of Educational Assessment. <https://www.nciea.org/blog/trying-to-serve-multiple-uses-with-through-year-assessments>.
- Marsh, J. A. (2002). How districts relate to states, schools, and communities: A review of emerging literature. In A. M. Hightower, M. S. Knapp, J. A. Marsh, & M. W. McLaughlin (Eds.), *School districts and instructional renewal* (pp. 25–40). Teachers College Press.
- Massell, D., & Goertz, M. E. (2002). District strategies for building instructional capacity. In A. M. Hightower, M. S. Knapp, J. A. Marsh, & M. W. McLaughlin (Eds.), *School districts and instructional renewal* (pp. 43–60). Teachers College Press.
- McLaughlin, M., & Talbert, J. (2002). Reforming districts. In A. M. Hightower, M. S. Knapp, J. A. Marsh, & M. W. McLaughlin (Eds.), *School districts and instructional renewal* (pp. 173–192). Teachers College Press.
- McMillan, J. H. (2001). Secondary teachers' classroom assessment and grading practices. *Educational Measurement: Issues and Practice*, 20(1), 20–32. <https://doi.org/10.1111/j.1745-3992.2001.tb00055.x>.
- McMillan, J. H., Myran, S., & Workman, D. (2002). Elementary teachers' classroom assessment and grading practices. *The Journal of Educational Research*, 95(4), 203–213. <https://doi.org/10.1080/00220670209596593>.
- McTighe, J., & Gareis, C. (2021, June 10). *Assessing deeper learning after a year of change*. ASCD. <https://www.ascd.org/el/articles/assessing-deeper-learning-after-a-year-of-change>.

- Mehta, J., & Fine, S. (2015). Bringing values back in: How purposes shape practices in coherent school designs. *Journal of Educational Change*, 16, 483–510. <https://doi.org/10.1007/s10833-015-9263-3>.
- Meyers, C. V. (2020). An urban district's struggle to preserve school turnaround change. *Urban Education*. <https://doi.org/10.1177/0042085920966031>.
- Milosovic, S. (2007). Building a case against scripted reading programs. *Education Digest: Essential Readings Condensed for Quick Review*, 73(1), 27–30.
- Muñoz, M. A., & Guskey, T. R. (2015). Standards-based grading and reporting will improve education. *Phi Delta Kappan*, 96(7), 64–68. <https://doi.org/10.1177/0031721715579043>.
- Next Generation Science Storylines. (n.d.). *What are storylines?* <https://www.nextgenstorylines.org/what-are-storylines>.
- Nicola, M., Alsafi, Z., Sohrabi, C., Kerwan, A., Al-Jabir, A., Iosifidis, C., Agha, M., & Agha, R. (2020). The socio-economic implications of the coronavirus pandemic (COVID-19): A review. *International Journal of Surgery*, 78, 185–193. <https://doi.org/10.1016/j.ijisu.2020.04.018>.
- Nieto, S. (2013). *Finding joy in teaching students of diverse backgrounds: Culturally responsive and socially just practices in U.S. classrooms*. Heinemann.
- No Child Left Behind Act of 2001, 20 U.S.C. 6301. (2001). <https://www.congress.gov/bill/107th-congress/house-bill/1/text>.
- Olson, L., & Jerald, C. (2020). *The big test: The future of statewide standardized assessments*. FutureEd. <https://www.future-ed.org/the-big-test-the-future-of-statewide-standardized-testing>.
- Penuel, W. R. (2019). Infrastructuring as a practice of design-based research for supporting and studying equitable implementation and sustainability of innovations. *Journal of the Learning Sciences*, 28(4–5), 659–677. <https://doi.org/10.1080/10508406.2018.1552151>.
- Penuel, W. R., Briggs, D. C., Davidson, K. L., Herlihy, C., Sherer, D., Hill, H. C., Farrell, C., & Allen, A.-R. (2017). How school and district leaders access, perceive, and use research. *AERA Open*, 3(2). <https://doi.org/10.1177/2332858417705370>.
- Penuel, W. R., Bell, P., & Neill, T. (2020a). Creating a system of professional learning that meets teachers' needs. *Phi Delta Kappan*, 101(8), 37–41. <https://doi.org/10.1177/0031721720923520>.
- Penuel, W. R., Riedy, R., Barber, M. S., Peurach, D. J., LeBouef, W. A., & Clark, T. (2020b). Principles of collaborative education research with stakeholders: Toward requirements for a new research and development infrastructure. *Review of Educational Research*, 90(5), 627–674. <https://doi.org/10.3102/0034654320938126>.
- Perie, M., Marion, S., & Gong, B. (2009). Moving toward a comprehensive assessment system: A framework for considering interim assessments. *Educational Measurement: Issues and Practice*, 28(3), 5–13. <https://doi.org/10.1111/j.1745-3992.2009.00149.x>.
- Polikoff, M. (2021). *Beyond standards: The fragmentation of education governance and the promise of curriculum reform*. Harvard Education Press.
- Randall, J., Slomp, D., Poe, M., & Oliveri, M. E. (2022). Disrupting white supremacy in assessment: Toward a justice-oriented, antiracist validity framework. *Educational Assessment*, 27(2), 170–178. <https://doi.org/10.1080/10627197.2022.2042682>.
- Reiser, B. J., Novak, M., McGill, T. A. W., & Penuel, W. R. (2021). Storyline units: An instructional model to support coherence from the students' perspective. *Journal of Science Teacher Education*, 32(7), 805–829. <https://doi.org/10.1080/1046560X.2021.1884784>.
- Ruiz-Primo, M. A., Li, M., Wills, K., Giamellaro, M., Lan, M.-C., Mason, H., & Sands, D. (2012). Developing and evaluating instructionally sensitive assessments in science. *Journal of Research in Science Teaching*, 49(6), 691–712. <https://doi.org/10.1002/tea.21030>.
- Schoenfeld, A. H. (2013). Classroom observations in theory and practice. *ZDM*, 45, 607–621. <https://doi.org/10.1007/s11858-012-0483-1>.
- Schoenfeld, A. H. (2017). Uses of video in understanding and improving mathematical thinking and teaching. *Journal of Mathematics Teacher Education*, 20, 415–432. <https://doi.org/10.1007/s10857-017-9381-3>.
- Shepard, L. A. (2017). Formative assessment: Caveat emptor. In C. A. Dwyer (Ed.), *The future of assessment: Shaping teaching and learning* (pp. 279–303). Routledge. <https://doi.org/10.4324/9781315086545>.
- Shepard, L. A. (2019). Classroom assessment to support teaching and learning. *ANNALS of the American Academy of Political and Social Science*, 683(1), 183–200. <https://doi.org/10.1177/0002716219843818>.

- Shepard, L. A., Penuel, W. R., & Pellegrino, J. W. (2018). Using learning and motivation theories to coherently link formative assessment, grading practices, and large-scale assessment. *Educational Measurement: Issues and Practice*, 37(1), 21–34. <https://doi.org/10.1111/emip.12189>.
- Snipes, J., Doolittle, F., & Herlihy, C. (2002). *Foundations for success: Case studies of how urban school systems improve student achievement*. MDRC. https://www.mdrc.org/sites/default/files/foundations_for_success_summary.pdf.
- Spillane, J. P., Healey, K., Parise, L. M., & Kenney, A. (2011). A distributed perspective on learning leadership. In J. Robertson & H. Timperley (Eds.), *Leadership and learning* (pp. 159–171). SAGE Publications. <https://doi.org/10.4135/9781446288931.n12>.
- Stosich, E. L., & Bae, S. (2018). Engaging diverse stakeholders to strengthen policy. *Phi Delta Kappan*, 99(8), 8–12. <https://doi.org/10.1177/0031721718775670>.
- Suárez, E., Bell, P., McCulloch, A., & Starr, M. (2020). *Why you should stop pre-teaching science vocabulary and focus on students developing conceptual meaning first*. STEM Teaching Tools. <https://stemteachingtools.org/brief/66>.
- Supovitz, J. A. (2018). Teacher leaders' work with peers in a quasi-formal teacher leadership model. *School Leadership & Management*, 38(1), 53–79. <https://doi.org/10.1080/13632434.2017.1389718>.
- Teoh, M., Coggins, C., Guan, C., & Hiler, T. (2014). *The student & the stopwatch: How much time do American students spend on testing?* Teach Plus. https://teachplus.org/wp-content/uploads/files/publication/pdf/the_student_and_the_stopwatch.pdf.
- Thompson, J., Mawyer, K., Johnson, H., Scipio, D., & Luehmann, A. (2021). *C²AST (Critical and Cultural Approaches to Ambitious Science Teaching): From responsive teaching toward developing culturally and linguistically sustaining science teaching practices*. National Science Teaching Association. <https://www.nsta.org/science-teacher/science-teacher-septemberoctober-2021/c2ast-critical-and-cultural-approaches>.
- Togneri, W., & Anderson, S. E. (2003). *Beyond islands of excellence: What districts can do to improve instruction and achievement in all schools—A leadership brief*. Learning First Alliance. <https://files.eric.ed.gov/fulltext/ED475875.pdf>.
- Townsend, M. (2019). Considering standards-based grading: Challenges for secondary school leaders. *Journal of School Administration Research and Development*, 4(1), 35–38.
- United Nations Educational, Scientific and Cultural Organization, UNESCO Institute for Statistics, United Nations Children's Fund, UNICEF Office of Research, The World Bank, & Organisation for Economic Co-operation and Development. (2021). *What's next? Lessons on education recovery: Findings from a survey of ministries of education amid the COVID-19 pandemic*. <https://www.oecd.org/education/what-s-next-lessons-on-education-recovery-697bc36e-en.htm>.
- Wang, E. L., Tuma, A. P., Doan, S., Henry, D., Lawrence, R. A., Woo, A., & Kaufman, J. H. (2021). *Teachers' perceptions of what makes instructional materials engaging, appropriately challenging, and usable: A survey and interview study*. RAND Corporation. <https://doi.org/10.7249/RRA134-2>.
- Watson, C. (2014). Effective professional learning communities? The possibilities for teachers as agents of change in schools. *British Educational Research Journal*, 40(1), 18–29. <https://doi.org/10.1002/berj.3025>.
- Wiggins, G., & McTighe, J. (2005). *Understanding by design* (2nd ed.). Association for Supervision and Curriculum Development.
- Willis, L., Badrinarayan, A., & Martinez, M. (2022). *Quality criteria for systems of performance assessment for school, district, and network leaders*. Learning Policy Institute. <https://doi.org/10.54300/439.730>.
- Windschitl, M., Thompson, J., Braaten, M., & Stroupe, D. (2012). Proposing a core set of instructional practices and tools for teachers of science. *Science Education*, 96(5), 878–903. <https://doi.org/10.1002/sce.21027>.
- Windschitl, M., Thompson, J., & Braaten, M. (2018). *Ambitious science teaching*. Harvard Education Press.