Chapter 11
Best Practice Works: The Proof and the Pendulum

It's easy to become cynical and view new (or revitalized) ideas as mere fads, the latest swing of the never-ending pendulum of reform strategies pushed by the latest set of academics trying to build their careers. But as we've shown throughout this book, Best Practice is far more than that. It represents a broad consensus across the subject areas about what is important for children to learn and how they can best learn it. Best Practice draws on a long history of inquiry into the psychology and dynamics of learning. It's confirmed by research and testing that shows greater student achievement and learning as a result of Best Practice strategies. And it reflects the concrete everyday experience of talented teachers across the country.

Nevertheless, advocating for Best Practice approaches—as well as refining and further advancing our understanding of them—requires that we collect the best data possible. We need a steady flow of reliable information—data about individual students, data about whole schools, and research on the effectiveness of particular educational strategies. And we need this data not just to advance our own argument. Rather, each child and each school embodies a fresh human experience. No educational theory or plan or curriculum, however strongly it's shown to be effective, is automatically a success. It's a tool in the hands of thinking people who must always figure out how to make it work with these children and these teachers in this particular culture and setting.

So as we conclude this book, we need to reflect on ways to provide ourselves with the best information on how Best Practice strategies are working with particular kids and schools—and the best research that tests their effectiveness in general. It should be information that not only helps us understand if something is effective, but also what is needed if it is not.
DATA ON STUDENTS

To be very direct about it, we think teachers and schools evaluate students badly, unfairly, and far too much. American schoolchildren, teachers, parents, taxpayers, politicians, and policymakers are downright fixated on grades and tests. Everyone involved spends far more time worrying about test scores than thinking up ways to increase student learning, which might actually raise achievement scores. We are a country full of measurement-obsessed people who seem to believe that you can raise the temperature by improving your thermometer.

What’s so bad about the way we evaluate kids in school? To begin with, the socioeconomic function of evaluation in American education has always been problematic and unsavory. Grading and testing historically have been harnessed to the screening, sorting, and classifying of children into categories of “merit” or “intelligence.” These certified categories of students are then allocated certain current or future rewards, such as school prizes, invitations to honors classes, admission to good colleges, or entry to high-paying careers. As scholars like Alfie Kohn (1993), Michael Katz (1968), and Joel Spring (1972) have convincingly shown for decades, this vaunted American meritocracy is largely a sham. School tests and grades are part of a system that camouflages the replication of the existing social hierarchy: kids from wealthy, culturally mainstream homes are certified by schools as “deserving” rewards, while students from poor, culturally different homes are proven by tests and grades to “need” a vocational education or to be “unable to benefit” from a college preparatory program.

The school grading system has been abused, co-opted, and enlisted in the service of some shamefully undemocratic arrangements in our culture. Even today, after the wide distribution of work by social historians and the many exposés of standardized test bias, our two most famous educational exams—the Scholastic Aptitude Test and the American College Test—brazenly continue to show a near-perfect correlation between family income and score levels, and still deliver a huge score penalty for being African American or Hispanic. As the New York Times once pointed out, Martin Luther King, “a man who is now viewed as among the nation’s greatest orators ever, was in the third quartile or below average [in verbal aptitude on the Graduate Records Exam]. His quantitative score was in the bottom 10 percent, and he was in the bottom quarter for tests in physics, chemistry, biology, social studies, and the fine arts” (November 8, 1997, p. A8).
But even if we admit that educational assessment has been often misused by the society at large, don’t we use it more responsibly within the institution of school? Sadly, most teachers are still wedded to evaluation procedures that are ineffective, time-consuming, and hurtful to students. One example from the field of writing is particularly illustrative. Everyone is familiar with the deep-rooted school tradition called “intensive correction,” where the teacher marks every error in every paper that every student ever writes. Indeed, in American schools this practice is often considered to be the basic, standard treatment for responding to student writing. But George Hillocks’ meta-analysis of research (1986) showed that such intensive correction is completely useless. Marking all the errors in a student paper is no more effective, in terms of future growth or improvement, than marking none of them. The only difference is the huge expenditure of teacher time and the student demoralization that accompany this practice.

Under pressure to “justify” grades with copious scores and marks in their gradebooks, teachers devote enormous energy to feeding the grading machine—finding ways to quantify, measure, score, compute, and record assorted aspects of kids’ behaviors. Coauthor Harvey Daniels has worked to popularize the classroom structure called “literature circles” (2002) and reports that in workshops, the number-one question raised by teachers is: How can I get a grade out of this activity? This frantic quest for quantification should remind us that the main legitimate purpose of evaluation in education ought to be to guide instruction, not to rack up numbers. Anything we do to gather and interpret information about kids’ learning should provide accurate, helpful input for nurturing children’s further growth.

Back in their preservice educational psychology courses, teachers learned the distinction between summative and formative evaluation. Formative evaluation is the basic, everyday kind of assessment that we continually use to understand students’ growth and help “form” their further learning. Summative evaluation doesn’t aim to nurture learning at all, but merely quantifies what has been learned up to a given point, translating it into a score or symbol that allows students to be ranked against each other. Summative evaluation isn’t actually educational; it is just a way of reporting periodically to outsiders about what has been studied or learned.

It’s problem enough when tests are valid but overused or misused, but it’s much worse when they don’t reflect Best Practice teaching or the skills that the kids are, in fact, learning. Yet over and over, even when communities and news reporters are confronted with studies showing that standardized tests
are inaccurate, discriminate against minorities, and reward memorization instead of thinking, school districts opt to retain them anyway. Better, more complex tests are too expensive, they conclude, and besides, how can we compare ourselves to other communities if we don't use the same tests?

Meanwhile, we simply evaluate kids' work too much. We have a norm of grading every piece of work that students ever attempt in any school subject, duly placing a carefully computed number in the gradebook after each attempt. Indeed, in many classrooms the compulsion to evaluate every piece of student work actually becomes an instructional bottleneck, limiting the amount of student practice to a level that the teacher has time to grade. The sad irony here, of course, is that practice—unmonitored practice—is the main way in which humans learn almost every valuable activity in life, from piano playing to roof shingling.

We grade and test and score kids far more than is needed to effectively guide instruction—and ironically, we too often fail to use the data to actually guide the successive help we provide for individual students. In classrooms where teachers are constantly watching, talking, and working with kids, elaborate grading systems are unnecessary, unhelpful, redundant, and sometimes contradictory. As far as the demand for official grades and records is concerned, teachers can produce a perfectly adequate documentation of students' growth through the occasional sampling of their work, periodic observations, and once-in-a-while examination of their products. Especially when records are backed up by a portfolio of students' actual work—the raw material upon which any grade ought to be based—there should be no problem in explaining a given grade. When teachers make this change, substituting descriptive evaluation for grading, they are essentially making a trade: they are swapping time previously spent on scoring, computing, recording, averaging, and justifying grades, in exchange for time to collect, save, discuss, and reflect on kids' real work.

However, in the current climate of preoccupation with test scores and mistrust of teachers and schools, teachers are understandably susceptible to the evaluation obsession themselves. When we lead a workshop with teachers, we always begin by asking what concerns or topics they would like to discuss during the course. Evaluation is usually the first topic to be mentioned, and it almost always ends up being listed as a top priority of every group. To be fair, part of this obsession simply reflects the pressure that teachers feel from the public, taxpayers, the media, state assessments, and so forth. But it also reveals that teachers are, finally, just another group of
Americans—and they have acquired the evaluation fixation just as deeply, and in much the same way, as any other citizens.

But let's return to the bright side. As teachers and schools move toward Best Practice, there is a clear mandate for new forms of assessment, evaluation, grading, and reporting student progress. Across subject fields, Best Practice in evaluation means that teachers assess student learning according to the principles outlined below.

**Best Practice Evaluation**

- focuses on the knowledge and abilities that are key to Best Practice learning, and on complex whole outcomes and performances of writing, reading, researching, and problem solving, rather than only on isolated subskills
- most of the time, uses assessment that is formative, not summative—and then applies the data to guide individual students’ further learning and to adjust our own teaching
- employs evaluation that is descriptive or narrative, not scored and numerical
- involves students in developing meaningful assessments (for example, asks students to describe what makes a good research report), and will call on them to keep track of and judge their own work
- triangulates assessments, looking at each child from several angles, by drawing on observation, conversation, artifacts, and performances, and by looking at learning over time
- operates as a part of instruction (as in teacher-student conferences), rather than separate from it
- allots a moderate amount of time to evaluation and assessment, not allowing it to rule a teacher's professional life or consume lots of instruction time
- where possible, abolishes or de-emphasizes competitive grading systems
- employs parent-education programs to help community members understand the value of new approaches, and then invites parents to participate in the process
DATA ON SCHOOLS

The coin of the realm nowadays is the state-mandated standardized test. Schools that do well or improve markedly are praised and those that don’t are threatened with sanctions or even state takeover (though the latter puts us in mind of the dog that finally catches the neighbor’s car he’s been chasing for years—now what does he do with it?). We’ve discussed throughout this book the research that unequivocally shows how Best Practice teaching strategies actually improve student performance on such tests. But that doesn’t mean we find the tests particularly accurate or meaningful in judging a school’s overall performance, or likely to encourage real improvement. In Chicago, for example, high schools are judged on the number of students scoring at or above the 50th percentile on Illinois’ “Prairie State” test, which combines the ACT and a test called “Work Keys.” Using the test this way means that if your faculty work like demons and improve all their lowest-scoring kids from the 10th to the 49th percentile, they get exactly no credit for their effort; but if you move a bunch of kids from the 49th to the 50th, you look like a miracle worker. You can easily guess which path a principal who is under the microscope is likely to choose, which students will get extra attention and which are likely to be, uh, left behind.

In this atmosphere, it is important for schools to collect and publish data of other kinds, to show parents and the community what they are really accomplishing. Figures on attendance, dropouts versus kids who stay in school, graduation rate, rate of college attendance for high schools, rate of success in college (or success in high school, for graduates of elementary schools)—all reflect important aspects of a school’s performance. Schools that use end-of-semester or yearly performance assessments (with jury panels that include community members) can calculate rates of acceptable and high performance and compare them year to year. Regularly publicizing this information can gradually give it credibility and recognition in the community. We know schools that track the number of books and even the number of pages that students read over the course of a year. While this raw data doesn’t tell what the student learned, it’s still valuable information, compared to the school that doesn’t talk publicly about reading at all.

It’s also important for schools to seek meaningful program evaluation—in other words, to conduct thoughtful reflection on just what sort of teaching is in fact taking place, and whether the school is effectively enacting the programs it has committed to. After all, we’ve seen so many cases in our profession in which a reform was supposedly “tried” and pronounced a failure,
while proponents maintain that it wasn’t well enough executed to really determine its efficacy.

One traditional mechanism for program evaluation is the principal’s evaluation of individual teachers. But even at its best, this process does not provide a larger whole-school picture of the teaching and learning that is taking place. At the high school level, the regional accrediting agencies around the country make their once-every-five-year visits, but we know of few schools that have truly rethought their curriculum as a result.

One valuable effort to address this has been the “critical friends” strategy. We wish it bore a more felicitous label because the phrase seems uncomfortably self-contradictory and negative. Nevertheless, it’s a powerful approach that can both help schools improve and inform people about where they are in the effort. Actually, there are two forms of critical friends work, one internal and the other external. The internal version is a sophisticated form of self-guided reflection: teachers in a school meet regularly and follow a protocol to share and examine their practice (Bambino 2002; Dunne et al. 2000). In the external version, a visiting team provides outside eyes and ears for this process. Both versions receive high praise from teachers who have participated in them.

We’ll focus for a moment here on the external version, since it provides more of an explicit report on the school’s work. When conducted well, a visiting team asks the school’s teachers and administrators to describe their goals and the programs and strategies they are using. Then they spend several days observing classrooms, examining student work, and interviewing kids, teachers, and parents. Finally, the team provides a written report and meets with the school staff to discuss their observations. As a Coalition of Essential Schools guide for this work explains, “The purpose of our evidence-gathering is not to compare or evaluate, but to provide the school with accurate information they can use to decide ‘next steps’ in their quest to become an excellent school” (Cushman 1998). Clearly, schools that employ such self-reflection can readily demonstrate to the community that whatever the standardized test scores, the school is serious about strengthening teaching and learning for all students.

DATA ON BEST PRACTICE APPROACHES

In many places in this book, we’ve referred to the research that supports Best Practice, as well as the controversy that has arisen around at least some of that research. We’ve described some key issues in the Reading Wars, which
Best Practice Works: The Proof and the Pendulum • 319

focus on the narrow question of how to best help beginning readers. We’ve noted how proponents of differing approaches focus their research on differing aspects of reading or different kinds of reading tests. And we’ve referred to the important research by Fred Newmann and his associates showing that authentic assignments, interactive teaching, and strong social support in the classroom all lead to higher standardized test scores (see Chapters 1 and 10).

We want to round out the picture by recalling that there is a long history of such research, covering a wide range of Best Practice strategies. Those of us who work with the research understand that all of it is limited because so many variables are beyond control in any classrooms studied, and we don’t have much faith in many of the measures for studying them. However, in spite of the more politically motivated statements sometimes heard, the dozens of research studies done and repeated over many years have confirmed over and over the positive outcomes for a wide range of progressive teaching strategies compared to more traditional ones.

For years, for example, cooperative education researchers have published studies confirming significant achievement gains in a wide range of content areas when classrooms include ample cooperative activity—one of the fundamental components of the Best Practice paradigm (Johnson, Johnson, and Holubec 1998). In the teaching of writing, a meta-analysis of numerous statistical studies by George Hillocks (1986) showed that while activities for engaging students with material and ideas for writing have a strong effect on writing quality, lecture presentations and grammar drills are of little use, or even bring down writing scores.

Since reading is the most heavily researched area of the curriculum, we’ll briefly list the books and articles that review this vast literature. The learning principles are very similar for other subjects, as this entire book has illustrated. A decades-long body of research on literature-based reading programs shows standardized achievement score gains for students in progressive programs, not just in regular education but among students with ESL, special education, or disadvantaged backgrounds. Michael Tunnell and James Jacobs, surveying studies up to 1989, listed one after another that reached the same conclusion.

Jane Braungar and Jan Lewis compiled a well-organized account of strategies for teaching reading that are supported by research ranging from the 1970s to the mid 1990s. The studies covered in their book, Building a Knowledge Base in Reading (1997), confirm that reading is a constructive, active process that involves complex thinking. In 1999, Harvey and Steve,
with their partner, Marilyn Bizar, published a sixty-year review of research supporting literature-based teaching of reading, use of independent reading, teaching of the writing process, and other related approaches. Richard Allington’s *What Really Matters for Struggling Readers* (2001) marshals the research that supports a simple and clear list of instructional needs:

- Kids need to read a lot.
- Kids need books they *can* read.
- Kids need to learn to read fluently.
- Kids need to develop thoughtful literacy.

Alan Farstrup and Jay Samuels’ thorough guide, *What Research Has to Say About Reading Instruction* (2002), presents a range of reading experts who explain the implications of research for teaching reading. The central role of best practice strategies as described in the IRA standards is very clear. Michael Pressley’s chapter, for example, examines the importance of above-the-word-level comprehension strategies for good reading. Constance Weaver and coauthors created a very handy research guide in *Creating Support for Effective Literacy Education* (1996), with brief summaries and bibliographies of the studies on topics such as teaching skills in context, teaching phonics, and spelling. However, the third edition of Weaver’s *Reading Process and Practice* (2002) is now more up to date, and is encyclopedic in its coverage of the concepts behind a progressive approach to reading and the research that supports each one. Finally, Daniels and Zemelman’s *Subjects Matter: Every Teacher’s Guide to Content-Area Reading* (2004) contains a summary of research implications for reading in the secondary grades.

But what about the research behind the arguments from the so-called other side in the Reading Wars? The National Reading Panel (2000) has supported a much more conservative view, particularly by defining the acceptability of research so narrowly that many important studies are left out, thereby emphasizing the importance of phonics and direct instruction. Other researchers, like Richard Allington (2001), explain the flaws in this approach. And yet even with all this debate, many of the panel’s recommendations firmly support the Best Practice approaches described in this book. The panel affirms that the best research roundly endorses the learning of vocabulary “in rich contexts,” the importance of self-monitoring of comprehension, the value of cooperative learning, and the effectiveness of students generating their own questions.
We recognize that no amount of quoted research is likely, by itself, to change educational practice in American schools. Here in Chicago, for example, a major scientific study showed that the effects of retaining students at a grade level were clearly negative (Nagaoka and Roderick 2004), and yet school officials here (and in other cities, like New York) only reaffirmed their commitment to the strategy. Nevertheless, responsible educators want to learn all that they can about what works in their profession, and need to know that they are on the right track. When it comes to Best Practice approaches, we are clearly reassured. And even the schools that try narrower, more traditional strategies frequently end up rethinking and asking for help with new ideas a few years later.

FADS?

Sometimes teachers wonder: Isn't Best Practice just another educational fad? And is any of this stuff really new? It sounds just like the open classrooms of the 1960s or Whole Language in the early 1990s, and those trends died away, didn't they? What we are calling the Best Practice movement certainly has a very familiar ring to school veterans from the past thirty years. It's not surprising that earlier waves of progressive reform are scoffed at by some educators, just as newer incarnations are attacked. But we think the educational innovations of the late 1960s and early 1990s were important precursors to today's developments, and it's worth understanding what really happened back then, to comprehend our struggles now.

Let's look at one controversial innovation from this era to help us revalue our heritage. In the 1970s, some American schools plunged into the open-classroom experiment, tearing down walls and offering teachers a one-day in-service at the end of the summer, in which some administrator or outside consultant essentially announced: "OK, next week we want you to throw away the one model of teaching that you were trained for and are experienced in, and instead run your classroom in ten other ways you've never tried and we've never trained you for. Have a nice year." The great open-education movement inevitably collapsed because teachers did not have in their professional repertoire the structures and strategies to run a variety of student-directed, independent, and small-group activities. Kids went bananas, and many teachers who were working in huge rooms with hundreds of kids started sneaking in cardboard boxes, shelves, and other large objects with which they could gradually rebuild a classroomlike space inside the trackless waste of the "pod."
But if you visited Joanne Trahanas’ eighth-grade language arts class in Glenview, Illinois, in recent years (as a person who’s always growing, she’s now moved on to become a principal), you’d have seen something that looks very much like an open classroom. You’d have seen thirteen-year-olds—an age group not usually noted for self-discipline—working industriously and without overt supervision on their writing skills every day. They’d come into Joanne’s room, get out their writing folders, and get ready to work. During a quick round of the class, each kid would announce what he or she was planning to work on that day, and then everyone, including the teacher, would go to work. Some were drafting new pieces, others were editing ongoing drafts. A few would quietly seek out a partner for a quick conference. After a while, Joanne would start to see kids one at a time on a schedule she kept, and as she conferred with each student about his or her writing, she would jot a phrase or two in the book where she tracked progress. A couple of times a week, the group would gather for ten minutes to read passages aloud and discuss the progress of their writing.

This was an open classroom—an open classroom that worked. It worked because Joanne set up the structure, the norms, the schedule, the procedures, and the materials in such a way that kids quickly grew into responsible use of this special time and space. What’s even more impressive to us is that Joanne had six writing workshops each day, with six different sets of kids. Joanne’s workshop reminds us that there are important links between contemporary Best Practice ideas and past innovations, but it also demonstrates that we’ve learned a lot about how to make things really stick this time around.

And here’s some more history that helps us understand how change works, or doesn’t, in American schools. All those open classrooms of the early seventies were doing open education wrong in the first place. In borrowing the idea from British education (where it was more commonly called “integrated day”), many American educators misconstrued open education as an architectural rather than an instructional innovation, believing that the process required big open spaces. They focused on creating huge spaces rather than helping teachers orchestrate the inquiry-oriented, experiential learning that was the true core of the model. Harvey remembers taking one English educator through Chicago’s first gigantic open-classroom school in 1972, and the visitor shaking his head in amazement: “They’ve got this completely wrong,” he moaned.

We’ll close with a happier but equally important history lesson. At Barton Open School in Minneapolis, progressive reformers got it right the
first time. They did not tear down walls or build vast “pods.” Instead, inside normal-sized classrooms, they set up spaces where kids’ interests were piqued, where exploration and expression were valued, where individual differences were prized, and where teachers could operate as thoughtful facilitators of learning. As a result of this faithful implementation, Barton is still there, still thriving, and has become one of the most celebrated schools in the Midwest. Today, when we visit Barton, we see a Best Practice school in the truest sense. Principal Steve DeLapp and his amazing staff have simply been doing progressive education steadily and genuinely for more than thirty years. No pendulums have swung through Barton.

**DOWN WITH THE PENDULUM METAPHOR**

One tendency we have pretty successfully avoided in this book is quoting ourselves. But on this subject, we are unlikely to improve on what we wrote in 1988 in *A Community of Writers*:

A pendulum swinging back and forth . . . seems to be the standard metaphor for changes in our field. How impartial and content free this pendulum image is, tempting us to believe that the fluctuations in educational practice are merely the results of some pointless, random, eternal variation. In fact, this is no impartial pendulum swinging; it is more like the battlefront in a war that moves back and forth with assaults and retreats. This is a historical struggle of one set of ideas against another, continually being fought out in close relation to the social-political-economic issues outside of schools . . . the playing out, over a huge span of time, of a war for the soul of schooling in our society. In the end, the student-centered, humanistic, developmental approach will win out over the authoritarian model because it parallels the direction in which civilization itself progresses. If we look broadly enough, we can see evidence that this direction is already well established: in matters of discipline, teachers no longer whip students in school or crown them with dunce caps; in language arts, we no longer require endless copying of great authors’ texts; we no longer have kids stand beside their desks for recitation and enunciation; and sentence diagramming, though far from dead, now occupies far less of the average school day than it did just one generation ago.

Of course there will always be regressions and short-term backslides. Change in schools never follows a straight, steady path, but is more like
three steps forward and two-and-a-half steps back. Perhaps tomorrow a
movement will spring up to restore the teaching of sentence diagramming
to its “rightful, central place in the English curriculum,” and such a trend
might even catch on for a few years. Indeed, it is just this sort of event that
misleads us, as individuals living in a particular brief here and now, into be-
lieving that there’s always a pendulum swinging back and forth between
two eternal, fixed points, ensuring that nothing ever really changes.
(Zemelman and Daniels 1988, 17, 270)

But each time the progressive set of ideas comes back, it gains strength and
coherence from the new research and practice that connects with it, and each
time it appears, it exerts more influence on the schools. Best Practice, class-
room workshop, interdisciplinary studies—if taught in their true, genuine
forms, adaptable to the needs of the students involved in them—reflect a set
of deep educational ideas. Ideas that are partisan, yes—but they are deeply
rooted in an ever-deepening understanding of how human beings learn, and
so they will continue to spread. In the words of a courageous Chicago
teacher, Joe Perlstein:

It’s very difficult to change. It takes a lot out of a person. However, you
don’t mind if there’s a payback, if you feel that the children are growing,
that they appreciate what is going on. If you see a light at the end of the
tunnel, then you say to yourself, “Don’t stop now. Keep pushing, keep
pressing.” Because it will all be worth it in the end.

WORKS CITED

Research-Based Programs. New York: Addison Wesley.
Braunger, Jane, and Jan Lewis. 1997. Building a Knowledge Base in Reading. Port-
land, OR: Northwest Regional Educational Laboratory.
Cushman, Kathleen. 1998. How Friends Can Be Critical as Schools Make Essential
Changes. Horace 14:5.
Daniels, Harvey. 2001. Literature Circles: Voice and Choice in Book Clubs and
Reading Groups. York, ME: Stenhouse.
Daniels, Harvey, and Steven Zemelman. 2004. Subjects Matter: Every Teacher’s


National Reading Panel. 2000. *Teaching Children to Read: An Evidence-Based Assessment of the Scientific Research Literature on Reading and Its Implications for Reading Instruction*. Washington, DC: National Institutes of Health.


