Response to “The Case Against Grades”

David M. Harrison
Senior Lecturer (Emeritus), Dept. of Physics, Univ. of Toronto
david.harrison@utoronto.ca

August, 2022

Recently Alfie Kohn’s paper “The Case Against Grades” has been uploaded to this group by Eugenia. Here I discuss this document, ask some questions, and offer some opinions. I hope it stimulates a respectful and useful discussion.

What’s Politics Got To Do With It?

In 2012 Bar & Zussman published the results of a U.S. study of the grades given to students by professors as a function of the political persuasion of the professors (Democrat-liberal or Republican-conservative). The professors were roughly almost equally distributed between the humanities, social sciences, and physical and mathematical sciences. The school was an "elite" U.S. university (probably Cornell).

The range of grades, i.e. the standard deviation σ, was less for the Democrats than for the Republicans. It seems that the liberal profs were more inclined to grade all students as the same, while the conservative ones were more inclined to reward the top performers with high grades and punish the bottom of the distribution with low grades.

These results are consistent with the liberal/conservative difference in values explored by, for example, Haidt in *The Righteous Mind.*

Eliminating grades entirely as Kohn advocates essentially is reducing the standard deviation to zero. Does this mean that I can make a good prediction of who Kohn voted in the last election based solely on his advocacy for eliminating grades? I find this idea disturbing.

Below I will have a few comments about testing theory, but here will just point out that a professionally designed test strives for a large standard deviation in the scores. So at least on this basis the more conservative professors have it right.

What's Physics Got To Do With It?

People in general are lazy, and high school and first year university students are particularly lazy. Further, being young, inexperienced, and with not fully developed pre-frontal cortexes, they often make poor choices. One of our jobs as educators is to try to guide our students to make better choices and hopefully help them overcome their laziness. I will have a few more remarks about this below, but for now will point out that physics teachers have a particular challenge: physics is widely considered by our students to be not only irrelevant but also very difficult.

For example, in a 1000-student first year university physics course for life science majors at the Univ. of Toronto, only 16% of the class are taking the course for their own interest.

Part of the issue is the large number of people who are math-phobic. Examples include a 19-year old Charles Darwin, and psychologist C.G. Jung for his entire life. Some math-phobic students end up in our physics class because it is required for some purpose, but they do not want to be there.

For us as teachers to think we have such great talent and charisma that we can get all of these students to completely change their attitudes and embrace the joys of physics strikes me as nearly delusional.

What's Equality Got To Do With It?

The U.S. Declaration of Independence states that all people "are created equal." But this is clearly not true. A combination of genetics, family socio-economic status, and perhaps some random factors mean that some people can, for example, run 100 meters much much faster than I can. Is the ability to do physics similarly influenced by such factors? It would be pretty surprising if it were not. What data are available also indicate that academic ability correlates with genetic and socio-economic factors.

So there is a big difference between equality of opportunities vs. equality of outcomes. In an ideal world, all people get to run a 100 m dash as fast as they can, and also get to learn as much as possible in a physics class. But some people will run faster than others, and some will learn more physics than others.

Some seem to think that this difference in outcomes in, for example, physics means that the system is not equitable. Nonetheless, when putting together an Olympic sprint team we choose the fastest runners. When putting together a team of physicists to, for example, build a quantum computer should we not similarly choose the people who have demonstrated that they know the most physics? How to define and determine who knows the most physics will be discussed at some length below.

Of course, choosing the “best” physicists for the quantum computer team is a form of meritocracy, which is under some attack by those who insist on equality of outcomes. I find Adrian Wooldridge’s arguments in favour of meritocracies in *The Aristocracy of Talent* to be compelling. This book was the *Times* (UK) book of the year for 2021,

Personally, if I am about to drive over a bridge, I want the designer to have been the best engineer on the planet.

What Do Tests Have To Do With It?

Priscilla Laws, whose work is one of the foundations of Physics Education Research (PER), says repeatedly, “If you don’t test for it, you don’t get it.” Some critics seem to believe that this fact shows a flaw in grading in general. Many of these same people, including Kohn, bristle when a student asks “Will this be on the test?” For me, however, what the student is asking is whether I, the teacher, thinks this particular thing is important. A fair question.

Certainly conventional tests and exams are far from perfect tools for assessing student knowledge. However, a large body of research has given us a number of tools for estimating just how effective a test is at the assessment. These include statistics for the reliability of the test, such as Cronbach’s α or the Kuder-Richardson Formula 20. There are also measures of the consistency of individual questions compared to the overall test, such as the Pearson correlation coefficient.

I have used these tools to investigate the statistical uncertainty in the grades generated by my tests. I like to believe that the tools have helped me to be more effective in designing reliable tests. Typically, however, the uncertainty in the grade on my tests, i.e. σ, is still a little more than one full letter grade (± 11%).

However, as physicists we are all used to always having experimental uncertainties, and have learned methods of dealing with them. Substituting a completely subjective narrative assessment, as Kohn and others advocate, seems to me to be in considerable conflict with our attempts to apply the scientific method to our pedagogy.

Finally, one of the most important tools in leading the reformation in teaching fostered by PER is the Force Concept Inventory (FCI) test. Although this instrument is not typically used to generate a grade for the student, the results and in particular the normalized gain in the FCI scores from the beginning to the end of the class is very much an assessment or grade of the teacher’s effectiveness. So I use conventional tests to grade my students, and the FCI to grade myself.

References

Talia Bar and Asaf Zussman, “Partisan Grading”, American Economic Journal, Applied Economics **4**(1),30 (January 2012,).
https://www.aeaweb.org/articles?id=10.1257/app.4.1.30

Jonathan Haidt, *The Righteous Mind: Why Good People Are Divided by Politics and Religion* (Pantheon, 2012).

David M. Harrison, “The Uncertainty of Grades in Physics Courses is Surprisingly Large”, Physics in Canada **70**(2), 75 (2014).
http://www.upscale.utoronto.ca/PVB/Harrison/UncertaintyOfGrades/UncertaintyOfGrades.pdf

Adrian Wooldridge, *The Aristocracy of Talent: How Meritocracy Made the Modern World* (Simon and Schuster, 2021).