

# The Case Against Grades

EDUCATIONAL LEADERSHIP November 2011

The Case Against Grades

By Alfie Kohn

*[This is a slightly expanded version of the published article.]*

"I remember the first time that a grading rubric was attached to a piece of my writing... Suddenly all the joy was taken away. I was writing for a grade — I was no longer exploring for me. I want to get that back. Will I ever get that back?" — Claire, a student (in Olson, 2006)

By now enough has been written about academic assessment to fill a library, but when you stop to think about it, the whole enterprise really amounts to a straightforward two-step dance. We need to collect information about how students are doing, and then we need to share that information (along with our judgments, perhaps) with the students and their parents. Gather and report — that's pretty much it.

You say the devil is in the details? Maybe so, but I'd argue that too much attention to the particulars of implementation may be distracting us from the bigger picture -- or at least from a pair of remarkable conclusions that emerge from the best theory, practice, and research on the subject: *Collecting information doesn't require tests, and sharing that information doesn't require grades*. In fact, students would be a lot better off without either of these relics from a less enlightened age.

Why tests are not a particularly useful way to assess student learning (at least the kind that matters), and what thoughtful educators do instead, are questions that must wait for another day. Here, our task is to take a hard look at the second practice, the use of letters or numbers as evaluative summaries of how well students have done, regardless of the method used to arrive at those judgments.

## **The Effects of Grading**

Most of the criticisms of grading you'll hear today were laid out forcefully and eloquently anywhere from four to eight decades ago (Crooks, 1933; De Zouche, 1945; Kirschenbaum, Simon, & Napier, 1971; Linder, 1940; Marshall, 1968), and these early essays make for eye-opening reading. They remind us just how long it's been clear there's something wrong with what we're doing as well as just how little progress we've made in acting on that realization.

In the 1980s and '90s, educational psychologists systematically studied the effects of grades. As I've reported elsewhere (Kohn, 1999a, 1999b, 1999c), when students from elementary school to college who are led to focus on grades are compared with those who aren't, the results support three robust conclusions:

\* *Grades tend to diminish students' interest in whatever they're learning.* A "grading orientation" and a "learning orientation" have been shown to be inversely related and, as far as I can tell, every study that has ever investigated the impact on intrinsic motivation of receiving grades (or instructions that emphasize the importance of getting good grades) has found a negative effect.

\* *Grades create a preference for the easiest possible task.* Impress upon students that what they're doing will count toward their grade, and their response will likely be to avoid taking any unnecessary

intellectual risks. They'll choose a shorter book, or a project on a familiar topic, in order to minimize the chance of doing poorly — not because they're "unmotivated" but because they're rational. They're responding to adults who, by telling them the goal is to get a good mark, have sent the message that success matters more than learning.

\* *Grades tend to reduce the quality of students' thinking.* They may skim books for what they'll "need to know." They're less likely to wonder, say, "How can we be sure that's true?" than to ask "Is this going to be on the test?" In one experiment, students told they'd be graded on how well they learned a social studies lesson had more trouble understanding the main point of the text than did students who were told that no grades would be involved. Even on a measure of rote recall, the graded group remembered fewer facts a week later (Grolnick and Ryan, 1987).

Research on the effects of grading has slowed down in the last couple of decades, but the studies that are still being done reinforce the earlier findings. For example, a grade-oriented environment is associated with increased levels of cheating (Anderman and Murdock, 2007), grades (whether or not accompanied by comments) promote a fear of failure even in high-achieving students (Pulfrey et al., 2011), and the elimination of grades (in favor of a pass/fail system) produces substantial benefits with no apparent disadvantages in medical school (White and Fantone, 2010). More important, no recent research has contradicted the earlier "big three" findings, so those conclusions still stand.

### **Why Grading Is Inherently Problematic**

A student asked his Zen master how long it would take to reach enlightenment. "Ten years," the master said. But, the student persisted, what if he studied very hard? "Then 20 years," the master responded. Surprised, the student asked how long it would take if he worked very, *very* hard and became the most dedicated student in the Ashram. "In that case, 30 years," the master replied. His explanation: "If you have one eye on how close you are to achieving your goal, that leaves only one eye for your task."

To understand why research finds what it does about grades, we need to shift our focus from educational measurement techniques to broader psychological and pedagogical questions. The latter serve to illuminate a series of misconceived assumptions that underlie the use of grading.

*Motivation:* While it's true that many students, after a few years of traditional schooling, could be described as motivated by grades, what counts is the nature of their motivation. Extrinsic motivation, which includes a desire to get better grades, is not only different from, but often undermines, intrinsic motivation, a desire to learn for its own sake (Kohn 1999a). Many assessment specialists talk about motivation as though it were a single entity — and their recommended practices just put a finer gloss on a system of rewards and punishments that leads students to chase marks and become less interested in the learning itself. If nourishing their *desire* to learn is a primary goal for us, then grading is problematic by its very nature.

*Achievement:* Two educational psychologists pointed out that "an overemphasis on assessment can actually undermine the pursuit of excellence" (Maehr and Midgley, 1996, p. 7). That unsettling conclusion — which holds regardless of the quality of the assessment but is particularly applicable to the use of grades -- is based on these researchers' own empirical findings as well as those of many others, including Carol Dweck, Carole Ames, Ruth Butler, and John Nicholls (for a review, see Kohn 1999b, chapter 2). In brief: the more students are led to focus on *how well* they're doing, the less engaged they tend to be with

*what* they're doing.

It follows that all assessment must be done carefully and sparingly lest students become so concerned about their achievement (how good they are at doing something -- or, worse, how their performance compares to others') that they're no longer thinking about the learning itself. Even a well-meaning teacher may produce a roomful of children who are so busy monitoring their own reading skills that they're no longer excited by the stories they're reading. Assessment consultants worry that grades may not accurately reflect student performance; educational psychologists worry because grades fix students' attention *on* their performance.

*Quantification:* When people ask me, a bit defensively, if it isn't important to measure how well students are learning (or teachers are teaching), I invite them to rethink their choice of verb. There is certainly value in *assessing* the quality of learning and teaching, but that doesn't mean it's always necessary, or even possible, to *measure* those things — that is, to turn them into numbers. Indeed, "measurable outcomes may be the least significant results of learning" (McNeil, 1986, p. xviii) — a realization that offers a refreshing counterpoint to today's corporate-style "school reform" and its preoccupation with data.

To talk about what happens in classrooms, let alone in children's heads, as moving forward or backward in specifiable degrees, is not only simplistic because it fails to capture much of what is going on, but also destructive because it may change what is going on for the worse. Once we're compelled to focus only on what can be reduced to numbers, such as how many grammatical errors are present in a composition or how many mathematical algorithms have been committed to memory, thinking has been severely compromised. And that is exactly what happens when we try to fit learning into a four- or five- or (heaven help us) 100-point scale.

*Curriculum:* "One can have the best assessment imaginable," Howard Gardner (1991, p. 254) observed, "but unless the accompanying curriculum is of quality, the assessment has no use." Some people in the field are candid about their relativism, offering to help align your assessment to whatever your goals or curriculum may be. The result is that teachers may become more adept at measuring how well students have mastered a collection of facts and skills whose value is questionable — and never questioned. "If it's not worth teaching, it's not worth teaching well," as Eliot Eisner (2001, p. 370) likes to say. Nor, we might add, is it worth assessing accurately.

Portfolios, for example, can be constructive if they replace grades rather than being used to *yield* them. They offer a way to thoughtfully gather a variety of meaningful examples of learning for the students to review. But what's the point, "if instruction is dominated by worksheets so that every portfolio looks the same?" (Neill et al. 1995, p. 4). Conversely, one sometimes finds a mismatch between more thoughtful forms of pedagogy -- say, a workshop approach to teaching writing — and a depressingly standardized assessment tool like rubrics (Wilson, 2006).

### **Improving Grading: A Fool's Errand?**

*"I had been advocating standards-based grading, which is a very important movement in its own right, but it took a push from some great educators to make me realize that if I wanted to focus my assessment around authentic feedback, then I should just abandon grades altogether."*

— New Jersey middle school teacher Jason Bedell (2010)

Much of what is prescribed in the name of "assessing for learning" (and, for that matter, "formative assessment") leaves me uneasy: The recommended practices often seem prefabricated and mechanistic; the imperatives of data collection seem to upstage the children themselves and the goal of helping them become more enthusiastic about what they're doing. Still, if it's done only occasionally and with humility, I think it's possible to assess for learning. But *grading* for learning is, to paraphrase a 1960's-era slogan, rather like bombing for peace. Rating and ranking students (and their efforts to figure things out) is inherently counterproductive.

If I'm right — more to the point, if all the research to which I've referred is taken seriously — then the absence of grades is a necessary, though not sufficient, condition for promoting deep thinking and a desire to engage in it. It's worth lingering on this proposition in light of a variety of efforts to sell us formulas to improve our grading techniques, none of which address the problems of grading, per se.

\* It's not enough to replace letters or numbers with labels ("exceeds expectations," "meets expectations," and so on). If you're sorting students into four or five piles, you're still grading them. Rubrics typically include numbers as well as labels, which is only one of several reasons they merit our skepticism (Wilson, 2006; Kohn, 2006).

\* It's not enough to tell students in advance exactly what's expected of them. "When school is seen as a test, rather than an adventure in ideas," teachers may persuade themselves they're being fair "if they specify, in listlike fashion, exactly what must be learned to gain a satisfactory grade... [but] such schooling is unfair in the wider sense that it prepares students to pass other people's tests without strengthening their capacity to set their own assignments in collaboration with their fellows" (Nicholls and Hazzard, 1993, p. 77).

\* It's not enough to disseminate grades more efficiently -- for example, by posting them on-line. There is a growing technology, as the late Gerald Bracey once remarked, "that permits us to do in nanoseconds things that we shouldn't be doing at all" (quoted in Mathews, 2006). In fact, posting grades on-line is a significant step backward because it enhances the salience of those grades and therefore their destructive effects on learning.

\* It's not enough to add narrative reports. "When comments and grades coexist, the comments are written to justify the grade" (Wilson, 2009, p. 60). Teachers report that students, for their part, often just turn to the grade and ignore the comment, but "when there's only a comment, they read it," says high school English teacher Jim Drier. Moreover, research suggests that the harmful impact of grades on creativity is no less (and possibly even more) potent when a narrative accompanies them. Narratives are helpful only in the absence of grades (Butler, 1988; Pulfrey et al, 2011).

\* It's not enough to use "standards-based" grading. That phrase may suggest any number of things -- for example, more consistency, or a reliance on more elaborate formulas, in determining grades; greater specificity about what each grade signifies; or an increase in the number of tasks or skills that are graded. At best, these prescriptions do nothing to address the fundamental problems with grading. At worst, they exacerbate those problems. In addition to the simplistic premise that it's always good to have more data, we find a penchant shared by the behaviorists of yesteryear that learning can and should be broken down into its components, each to be evaluated separately. And more frequent temperature-taking produces exactly the kind of disproportionate attention to performance (at the expense of learning) that

researchers have found to be so counterproductive.

The term "standards-based" is sometimes intended just to mean that grading is aligned with a given set of objectives, in which case our first response should be to inquire into the value of those objectives (as well as the extent to which students were invited to help formulate them). If grades are based on state standards, there's particular reason to be concerned since those standards are often too specific, age-inappropriate, superficial, and standardized by definition. In my experience, the best teachers tend to be skeptical about aligning their teaching to a list imposed by distant authorities, or using that list as a basis for assessing how well their students are thinking.

Finally, "standards-based" may refer to something similar to criterion-based testing, where the idea is to avoid grading students on a curve. (Even some teachers who don't do so explicitly nevertheless act as though grades ought to fall into something close to a normal distribution, with only a few students receiving As. But this pattern is not a fact of life, nor is it a sign of admirable "rigor" on the teacher's part. Rather, "it is a symbol of failure — failure to teach well, failure to test well, and failure to have any influence at all on the intellectual lives of students" [Milton, Pollio, & Eison, 1986].) This surely represents an improvement over a system in which the number of top marks is made artificially scarce and students are set against one another. But here we've peeled back the outer skin of the onion (competition) only to reveal more noxious layers beneath: extrinsic motivation, numerical ratings, the tendency to promote achievement at the expense of learning.

If we begin with a desire to assess more often, or to produce more data, or to improve the consistency of our grading, then certain prescriptions will follow. If, however, our point of departure isn't mostly about the grading, but about our desire for students to understand ideas from the inside out, or to get a kick out of playing with words and numbers, or to be in charge of their own learning, then we will likely end up elsewhere. We may come to see grading as a huge, noisy, fuel-guzzling, smoke-belching machine that constantly requires repairs and new parts, when what we should be doing is pulling the plug.

### **Deleting — or at Least Diluting — Grades**

"Like it or not, grading is here to stay" is a statement no responsible educator would ever offer as an excuse for inaction. What matters is whether a given practice is in the best interest of students. If it isn't, then our obligation is to work for its elimination and, in the meantime, do what we can to minimize its impact.

Replacing letter and number grades with narrative assessments or conferences --qualitative summaries of student progress offered in writing or as part of a conversation — is not a Utopian fantasy. It has already been done successfully in many elementary and middle schools and even in some high schools, both public and private (Kohn, 1999c). It's important not only to realize that such schools exist but to investigate *why* they've eliminated grades, how they've managed to do so (hint: the process can be gradual), and what benefits they have realized.

Naturally objections will be raised to this — or any -- significant policy change, but once students and their parents have been shown the relevant research, reassured about their concerns, and invited to participate in constructing alternative forms of assessment, the abolition of grades proves to be not only realistic but an enormous improvement over the status quo. Sometimes it's only after grading has ended that we realize just how harmful it's been.

To address one common fear, the graduates of grade-free high schools are indeed accepted by selective private colleges and large public universities --on the basis of narrative reports and detailed descriptions of the curriculum (as well as recommendations, essays, and interviews), which collectively offer a fuller picture of the applicant than does a grade-point average. Moreover, these schools point out that their students are often more motivated and proficient learners, thus better prepared for college, than their counterparts at traditional schools who have been preoccupied with grades.

In any case, college admission is surely no bar to eliminating grades in elementary and middle schools because colleges are largely indifferent to what students have done before high school. That leaves proponents of grades for younger children to fall back on some version of an argument I call "BGUTI": Better Get Used To It (Kohn, 2005). The claim here is that we should do unpleasant and unnecessary things to children now in order to prepare them for the fact that just such things will be done to them later. This justification is exactly as absurd as it sounds, yet it continues to drive education policy.

Even when administrators aren't ready to abandon traditional report cards, individual teachers can help to rescue learning in their own classrooms with a two-pronged strategy to "neuter grades," as one teacher described it. First, they can stop putting letter or number grades on individual assignments and instead offer only qualitative feedback. Report cards are bad enough, but the destructive effects reported by researchers (on interest in learning, preference for challenge, and quality of thinking) are compounded when students are rated on what they do in school day after day. Teachers can mitigate considerable harm by replacing grades with authentic assessments; moreover, as we've seen, any feedback they may already offer becomes much more useful in the absence of letter or number ratings.

Second, although teachers may be required to submit a final grade, there's no requirement for them to decide unilaterally what that grade will be. Thus, students can be invited to participate in that process either as a negotiation (such that the teacher has the final say) or by simply permitting students to grade themselves. If people find that idea alarming, it's probably because they realize it creates a more democratic classroom, one in which teachers must create a pedagogy and a curriculum that will truly engage students rather than allow teachers to coerce them into doing whatever they're told. In fact, negative reactions to this proposal ("It's unrealistic!") point up how grades function as a mechanism for controlling students rather than as a necessary or constructive way to report information about their performance.

I spoke recently to several middle and high school teachers who have de-graded their classes. Jeff Robbins, who has taught eighth-grade science in New Jersey for 15 years, concedes that "life was easier with grades" because they take so much less time than meaningful assessment. That efficiency came at a huge cost, though, he noticed: Kids were stressed out and also preferred to avoid intellectual risks. "They'll take an easier assignment that will guarantee the A."

Initially Robbins announced that any project or test could be improved and resubmitted for a higher grade. Unfortunately, that failed to address the underlying problem, and he eventually realized he had to stop grading entirely. Now, he offers comments to all of his 125 students "about what they're doing and what they need to improve on" and makes abbreviated notes in his grade book. At the end of the term, over a period of about a week, he grabs each student for a conversation at some point -- "because the system isn't designed to allow kids this kind of feedback" -- asking "what did you learn, how did you learn it. Only at the very end of the conversation [do] I ask what grade will reflect it... and we'll collectively arrive

at something." Like many other teachers I've spoken to over the years, Robbins says he almost always accepts students' suggestions because they typically pick the same grade that he would have.

Jim Drier, an English teacher at Mundelein High School in Illinois who has about 90 students ranging "from at-risk to A.P.," was relieved to find that it "really doesn't take that long" to write at least a brief note on students' assignments — "a reaction to what they did and some advice on how they might improve." But he never gives them "a number or grade on anything they do. The things that grades make kids do are heartbreaking for an educator": arguing with teachers, fighting with parents, cheating, memorizing facts just for a test and then forgetting them. "This is not why I became a teacher."

Without grades, "I think my relationships with students are better," Drier says. "Their writing improves more quickly and the things they learn stay with them longer. I've had lots of kids tell me it's changed their attitude about coming to school." He expected resistance from parents but says that in three years only one parent has objected, and it may help that he sends a letter home to explain exactly what he's doing and why. Now two of his colleagues are joining him in eliminating grades.

Drier's final grades are based on students' written self-assessments, which, in turn, are based on their review of items in their portfolios. He meets with about three-quarters of them twice a term, in most cases briefly, to assess their performance and, if necessary (although it rarely happens) to discuss a concern about the grade they've suggested. Asked how he manages without a grade book full of letters or numbers, Drier replies, "If I spend 18 weeks with them, I have a pretty good idea what their writing and reasoning ability is."

A key element of authentic assessment for these and other teachers is the opportunity for students to help design the assessment and reflect on its purposes — individually and as a class. Notice how different this is from the more common variant of self-assessment in which students merely monitor their progress toward the teacher's (or legislature's) goals and in which they must reduce their learning to numerical ratings with grade-like rubrics.

Points of overlap as well as divergence emerge from the testimonies of such teachers, some of which have been collected by Joe Bower (n.d.), an educator in Red Deer, Alberta. Some teachers, for example, *evaluate* their students' performance (in qualitative terms, of course), but others believe it's more constructive to offer only *feedback* — which is to say, information. On the latter view, "the alternative to grades is description" and "the starting point for description is a plain sheet of paper, not a form which leads and homogenizes description" (Marshall, 1968, pp. 131, 143).

Teachers also report a variety of reactions to de-grading not only from colleagues and administrators but also from the students themselves. John Spencer (2010), an Arizona middle school teacher, concedes that "many of the 'high performing' students were angry at first. They saw it as unfair. They viewed school as work and their peers as competitors.... Yet, over time they switch and they calm down. They end up learning more once they aren't feeling the pressure" from grades.

Indeed, research suggests that the common tendency of students to focus on grades doesn't reflect an innate predilection or a "learning style" to be accommodated; rather, it's due to having been led for years to work for grades. In one study (Butler, 1992), some students were encouraged to think about how well they performed at a creative task while others were just invited to be imaginative. Each student was then taken to a room that contained a pile of pictures that other people had drawn in response to the same

instructions. It also contained some information that told them how to figure out their "creativity score." Sure enough, the children who were told to think about their performance now wanted to know how they had done relative to their peers; those who had been allowed to become immersed in the task were more interested in seeing *what* their peers had done.

Grades don't prepare children for the "real world" — unless one has in mind a world where interest in learning and quality of thinking are unimportant. Nor are grades a necessary part of schooling, any more than paddling or taking extended dictation could be described that way. Still, it takes courage to do right by kids in an era when the quantitative matters more than the qualitative, when meeting (someone else's) standards counts for more than exploring ideas, and when anything "rigorous" is automatically assumed to be valuable. We have to be willing to challenge the conventional wisdom, which in this case means asking not how to improve grades but how to jettison them once and for all.

## References

- Anderman, E.M., & Murdock, T.B., eds. (2007). *Psychology of academic cheating*. Burlington, MA: Elsevier Academic Press.
- Bedell, J. (2010, July). Blogpost M.
- Bower, J. (2010, March 28). Blogpost I<sup>2</sup>].
- Bower, J. (n.d.). Blog post t<sup>3</sup>]. [Grading moratorium list]
- Butler, R. (1988). Enhancing and undermining intrinsic motivation: The effects of task-involving and ego-involving evaluation on interest and performance. *British Journal of Educational Psychology*, 58, 1-14.
- Crooks, A.D. (1933). Marks and marking systems: A digest. *Journal of Educational Research*, 27(4), 259-72.
- De Zouche, D. (1945). "The wound *is* mortal": Marks, honors, unsound activities. *The Clearing House*, 19(6), 339-44.
- Eisner, E.W. (2001, Jan.). What does it mean to say a school is doing well? *Phi Delta Kappan*, pp. 367-72.
- Gardner, H. (1991). *The unschooled mind: How children think and how schools should teach*. New York: Basic Books.
- Grolnick, W.S., & Ryan, R.M. (1987). Autonomy in children's learning: An experimental and individual difference investigation. *Journal of Personality and Social Psychology*, 52, 890-98.
- Kirschenbaum, H., Simon, S.B., & Napier, R.W. (1971). *Wad-ja-get?: The grading game in American education*. New York: Hart.
- Kohn, A. (1999a). *Punished by rewards: The trouble with gold stars, incentive plans, A's, praise, and other bribes*. Rev. ed. Boston: Houghton Mifflin.
- Kohn, A. (1999b). *The schools our children deserve: Moving beyond traditional classrooms and "tougher standards."* Boston: Houghton Mifflin.
- Kohn, A. (1999c, March). From degrading to de-grading M. *High School Magazine*, pp. 38-43.
- Kohn, A. (2001, Sept. 26). Beware of the standards, not just the tests I<sup>5</sup>]. *Education Week*, pp. 52, 38.
- Kohn, A. (2005, Sept. 7). Getting hit on the head lessons t<sup>6</sup>J. *Education Week*, pp. 52, 46-47.
- Kohn, A. (2006, March). The trouble with rubrics t<sup>7</sup>]. *Language Arts*, pp. 12-15.
- Linder, I.H. (1940, July). Is there a substitute for teachers' grades? *School Board Journal*, pp. 25, 26, 79.
- Maehr, M.L., & Midgley, C. (1996). *Transforming school cultures*. Boulder, CO: Westview.
- Marshall, M.S. (1968). *Teaching without grades*. Corvallis, OR: Oregon State University Press.
- Matthews, J. (2006, Nov. 14). Just whose idea was all this testing? *Washington Post*.
- McNeil, L. M. (1986). *Contradictions of control: School structure and school knowledge*. New York: Routledge & Kegan Paul.



- Milton, O., Pollio, H. R., & Eison, J. A. (1986). *Making sense of college grades*. San Francisco: Jossey-Bass.
- Neill, M., Bursh, P., Schaeffer, B., Thall, C., Yohe, M., & Zappardino, P. (1995). *Implementing performance assessments: A guide to classroom, school, and system reform*. Cambridge, MA: FairTest.
- Nicholls, J. G., & Hazzard, S. P. (1993). *Education as adventure: Lessons from the second grade*. New York: Teachers College Press.
- Olson, K. (2006, Nov. 8). The wounds of schooling. *Education Week*, pp. 28-29.
- Pulfrey, C., Buch, C., & Butera, F. (2011). Why grades engender performance-avoidance goals: The mediating role of autonomous motivation. *Journal of Educational Psychology*, 103, 683-700.
- Spencer, J. (2010, July). Blogpost<sup>8</sup>].
- White, C.B., & Fantone, J.C. (2010). Pass-fail grading: Laying the foundation for self-regulated learning. *Advances in Health Science Education*, 15, 469-11.
- Wilson, M. (2006). *Rethinking rubrics in writing assessment*. Portsmouth, NH: Heinemann.
- Wilson, M. (2009, Nov). Responsive writing assessment. *Educational Leadership*, pp.58-62.