

It's Time to Stop Averaging Grades

By: Rick Wormeli



You've always averaged grades. Your teachers averaged grades when you were in school and it worked fine. It works fine for your students.

Does it? Just as we teach our students, we don't want to fall for *Argumentum ad populum*: something is true or good just because a lot of people think it's true or good. Let's take a look at the case against averaging grades.

Hiding Behind the Math

Just because something is mathematically easy to calculate doesn't mean it's pedagogically sound. The 100-point scale makes averaging attractive to teachers, and averaging implies credible, mathematical objectivity. However, statistics can be manipulated and manipulative in a variety of ways.

One percentage point is the arbitrary cut-off between getting into or being denied entrance into graduate school. One student gets a 90% and another gets an 89%: the first is an A and the second is a B, yet we can't discern mastery of content to this level of specificity. These students are even in mastery of content, but we declare a difference based only on the single percentage point. The student with 90% gets scholarships and advanced class placements and the student with 89% is left to a lesser path. Something's wrong with this picture.

Early in my career, one of my students had a 93.4% in my class. Ninety-four to 100 was the A range set for that school, so he was 0.6% from achieving an A. The student asked if I would be willing to round the score up to the 94% so he could have straight As in all his classes. I reminded him that it was 93.4, not 93.5, so if I rounded anything, I would round down, not up. I told him that if it was 93.5, I could justify rounding up, but not with a 93.4.

I was hiding behind one-tenth of a percentage point. I should have interviewed the student intensely about what he had learned that grading period and made an executive decision about his grade based on the evidence of learning he presented in that moment. The math felt so safe, however, and I was weak. It wasn't one of my prouder moments.

We can't resort to averaging just because it feels credible by virtue of its mathematics. There's too much at stake.

Falsifying Grade Reports

Consider the teacher who gives Martin two chances to do well on the final exam, then averages the two grades. The first attempt results in an F grade, but after re-learning and a lot of hard work, the second attempt results in an A. We trust the exam to be a highly valid indicator of

student proficiency in the subject, and Martin has clearly demonstrated excellent mastery in the subject. When the two grades are averaged, however, the teacher records a C in the grade book—falsely reporting his performance against the standards.

This is strikingly inaccurate when using grading scale endpoints such as A and F, and it creates just as inaccurate “blow-to-grade-integrity” reporting as when we average grades closer to one another on the scale: B with D, B with F, A with C, etc.

Consider a sample with more data: Cheryl gets a 97, 94, 26, 35, and 83 on her tests, which correspond to an A, A, F, F, and a B on the school grading scale. When the numbers are averaged, however, everything is given equal weight, and the score is 67, which is a D. This is an incorrect report of her performance against individual standards.

Thankfully, many schools are moving toward disaggregation in which students receive separate grades for individual standards. This will cut down dramatically on the distortions caused by aggregate grades that combine everything into one small symbol and will help eliminate teacher concerns about students who “game” the system when their teachers re-declare zeroes as 50s on the 100-point scale. These students try to do just enough— skipping some assessments, scoring well on others—to pass mathematically. In classrooms where teachers do not average grades, students can't do this.

No more mind games; students have to learn the material.

Countering the Charge

“Average,” “above average,” and “below average” are norm references, but in today's successful classrooms, we claim to be standards- (outcomes-) based. This means that assessments and grading are evidentiary, criterion-referenced. A teacher declares Toby is above average, but we're not interested in that because it provides testimony of Toby's proficiencies only in relation to others' performance, which may be high or low, depending on the group. Instead, we want to know if Toby can write an expository essay, stretch correctly before running a long distance, classify cephalopods, and interpret graphs accurately. We don't need to know how well he's doing in relation to classmates nearly so much as how he's doing in relation to his own progress and to societal standards declared for this grade level and subject.

We can't make specific instructional decisions, provide descriptive feedback, or document progress without being criterion-referenced. Declarations of average-ness muddle our thinking and create a false sense of reporting against standards. We need grade reports to be accurate.

Distorting Averaging's Intention

One of the reasons we developed averaging in statistics was to limit the influence of any one sample error on experimental design. Let's see how that works in the classroom.

Consider a student taking a test on a particular topic and in a particular format. The student ate breakfast, or he did not. He slept well, or he did not. His parents are divorcing, or they are not. He has a girlfriend, or he does not. He studied for this test, or he did not. He is competing in a high-stakes drama/music/sports competition later this afternoon, or he is not. Whatever the combination, all these factors conspire to create this student's specific performance on this test on this day at this time of day.

Three weeks later, we give students another test about new material in our unit. Have students changed during three week? Yes, hormonally, if nothing else. Add that the second test is on a different topic and perhaps in a different format. On the first test date, the student ate well, but didn't study. He slept well, but his parents are arguing each night. The drama/music/sports performance came and went and he did well in it. He didn't have a girlfriend. For the second test, however, he has a girlfriend, and he studied. He didn't sleep well, however, nor did he eat breakfast, and his parents have stopped arguing which has calmed things down at home.

The second test situation is dramatically altered. The integrity of maintaining consistent experimental design is violated. We can no longer justify averaging the score of the first test with the score of the second test just to limit the influence of any one sample error.

The Electronic Gradebook

The only reason our electronic gradebooks average grades is because someone declared it a policy—not because it was the educationally wise thing to do—so the district uses the technology that supports that decision. Why don't we choose our grading philosophy first, then find the technology to support it rather than sacrificing good grading practices because we can't figure out a way to make the technology work?

How do we do what's right when we are asked by administrators or a school board to do something that we know is educationally wrong? This is a tough situation, but I suggest we do the ethical thing in the microcosm of our own classrooms, then translate that into the language of the school or district so we can keep our jobs.

We can experiment in our own classes by reporting a subset of students' grades with and without averaging them just to see how they align with standardized testing. Sometimes running the numbers/grades ourselves helps us see with greater clarity than just hearing about ideas second-hand.

We can read articles on grading and averaging, participate in online conversations on the topic, and start conversations with faculty members. We can also volunteer to be on the committee to revise the gradebook format.

We're working with real individuals, not statistics. Our students have deeply felt hopes and worries and wonderfully bright futures. They deserve thoughtful teachers who transcend conventional practices and recognize the ethical breach in knowingly falsifying grades. Let's live up to that charge and liberate the next generation from the oppression of averaging.

Previously published in Middle Ground magazine, October 2012

More on these topics

[Assessment](#) [Teaching](#)

Article tags

[Grades](#) [Rick Wormeli](#)



5 Comments

// This is an important article with a powerful message. In addition to Rick's outstanding work - I encourage readers to review the work of O'Connor. Thank you Rick. Keep the messages coming!

—Shelley Joan
10/4/2013 5:52 PM

// I'm interested in seeing what others think. Please add your comments!

—Shelley Joan
10/4/2013 5:52 PM

// This article has some interesting points, but to say average only applies to a group is short-sided. Algebra 1 content may slightly ebb and flow, but a "C", or average student today can help a teacher compare that student to the last 10 years of students. I sure hope teachers aren't basing their grades on a specific cohort. Some years I have 10 average students out of 25, while other years it may be 5.

I only have 6 years experience, but I've never seen a student with two 90s and a 20. That's just silliness to prove a point.

I am intrigued by standards based grading, but who is going to design and grade each assessment to identify the standards addressed by each question. It's not feasible to do by hand and have a family. Computers can do it, but then teachers lose class customization. What about single questions which address multiple standards, such as any system of equations word problem?

Isn't standards based grading, what state testing and MAP testing has been doing for years?

—Bryan
3/17/2014 10:18 PM

Advertisement



“ In this article, you take the stance that averaging grades is detrimental to the educational system. Instead, you advocate for standards-based grading. I agree with some of the points you make. I completely agree it's unfair to the student when he/she gets an A on a test the second time around and still gets a C in the grade book. It's also extremely wrong that one percentage point can determine the future of a child. Average-based grades have become so essential to future success that scant else matters to some students. Standards-based testing may be a feasible solution.

—Jennifer

10/7/2014 10:03 PM

“ What I am hearing you say is that averaging grades doesn't accurately reflect a student's comprehension and mastery of their assignments. One bad grade can ruin a student's average, making them appear to be doing poorly when they may have just been hungry, tired, or upset the day they received the grade. I agree that we should be evaluating what students know, not how they compare with others. I am in favor of rubric-based grading because students know exactly what to expect and can pre-determine their scores based on the effort they put forth.

—Brooke

10/13/2014 12:34 AM

<https://www.amle.org/BrowsebyTopic/WhatsNew/WNDet/TabId/270/ArtMID/888/ArticleID/21/Its-Time-to-Stop-Averaging-Grades.aspx>