

Effective Supervision

by Robert J. Marzano, Tony Frontier and David Livingston

Chapter 2. A Brief History of Supervision and Evaluation

Before providing the specific suggestions for the initiatives described in Chapter 1, it is useful to consider the history of teacher supervision and evaluation in the United States. We include the topic of evaluation in our review because it is so commonly linked to supervision.

The Early Days of Supervision and Evaluation

In the 1700s, education was not considered a professional discipline or field of study. Early towns in the United States turned to existing power structures, such as local government and the clergy, to hire teachers and make judgments about their teaching. Clergy were considered logical choices for this role because of their extensive education and presumed ability to guide religious instruction in schools (Tracy, 1995, p. 320). The teacher was considered a servant of the community. Individual supervisors or supervisory committees were charged with monitoring the quality of instruction. These supervisors had nearly unlimited power to establish criteria for effective instruction and to hire and fire teachers (Burke & Krey, 2005). Because there was no necessary agreement as to the importance or nature of pedagogical expertise, the quality and type of feedback to teachers was highly varied.

A rising industrial base and the common schooling movement that extended through the 1800s spawned large urban areas with more complex school systems. In these larger schools and districts, a demand grew for teachers who held expertise in specific disciplines and for administrators who could assume increasingly complex roles. One teacher within a building was often selected to assume administrative duties. This "principal" teacher ultimately grew into the role of building principal.

The trend toward specialized roles started in large urban districts and soon spread to smaller cities and rural areas (Tracy, 1995). About this time, it was acknowledged that clergy didn't necessarily have the knowledge base to make informed judgments about teacher effectiveness. Tracy explains, "Rather than simply understanding the mores of the community, the supervisor now needed to have subject area knowledge and teaching skills" (p. 323). Clearly, clergy were not trained for such a role.

By the mid-1800s, the view of teaching was that it was a complex endeavor requiring complex feedback if expertise was to be fostered. Blumberg (1985) notes that at this time supervision began to focus on improving instruction. He offers the following quote from an 1845 document titled *The Annual Report of the Superintendent of Common Schools of the State of New York*:
Too much reliance ought not to be placed upon visitation to the schools, to give method to the teacher and efficacy to his instructions. Instruction is the primary object of visitation, and ... more instruction can be given to teachers of a town when assembled together in one day. (p. 63, as cited in original source of 1845, p. 131)

Blumberg asserts that although supervisors were no longer clergy, they were no less evangelical. Within a given county, superintendents traveled from community to community and school to school, proselytizing for more effective instructional practices. As one superintendent stated, "The only salvation for the republic is to be sought for in our schools" (1845, p. 19, as cited in Blumberg, 1985).

The period from the beginning of formal education in the United States up to the mid-1800s saw the dawning of the awareness that pedagogical skills are a necessary component of effective teaching. Although there was little or no formal discussion about the specifics of these skills, the acknowledgment of their importance might be considered the first step in the journey to a comprehensive approach to developing teacher expertise.

The Period of Scientific Management

The latter part of the 19th century and the early part of the 20th century were dominated by two competing views of education. One was embodied in the writings of John Dewey. Dewey was one of the most prolific writers and thinkers in the field of education in the early 20th century. He saw democracy, not scientific management, as the conceptual underpinning of human progress. He argued that schools should be organized in such a way that students can practice citizenship and further develop the ideals of democracy (Dewey, 1938, 1981). Progressive ideas such as a student-centered education, connecting the classroom to the real world, differentiation based on student learning needs, and integration of content areas were espoused by Dewey as ways of bridging the gap between students' passive role as learners and the active role they would need to play as citizens.

The second view of education was embodied in the work of Frederick Taylor. Taking a scientific view of management, Taylor believed that measurement of specific behaviors of factory workers was perhaps the most powerful means to improve production. He argued that if there were 100 ways to perform a task, some methods would be more efficient than others. By studying the various ways a task such as shoveling coal could be performed, the *one best method* could be determined. According to Taylor (1911), these principles could be applied to discrete tasks such as shoveling coal and to more systemic tasks such as the selection of workers, development of training programs, and processes for dividing labor. Taylor's ideas resonated with engineers and business owners, and colleges of engineering and business were well positioned to infuse his principles into their courses. Taylor's principles also began to have an impact on K–12 education. Led by Edward Thorndike, educators began to view measurement as the ultimate tool for a more scientific approach to schooling. Thorndike's theories were applied to administration by Ellwood Cubberley. Originally published in 1916, Cubberley's book *Public School Administration* (1929) described how Taylor's principles could be used to manage schools in the same way factories are managed:

Our schools are, in a sense, factories in which the raw products (children) are to be shaped and fashioned into products to meet the various demands of life. The specifications for manufacturing

come from the demands of twentieth century civilization and is the business of the school to build its pupils according to the specifications laid down. (p. 338)

Based on the factory metaphor, Cubberley laid out a set of principles for school administrators that emphasized measurement and analysis of data to ensure that teachers and schools were productive. In the third edition of his book, *Public School Administration* (1929), Cubberley provided specific examples of how a scientific approach could be applied when visiting teachers' classrooms. He described specific feedback that a supervisor might provide to a teacher. For example, on a scale from *A* to *F*, this 6th grade teacher was given a *D* for her arithmetic lesson. Cubberley's supervisory form stated:

Weak Points: Entirely wrong procedure for type of problems used. No attempt at problem-solving instruction. ...

Suggestions Made: Explained to her that, being a new teacher to our schools, she evidently did not know how we taught Arithmetic. Explained faults of the lesson, but commended her managerial ability. Told her how she should handle such work, and gave her Newcomb's *Modern Methods of Teaching Arithmetic* to take home and read designated chapters. (Cubberley, 1929, p. 327)

Building on Cubberley's work, William Wetzel (1929) proposed using measures of student learning to determine the effectiveness of a teacher or school. These measures were in addition to focusing on a teacher's use of specific strategies and behaviors. However, Wetzel distanced himself from the metaphor of schools as factories with a manufacturing function. He recommended three components as the basis for scientific supervision: the use of aptitude tests to determine the ability level of each child; the establishment of clear, measurable objectives for each course; and the use of reliable measures of student learning.

Through the 1930s, there was continued tension between the scientific approach to schooling, including a greater reliance on standardized tests, and the approach that focused on social development and democratic values. To some extent, this was a false dichotomy. The science of education as proposed by Cubberley and Wetzel dealt more with the feedback system used to determine if teachers, schools, and districts were being effective. To this extent, their emphasis was on data with which to make decisions about future actions. Considered from this perspective, some of Cubberley and Wetzel's recommendations might be considered precursors to some of our recommendations regarding the use of data for feedback. Dewey's focus was more on the ultimate goal of education. The two perspectives are not innately incompatible. One can use data for feedback but still maintain the goal of an education system that fosters democratic ideals. Nonetheless, the two perspectives were not described or perceived in a fashion that allowed for integration, and the tension between them continued through the Great Depression.

Post-World War II

The period immediately after World War II began with a swing away from the scientific approach to schooling. Rather than describing supervisory processes in terms of raw materials and products,

the literature began to focus on the teacher as *an individual*. Emphasis was placed on not only assisting the teacher to develop his or her unique skills, but also tending to his or her emotional needs. The January 1946 issue of *Educational Leadership* magazine, published only a few months after the conclusion of World War II, reflects this shift. In an article titled "The Supervisory Visit," Elsie Coleman (1945) stated that "the first fundamental in understanding the teacher is ... that the teacher is a person, different from every other person, living in an environment which affects and in turn is affected by that person" (p. 165). In the same issue of *Educational Leadership*, Lewis and Leps (1946) described the supervisory process as though it were an extension of efforts to liberate Europe. Guidelines for a successful supervisory model included (1) democratic ideals, (2) opportunities for initiative, (3) understanding human limitations, (4) shared decision making, and (5) delegation of responsibility (p. 163). In describing this new world of supervision, Lewis and Leps stated, "The school administrator, with the acceptance of the community, is gaining the courage to utilize the creative force to be gained in freeing the human beings who comprise the school situation to participate in the making of policies and plans for their execution; and, hence, to utilize the force and creativity inherent in the democratic process" (p. 161).

In spite of the emphasis on the teacher as an individual, the role of the supervisor during this era was defined in rather specific terms. Unfortunately, the list of supervisory responsibilities was quite long and broad. For example, Swearingen (1946) described the role of the supervisor as including the following areas: the curriculum, teaching personnel, the teaching/learning situation, the emotional quality of the classroom, resources and materials of instruction, auxiliary functions including working with the school lunch service, attendance, distribution of textbooks, public relations, and working with cooperative groups and agencies. In his text *Instructional Supervision: A Guide to Modern Practice*, William Melchoir (1950) described supervision as including individual meetings with teachers, faculty meetings, business meetings, social meetings, workshops and other committee meetings in addition to "classroom visitation for observation and study" (p. 51). While classroom visitation is discussed explicitly in Melchoir's text, its relative importance (based on page count in the book) seems to imply that the supervisor's role was more about management of the physical plant than instructional leadership. For example, 23 pages in the book were devoted to "Beautifying Grounds and Buildings" (pp. 107–130), while only 16 pages were devoted to classroom observation (pp. 364–380). Finally, in her article titled "So Begins—So Ends the Supervisor's Day," Ethel Thompson (1952) added to the growing list of responsibilities by describing the supervisor's role as attending student placement conferences, observing in a classroom, working with parents and principals, completing paperwork, meeting with various school committees, attending student conferences, recruiting new teachers, meeting with various professional organizations, doing demonstration lessons, and acting as a resource to others in the organization.

Although the proliferation of responsibilities for the supervisor was counterproductive at best, one positive outcome from this era was a consensus on the importance and utility of teacher observation. In his article "Teachers Look at Supervision," Matthew Whitehead (1952) described six broad areas of supervision and surveyed teachers as to their perceptions of the importance of each area. Noting the importance of effective classroom observation, he pointed out advances that

must be made in observational practices to capitalize on its potential: "Improvements were still needed in following up the visitation with a conference, and in having the principal see the importance of remaining the entire period. It is not fair to teachers to visit them and not hold a conference following the visitation nor is it just to visit in a 'piecemeal' fashion" (p. 102). Whitehead summarized his position by explaining that "administrators should pay more attention to the chief aim of education—effective teaching" (p. 106). It was the recognition of the importance of classroom observation that laid the foundations for one of the most influential movements in supervision.

The Era of Clinical Supervision

Few innovations in the field of education spread as quickly as clinical supervision. Developed in the late 1950s and described in detail in books published in the late 1960s and early 1970s, clinical supervisory models spread like wildfire. By 1980, one study found that about 90 percent of school administrators used some type of clinical supervisory model (Bruce & Hoehn, 1980). Few models in the entire field of education—let alone in the specific domain of educational supervision—have been as widely deployed, as widely disparaged, or as widely misunderstood.

Morris Cogan was a professor and supervisor of candidates in Harvard's Master's of Arts in Teaching (MAT) program in the 1950s. Over years of what might be described, at least in part, as trial and error, he and his colleagues developed a systematic approach to working with student teachers. By 1958, Cogan was lecturing on a process called the "cycle of clinical supervision" (Cogan, 1973). By 1962, a group of educational practitioners working with Cogan in the MAT program had further refined the clinical approach. According to one of those practitioners, Robert Goldhammer, the model was analogous to supervisory practices used in teaching hospitals. The process involved a purposeful, symbiotic relationship between practitioner and resident, where observation and discussion drove both parties to higher levels of growth and effectiveness (Goldhammer, 1969, p. 54).

The model that emerged from these efforts was published in a book by Goldhammer (1969) entitled *Clinical Supervision: Special Methods for the Supervision of Teachers*. Based on visits to hundreds of classrooms and hundreds of supervisory conferences, Goldhammer developed a five-phase process of clinical supervision that was designed to involve teachers and supervisors in a reflective dialogue.

- *Phase 1—Preobservation Conference:* This phase was designed to provide a conceptual framework for the observation. During this phase, the teacher and supervisor planned the specifics of the observation.
- *Phase 2—Classroom Observation:* During this phase, the supervisor observed the teacher using the framework articulated in Phase 1.
- *Phase 3—Analysis:* Data from the observation was organized by the supervisor with the intent of helping teachers participate "in developing evaluations of their own teaching" (p. 63).

- *Phase 4—A Supervision Conference:* The teacher and supervisor engaged in a dialogue about the data. The teacher was asked to reflect upon and explain his or her professional practice. This stage also could include providing "didactic assistance" (p. 70) to the teacher.
- *Phase 5—Analysis of the Analysis:* The supervisor's "practice was examined with all of the rigor and for basically the same purposes that Teacher's professional behavior was analyzed theretofore" (p. 71).

In 1973, Morris Cogan wrote the book *Clinical Supervision*. As mentioned previously, Cogan was one of Goldhammer's professors at Harvard. His focus was on specific classroom behaviors. He noted that supervisors should be looking for "critical incidents" that "impede desired learnings in striking fashion" (p. 172). He also emphasized the fact that the supervisory process should be viewed as a vital aspect of the process of continual improvement in teaching:

A cornerstone of the supervisor's work with the teacher is the assumption that clinical supervision constitutes a continuation of the teacher's professional education. This does not mean that the teacher is "in training," as is sometimes said of preservice programs. It means that he is continuously engaged in improving his practice, as is required of all professionals. In this sense, the teacher involved in clinical supervision must be perceived as a practitioner fulfilling one of the first requirements of a professional—maintaining and developing his competence. He must not be treated as a person being rescued from ineptitude, saved from incompetence, or supported in his stumblings. He must perceive himself to be engaged in the supervisory processes as a professional who continues his education and enlarges his competences. (p. 21)

One of the more interesting aspects of Cogan's perspective was his caution that a supervisor's personal model of teaching might impede his or her ability to provide effective feedback to teachers.

Most teachers have consciously and unconsciously constructed a personal model of the good teacher. Such conceptions generally grow by accretion rather than by critical examination and careful testing. The result is that too often the operating model of the teacher-turned-supervisor is pretty much what he himself does well. When teachers become supervisors, these personal preferences generally operate in full vigor, furnishing many of the criteria for viewing the teaching of others. (1973, p. 54)

It is instructive to contrast the original view of clinical supervision with that into which it evolved. Goldhammer was clear that what is to be observed is the holistic practice of teaching: the interaction of the teacher and student related to student learning. The five phases of the clinical supervision process were intended to be the vehicle to disclose effective instructional practices. However, over time, the five phases became an end in themselves. In some cases, the rich, trusting dialogue envisioned by Goldhammer was reduced to a ritualistic set of steps to be followed. Perhaps contributing to this problem was Goldhammer's resistance to defining any characteristics of effective instruction. In Goldhammer's view, the supervisor should have few if any preconceived

notions of what constitutes effective teaching:

Since I have deliberately not structured my observations in advance so that, for example, I should only record data in certain predetermined categories, and since I have collected as many data as possible in order to alleviate unconscious selectivity, I must now, ex post facto, invent categories of some kind. I must organize the data into classes of one sort or another in order to talk about them. ... Categories of behavior have no objective existence of their own; they do not exist independently in the real world; I make them up. (1969, p. 95)

Regardless of the reasons for its demise, Goldhammer's vision of supervision as a collegial, inquiry-driven quest for more effective instructional practices quickly disappeared. The five phases of the clinical model, absent the rich dialogue proposed by Goldhammer, became the de facto structure for the evaluation of teachers—clearly a purpose for which it was not intended.

The Hunter Model

The next major influence on supervision was the work of Madeline Hunter (1980, 1984). The centerpiece of her work was the seven-step model of a lesson depicted in Figure 2.1.

Figure 2.1. The Hunter Model of Lesson Design

Element	Description
Anticipatory set	A mental set that causes students to focus on what will be learned. It may also give practice in helping students achieve the learning and yield diagnostic data for the teacher. <i>Example:</i> "Look at the paragraph on the board. What do you think might be the most important part to remember?"
Objective and purpose	Not only do students <i>learn</i> more effectively when they know what they're supposed to be learning and why that learning is important to them, but teachers <i>teach</i> more effectively when they have that same information. <i>Example:</i> "Frequently people have difficulty in remembering things that are important to them. Sometimes you feel you have studied hard and yet don't remember some of the important parts. Today, we're going to learn ways to identify what's important, and then we'll practice ways we can use to remember important things."
Input	Students must acquire new information about the knowledge, process, or skill they are to achieve. To design the input phase of the lesson so that a successful outcome becomes predictable, the teacher must have analyzed the final objective to identify knowledge and skills that need to be acquired.
Modeling	"Seeing" what is meant is an important adjunct to learning. To avoid stifling creativity, showing several examples of the process or products that students are expected to acquire or produce is helpful.
Checking for	Before students are expected to do something, the teacher should determine that

understanding	they understand what they are supposed to do and that they have the minimum skills required.
Guided practice	Students practice their new knowledge or skill <i>under direct teacher supervision</i> . New learning is like wet cement; it is easily damaged. An error at the beginning of learning can easily "set" so that correcting it later is harder than correcting it immediately.
Independent practice	Independent practice is assigned only after the teacher is reasonably sure that students will not make serious errors. After an initial lesson, students are frequently not ready to practice independently, and the teacher has committed a pedagogical error if unsupervised practice is expected.
<p><i>Source:</i> Adapted from M. Hunter (1984), "Knowing, Teaching, and Supervising." In P. Hosford (Ed.), <i>Using What We Know About Teaching</i> (pp. 169–192). Alexandria, VA: ASCD.</p>	

Although the seven-step framework for a lesson is the most well-known aspect of Hunter's work, she contributed many other ideas to the process of supervision. For example, she championed the idea of using professional development to articulate a common language of instruction. She also identified a variety of purposes for supervisory conferences that included the following:

- To identify, label, and explain instructional behaviors as related to research;
- To encourage teachers to consider alternative approaches that are aligned to their style of teaching;
- To help teachers identify components of lessons that were not as effective as they had hoped;
- To identify and describe "less effective aspects of teaching that were not evident to the teacher" (1980, p. 410);
- To promote the continued growth of excellent teachers;
- To evaluate "what has occurred in and resulted from a series of instructional conferences" supportable by objective evidence rather than based on subjective opinion (1980, p. 412).

Observation and script taping were critical components of Hunter's process of supervision. During script taping, a supervisor recorded teaching behaviors and then later categorized them into those that "promoted learning; those that used precious time and energy, yet contributed nothing to learning; and those that, unintentionally, actually interfered with learning" (Hunter, 1980, p. 409). After script taping, supervisors conferred with teachers. During this postconference, the supervisor and teacher discussed the data from the script taping in depth.

In short order, Hunter's seven elements of an effective lesson became the prescription for teacher evaluation in many states (Fehr, 2001, p. 175). If clinical supervision was the prescribed structure of supervision, Hunter's seven-step model, referred to as *mastery teaching*, became the content of the preconference, observation, and postconference. Teachers described their lessons in terms of Hunter's model, and supervisors determined the effectiveness of observed lessons in terms of alignment to the model.

The Era of Developmental/Reflective Models

By the mid-1980s, researchers and theorists in supervision began to articulate alternative perspectives, primarily in reaction to the prescription applications of clinical supervision and mastery teaching. William Glatthorn promoted supervisory models that considered a teacher's career goals. In *Differentiated Supervision*, Glatthorn (1984) explained that as professionals, teachers should have input and some sense of control over their development. Through differentiation, supervisors were expected to focus clinical supervisory practices on staff members who would derive the greatest benefit from a clinical approach. Additionally, different opportunities and venues for professional growth were to be provided for teachers based on their individual needs.

In a similar vein, Thomas McGreal (1983) delineated a range of supervisory options based on teacher experience. These options ranged from intensive developmental supervision for nontenured teachers and teachers with significant instructional deficiencies to more self-directed professional development for experienced staff. For evaluation purposes, McGreal recommended that teachers be placed either in an intensive evaluation program designed to make high-stakes decisions related to continued employment or granting of tenure, or in a standard evaluation program designed for quality assurance.

Another proponent of the differentiated approach to supervision during this era was Carl Glickman. In the first edition of his book *Supervision of Instruction: A Developmental Approach*, Glickman (1985) affirmed that the most important goal of supervision was to improve instruction. In the fourth edition of his book (1998), he described a number of related actions that constitute a robust approach to supervision. They included "(1) direct assistance to teachers, (2) group development, (3) professional development, (4) curriculum development, and (5) action research" (p. xv). Glickman noted that to implement a robust model of supervision, educators must take a systemic approach to the supervisory process: "By understanding how teachers grow optimally in a supportive and challenging environment, the supervisor can plan the tasks of supervision to bring together organizational goals and teacher needs into a single fluid entity" (1998, p. 10).

Clearly this era saw substantive arguments against the rigid applications of clinical supervision and mastery teaching. This era also set the stage for an emphasis on teacher evaluation.

The RAND Study

Amid the debates about the proper approach to supervision in the 1980s, the RAND group engaged in a study to determine what types of supervisory and evaluation practices were actually occurring in school districts across the United States. Its report, titled *Teacher Evaluation: A Study of Effective Practices* (Wise, Darling-Hammond, McLaughlin, & Bernstein, 1984), found that many of the systems of supervision and evaluation in place at this time were quite didactic and formulaic in nature. One general finding from the study was that the supervisory and evaluative approaches that were more developmental and reflective were sometimes viewed as not specific enough to enhance pedagogical development. Indeed, the report stated that teachers were the

strongest advocates for more standardized processes. "In their view, narrative evaluation provided insufficient information about the standards and criteria against which teachers were evaluated and resulted in inconsistent ratings among schools" (Wise et al., 1984, p. 16). The models in place in most of the 32 districts they studied were adopted or developed through committees of teachers, administrators, union representatives, and principals.

Four consistent problems with supervision and evaluation were also identified in the study. Nearly all respondents felt that principals "lacked sufficient resolve and competence to evaluate accurately" (Wise et al., 1984, p. 22). Teacher resistance to feedback was the second most identified problem. A key source of this resistance was related to the third most identified problem: a lack of uniform evaluation practices. The hypothesized reason for this concern was the fact that of the 32 districts in the study, only one district had a system built on a set of established teacher competencies. The fourth problem was a lack of training for evaluators. The study authors summarized their findings in four conclusions and 12 recommendations. These are reported in Figure 2.2.

Figure 2.2. Conclusions and Recommendations from the RAND Study

Conclusion	Recommendation
<p>"To succeed, a teacher evaluation system must suit the educational goals, management style, conception of teaching, and community values of the school district" (Wise et al., 1984, p. 66).</p>	<ul style="list-style-type: none"> • Examine goals and purpose of educational system and align system to those ends. • States should not adopt highly prescriptive systems (Wise et al., 1984).
<p>"Top-level commitment to and resource for evaluation outweigh checklists and procedures" (Wise et al., 1984, p. 67).</p>	<ul style="list-style-type: none"> • Provide administrators with adequate time for evaluations. • The quality of evaluation and ability of evaluators should be monitored. • Training for evaluators is important, particularly with new systems (Wise et al., 1984).
<p>"The school district must decide the main purpose of its</p>	<ul style="list-style-type: none"> • Examine current systems to

teacher evaluation system and then match the process to the purpose" (Wise et al., 1984, p. 70).	<p>determine and align with primary purpose.</p> <ul style="list-style-type: none"> • Consider adopting multiple systems if there are different purposes (Wise et al., 1984).
"To sustain resource commitments and political support, teacher evaluation must be seen to have utility. Utility depends on the efficient use of resources to achieve reliability, validity, and cost effectiveness" (Wise et al., 1984, p. 73).	<ul style="list-style-type: none"> • Allocate resources as aligned to importance of purpose. • Target resources to achieve maximum results (Wise et al., 1984).
"Teacher involvement and responsibility improve the quality of teacher evaluation" (Wise et al., 1984, p. 76).	<ul style="list-style-type: none"> • Involve expert teachers in the supervision and assistance of peers. • Involve teacher organizations in the development of processes and ongoing monitoring. • Hold teachers accountable for instructional decisions (Wise et al., 1984).

The Danielson Model

In 1996, a seminal work on supervision and evaluation was published by Charlotte Danielson. *Enhancing Professional Practice: A Framework for Teaching*, which was updated in 2007, was based on her work with the Educational Testing Service that focused on measuring the competence of preservice teachers. Given its past and current popularity, the Danielson model must be the reference point for any new proposals regarding supervision and evaluation. Whereas Hunter had described steps in the teaching process and Goldhammer and Cogan had done the same for the supervisory process, Danielson sought to capture—in its full complexity—the dynamic process of classroom teaching.

As we briefly described in Chapter 1, Danielson's model included four domains: Planning and Preparation, the Classroom Environment, Instruction, and Professional Responsibilities. Within each of these domains, she described a series of components that further articulate the knowledge, skills, and dispositions required to demonstrate competence in the classroom. According to Danielson (1996), the intent of the framework was to accomplish three things. First, it sought to honor the complexity of teaching. Second, it constituted a language for professional conversation. Third, it provided a structure for self-assessment and reflection on professional practice. The framework was considered comprehensive by Danielson in that it included all phases of teaching—from planning to reporting achievement. Additionally, Danielson noted that the model was

grounded in research and that it is generic or flexible enough to be used across levels and disciplines.

One of the more powerful aspects of the Danielson framework was that each of the 76 elements of quality teaching was broken into four levels of performance (unsatisfactory, basic, proficient, and distinguished). An example of one of these elements and the corresponding levels of performance is reported in Figure 2.3.

Figure 2.3. Component from Danielson's Model

DOMAIN 2: THE CLASSROOM ENVIRONMENT				
Component 2b: Establishing a Culture for Learning				
Element	Unsatisfactory	Basic	Proficient	Distinguished
Expectations for learning and achievement	Instructional outcomes, activities and assignments, and classroom interactions convey low expectations for at least some students.	Instructional outcomes, activities and assignments, and classroom interactions convey only modest expectations for student learning and achievement.	Instructional outcomes, activities and assignments, and classroom interactions convey high expectations for most students.	Instructional outcomes, activities and assignments, and classroom interactions convey high expectations for all students. Students appear to have internalized these expectations.
Source: From <i>Enhancing Professional Practice: A Framework for Teaching</i> (p. 69) by C. Danielson, 2007, Alexandria, VA: ASCD. Copyright 2007 by ASCD.				

The level of specificity supplied in the Danielson model provided the foundation for the most detailed and comprehensive approach to evaluation to that time.

The Beginning of the 21st Century

Since the turn of the 21st century, emphasis has shifted from supervision to evaluation, as well as from teacher behavior to student achievement. In their 2005 book *Linking Teacher Evaluation and Student Learning*, Tucker and Stronge championed the importance of student achievement as a

criterion in the evaluation process. Specifically, they argued for evaluation systems that determine

criterion in the evaluation process. Specifically, they argued for evaluation systems that determine teacher effectiveness using evidence from student gains in learning as well as observations of classroom instruction. To study how both of these components can be valued concurrently, they examined the supervisory systems in four different school districts that used data on instructional practices and learning gains. They drew a series of recommendations supporting the use of both types of data. However, their recommendations regarding the use of student achievement data were the most forcefully stated: "Given the clear and undeniable link that exists between teacher effectiveness and student learning, we support the use of student achievement information in teacher assessment. Student achievement can, and indeed should be, an important source of feedback on the effectiveness of schools, administrators, and teachers" (p. 102).

In 2008, Toch and Rothman's report *Rush to Judgment* provided a provocative perspective on teacher evaluation. They critiqued current supervisory and evaluative practices, saying they are "superficial, capricious, and often don't even directly address the quality of instruction, much less measure students' learning" (p. 1). Specifically, they described teaching as a profession that focuses on formal credentials rather than on instructional effectiveness and student achievement. Furthermore, despite No Child Left Behind requirements around teacher quality, they found only 14 states that required school systems to do annual evaluations of teachers. They noted that some evaluation systems may not even reflect teacher effectiveness in the classroom. Michigan State professor Mary Kennedy is quoted as saying, "in most instances, it's nothing more than marking satisfactory or unsatisfactory" (p. 2).

In 2009, a similar study entitled *The Widget Effect* (Weisberg, Sexton, Mulhern, & Keeling, 2009) heavily criticized teacher evaluation practices in the United States. The report authors explained its unusual name in the following way:

The failure of evaluation systems to provide accurate and credible information about individual teachers' instructional performance sustains and reinforces a phenomenon that we have come to call the **Widget Effect**. The Widget Effect describes the tendency of school districts to assume classroom effectiveness is the same from teacher to teacher. This decades-old fallacy fosters an environment in which teachers cease to be understood as individual professionals, but rather as interchangeable parts. In its denial of individual strengths and weaknesses, it is deeply disrespectful to teachers; in its indifference to instructional effectiveness, it gambles with the lives of students. (p. 4)

The *Widget Effect* was the product of research into the evaluation practices in 12 districts across four states including approximately 15,000 teachers, 1,300 administrators, and more than 80 local and state education officials. Specific findings indicated major flaws in the teacher evaluation process:

The failure to assess variations in instructional effectiveness also precludes districts from identifying specific development needs in their teachers. In fact, 73 percent of teachers surveyed

said their most recent evaluation did not identify any development areas, and only 45 percent of

teachers who did have development areas identified said they received useful support to improve. (p. 6)

Final conclusions from the report suggested a complete overhaul of the teacher evaluation process: Evaluations are short and infrequent (most are based on two or fewer classroom observations totaling 60 minutes or less), conducted by untrained administrators, and influenced by powerful cultural forces—in particular, an expectation among teachers that they will be among the vast majority rated as top performers.

While it is impossible to know whether the system drives the culture or the culture the system, the result is clear—evaluation systems fail to differentiate performance among teachers. As a result, teacher effectiveness is largely ignored. Excellent teachers cannot be recognized or rewarded, chronically low-performing teachers languish, and the wide majority of teachers performing at moderate levels do not get the differentiated support and development they need to improve as professionals. (p. 6)

Clearly, by the end of the first decade of the 21st century, teacher evaluation practices were under siege.

Lessons from History

The history of supervision and evaluation in this country can be viewed as a gradual evolution to the recommendations we make in this book. A well-articulated knowledge base for teaching is supported by the successes of the Hunter model and the utility of the Danielson model. Their specificity was their strength. However, as evidenced by the misuses of clinical supervision, history has taught us that a well-articulated knowledge base should not be used as a prescription for teaching or teacher evaluation. Focused feedback and practice are supported by the development of reflective supervisory models proposed by Glatthorn, McGreal, and Glickman.

True pedagogical development comes from teacher self-reflection that results in clear goals for improvement. Clear criteria for success that involve both teacher behavior and student achievement have roots in the emphases in the first decade of the 21st century on student achievement as the ultimate criterion for teacher effectiveness with teacher behavior as a causal factor. Finally, recognizing expertise is also supported by the emphasis on teacher evaluation in the first decade of the 21st century. If student achievement is not linked to teacher evaluation, teachers have little incentive to develop into experts.

The only aspect of our model that is not supported by the history of supervision and evaluation is providing opportunities to observe and discuss expertise. Probably the closest support for this aspect of our model is found in Glickman's perspective that supervision should be a systemic process. Teaching occurs within the context of a community; supervision and evaluation should be supported by that community.

Summary

This chapter presented a brief discussion of the history of teacher supervision and evaluation in the United States. The early days of supervision and evaluation began in the 1700s and lasted until the mid-1800s. They were characterized by a reliance on clergy to provide guidance to and supervision of teachers. As school systems became more complex, the need for more specialized guidance for teachers gave rise to the principal teacher as leader and a growing awareness of the importance of pedagogy. The era of scientific management, from the late 1800s until right before World War II, was characterized by two competing views of education. One was the view that the purpose of education was the promotion of democratic ideals. The other was the view that schools function best when approached from the perspective of scientific management.

Throughout this era, the scientific approach gained strength and acceptance. The period after World War II saw a swing away from the scientific approach to an emphasis on developing the teacher as an individual. This period also saw a proliferation of the responsibilities of the supervisor.

The next era, lasting from the late 1960s to the early 1970s, saw the phenomenon of clinical supervision—one of the most influential movements in supervision and evaluation. The Hunter model was combined with clinical supervision to produce a widely used but oftentimes prescriptive approach to supervision. This period was followed by developmental/reflective models that were much less prescriptive. The RAND study provided a realistic look at the actual practice of supervision and evaluation in districts and schools and concluded that teachers preferred specific as opposed to general feedback.

The mid-1990s saw the introduction of the Danielson model to teacher supervision and evaluation. It was widely applied through K–12 education. Finally, the first decade of the 21st century witnessed heavy criticisms of current evaluation practices calling for major changes in tenure and compensation.