



Grade Inflation for Education Majors and Low Standards for Teachers: When Everyone Makes the Grade

By Cory Koedel

Students who take education classes at universities receive significantly higher grades than students who take classes in every other academic discipline. The higher grades cannot be explained by observable differences in student quality between education majors and other students, nor can they be explained by the fact that education classes are typically smaller than classes in other academic departments. The remaining reasonable explanation is that the higher grades in education classes are the result of low grading standards. These low grading standards likely will negatively affect the accumulation of skills for prospective teachers during university training. More generally, they contribute to a larger culture of low standards for educators.

A 2009 report from the New Teacher Project shows that teachers in K–12 schools receive overwhelmingly positive performance evaluations.¹ The report has brought much-needed attention to the low evaluation standards for K–12 teachers. This *Outlook* examines the standards by which prospective teachers are evaluated during university training. Grading standards in education departments at universities, where much of the teaching workforce is trained, are also strikingly low. In addition to documenting the low grading standards in education departments, I consider some of the likely consequences and discuss possible solutions.

The Grading Standards Discrepancy

A 1960 article by Robert M. Weiss and Glen R. Rasmussen provides early evidence of low grading standards in university education departments. Over fifty years ago, these authors showed that

undergraduate students taking education classes were twice as likely to receive an A compared with students taking classes in business or liberal arts departments. Low grading standards in education classes are still prevalent today.²

Figures 1 and 2 show distributions of classroom-level grades, on a four-point scale, at two state flagship universities during the 2007–2008 school year. The distributions are for all nonfreshman undergraduate classes in twelve fields, plus

Key points in this *Outlook*:

- Grades awarded in university education departments are consistently higher than grades in other disciplines.
- Similarly, teachers in K–12 schools receive overwhelmingly positive evaluations.
- Grade inflation in education departments should be addressed through administrative directives or external accountability in K–12 schools.

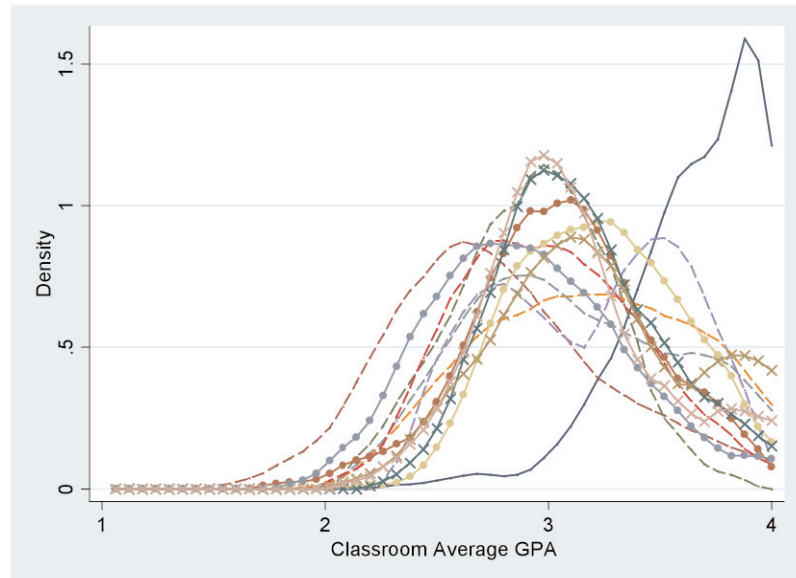
Cory Koedel (koedelc@missouri.edu) is an assistant professor of economics at the University of Missouri.

education.³ The outlying grade distribution in each figure belongs to the education department. The other distributions are cluttered, but this is largely the point: while all other university departments work in one space, education departments work in another. In figures 1 and 2, the average classroom-level grade point averages (GPAs) in the education departments are 3.66 and 3.80, respectively (the corresponding averages weighted by class size are 3.65 and 3.73). At the University of Missouri–Columbia, depicted in figure 2, every single student received an A (that is, 4.0) in one out of every five undergraduate education classes.⁴

In a recent article, I used additional data from other universities, provided by Myedu.com, to show that the grade distributions depicted in figures 1 and 2 are not unique.⁵ To the contrary, they appear to be quite typical. The data consistently show that education departments award exceptionally favorable grades to virtually all their students in all their classes.

The favorable grades awarded in education classes cannot be attributed to student quality or structural factors like smaller classes. With regard to the former, education majors score considerably lower than students in other academic departments on college entrance exams.⁶ And although education classes are typically smaller than other classes, this structural difference does not explain the

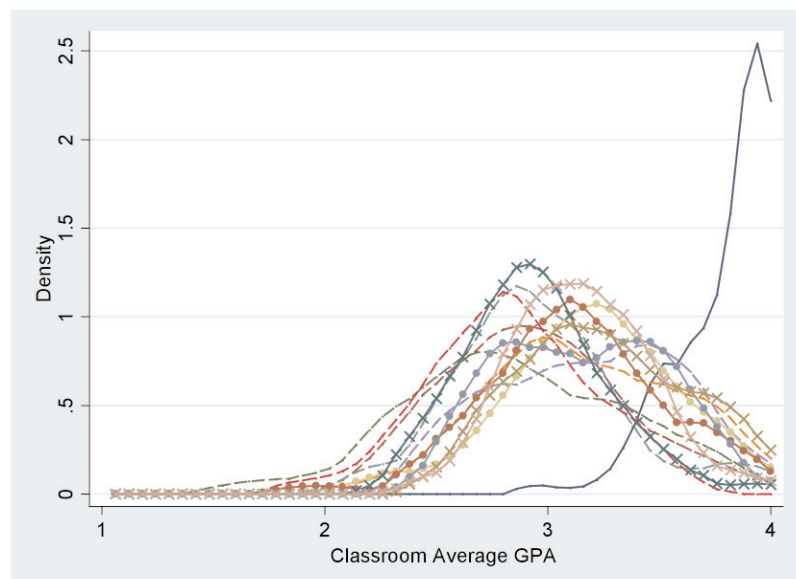
FIGURE 1
DISTRIBUTIONS OF CLASSROOM GRADES FOR TWELVE MAJOR ACADEMIC DEPARTMENTS, PLUS THE EDUCATION DEPARTMENT, 2007–2008 SCHOOL YEAR, NONFRESHMAN UNDERGRADUATE CLASSES AT INDIANA UNIVERSITY–BLOOMINGTON



SOURCE: Cory Koedel, “Grading Standards in Education Departments at Universities,” *Education Policy Analysis Archives* 19 (2011).

NOTE: The outlying distribution is for the education department. Other departments include math and science (dashed lines), social science (x’s), and humanities (circles).

FIGURE 2
DISTRIBUTIONS OF CLASSROOM GRADES FOR TWELVE MAJOR ACADEMIC DEPARTMENTS, PLUS THE EDUCATION DEPARTMENT, 2007–2008 SCHOOL YEAR, NONFRESHMAN UNDERGRADUATE CLASSES AT THE UNIVERSITY OF MISSOURI–COLUMBIA



SOURCE: Cory Koedel, “Grading Standards in Education Departments at Universities,” *Education Policy Analysis Archives* 19 (2011).

NOTE: The outlying distribution is for the education department. Other departments include math and science (dashed lines), social science (x’s), and humanities (circles).

grading discrepancy. To show this, I used data from multiple academic departments at three large universities to adjust classroom-level grades for differences in class size, and then I reestimated the GPA gaps between education and noneducation departments. Even after this adjustment, the large gap between the grades awarded in education and noneducation classes remains.⁷

Consequence 1: We Are Training Teachers Who Know Less. Grade inflation is associated with reduced student effort in college—put simply, students in classes where it is easier to get an A do not work as hard. This is not surprising, and a recent study by Philip Babcock quantifies the effect.⁸ He shows that in classes where the expected grade rises by one point, students respond by reducing effort, as measured by study time, by at least 20 percent.⁹ It is straightforward to apply Babcock's result to the data from the two schools depicted in figures 1 and 2. If the grading standards in each education department were moved to align with the average grading standards at their respective universities, student effort would rise by at least 11–14 percent.

Would it be valuable for prospective teachers to apply more effort in college? The answer is a qualified yes. For the increased effort to be beneficial, it must be the case that either the content of classes taught in education departments adds direct value in terms of teaching quality, or teachers gain other skills indirectly as a result of a more demanding college experience (for example, skills in time management or improved work ethics). Unfortunately, I am not aware of any rigorous evidence that explicitly links higher grading standards in education departments to improved teaching performance in K–12 schools. But this does not mean a link does not exist; instead, the lack of research appears to be largely due to a lack of data. Specifically, to evaluate a “treatment”—in this case, higher grading standards in education departments—researchers need to observe variability in the treatment. There seems to be little variability in grading standards in education departments, either across universities or within universities over time.

Of course, grades also play an important filtering role in most academic departments, deterring students with limited skills or with skills poorly matched to the discipline. Grades do not appear to play such a role in education departments.

Consequence 2: Education Departments Are Contributing to the Culture of Low Standards for Educators.

A superintendent asked a school principal to tell him how many of her teachers were performing well. The principal replied that they were all performing well. Puzzled, the superintendent reminded her that the vast majority of the children at the school were not reading even within a year of grade level, and he asked the question again. The principal's response was unchanged. He then asked the principal which of the teachers at her school would be suitable to teach her own granddaughter. She replied, “Well, if that's the bar, then none of them.”¹⁰

The overwhelmingly positive evaluations that teachers receive in K–12 schools look very similar to the favorable grades they receive during university training. In their 2008 study, Brian Jacob and Lars Lefgren asked principals from a Western school district to evaluate the teachers in their schools on a ten-point scale, with a one indicating “inadequate” and a ten indicating “exceptional.” Figure 3 plots the distribution of the principals' ratings.¹¹

Note that most teachers received an eight or better (in fact, even the thirtieth percentile teacher received an eight), and just like the grades awarded by education departments at universities, the ratings are highly compressed. Similarly favorable ratings were also observed in a Florida school district by Douglas N. Harris and Tim R. Sass, who surveyed school principals about teaching performance in a 2010 study.¹² These ratings are in stark contrast to one of the most consistent findings in the research literature on teacher quality—teacher effectiveness differs dramatically as measured by student outcomes.¹³

Undergraduate education majors become teachers, teachers become principals, and principals become district-level administrators. Ultimately, a sizable fraction of the workforce in the education sector is trained in education departments where evaluation standards are astonishingly low. Should we be surprised that low standards persist in K–12 schools?

The Underlying Problem

It seems difficult to argue with the notion that low grading standards in education departments at universities are bad for students in K–12 schools. But Weiss and Rasmussen documented these low standards over fifty years ago, so this has been an ongoing cultural norm for some time. What is causing the problem, and what can be done to fix it? Why have the outlying grading standards in education departments persisted for so long? Most university professors would probably prefer to give favorable grades to students because it is unpleasant to do otherwise, but what

keeps professors in other departments in line (relatively speaking, at least)?

One notable difference between education departments and other major departments at universities is that virtually all graduates from education departments move into a single sector of the labor market—education. If the education sector is less effective at identifying low-quality graduates than are other sectors of the labor market, this would help explain why professors in education departments are able to consistently award As to most students.

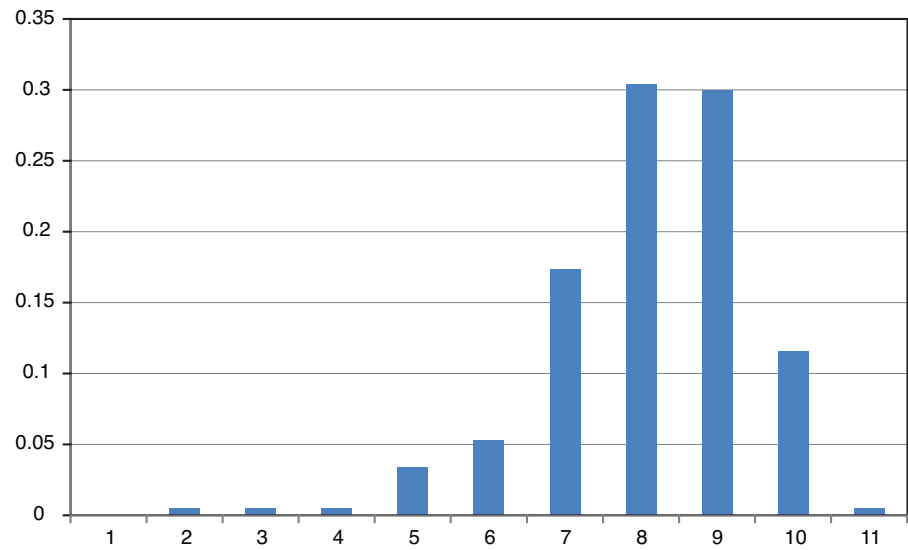
To illustrate, consider an example from outside education: suppose that engineering professors at University X greatly lowered their grading standards and began producing low-quality engineers. Engineering firms that hire the fresh engineering graduates are forced by the competitive nature of the market to pay workers commensurate with their skills. These firms would observe the decline in the quality of graduates from University X and respond by hiring fewer engineering students from University X or offering them lower wages. In turn, this would lower student demand for the program, which would put pressure on the professors to improve quality. The professors would respond by retightening standards.

The education sector is notoriously ineffective at identifying high- and low-quality workers, making it difficult for the labor market to penalize students from education departments that produce low-quality teachers. The culture of low standards in education is partly to blame, as those within the education establishment have shown little interest in distinguishing good teachers from mediocre teachers, but why can K–12 schools get away with not distinguishing high-quality workers from low-quality workers? What makes education different?

The answer is that there is not a competitive market forcing schools and districts to be efficient. For example, if a school hires mediocre teachers and produces mediocre outputs year after year, there is no mechanism to meaningfully penalize the school or its workers.¹⁴ In competitive

FIGURE 3

DISTRIBUTION OF PRINCIPALS' RATINGS OF TEACHERS ON A TEN-POINT SCALE



SOURCE: Cory Koedel, "Grading Standards in Education Departments at Universities," *Education Policy Analysis Archives* 19 (2011).

NOTE: The data for this figure were provided by Brian Jacob and Lars Lefgren.

industries, mediocre firms succumb to market pressure; they face consequences like lower stock prices, reduced market shares, or even going out of business altogether. Returning to the superintendent's conversation with the school principal, if the principal had to answer to shareholders about the students at her school who were not reading at grade level, and her job were on the line, it seems unlikely that she would be so quick to conclude that all her teachers were performing well.

The data consistently show that education departments award exceptionally favorable grades to virtually all their students in all their classes.

Solutions

The fundamental problem is simple: there is no pressure from competitive markets in education. The solution, as with any market failure, is external intervention. Two external forces with the potential to meaningfully intervene are university administrators and external accountability measures in K–12 schools.

University Administrators. University administrators are in a position to intervene directly by imposing more stringent grading standards on education departments. For example, administrators could externally dictate the grade distributions in education classes to bring them in line with those in other departments. Although this may seem like an extreme response, figures 1 and 2 show that what is occurring in education departments is *already extreme*. Also, such a move by administrators would not be unprecedented. Princeton University has been a stalwart in combating grade inflation, and it does so by issuing university-wide grading guidelines at the administrative level.

The overwhelmingly positive evaluations that teachers receive in K–12 schools look very similar to the favorable grades they receive during university training.

External Accountability in K–12 Schools. Several states have produced measures of teacher effectiveness linked to teacher-training institutions, and others are in the process of doing so. This work is partly motivated by increased demand from schools and districts, driven by increased external accountability. The results from these studies will allow schools and districts to identify teacher-training programs that produce high- and low-quality teachers. This information will facilitate private-market-like incentives for education departments similar to the incentives faced by other academic departments at universities, so long as K–12 schools are held accountable for performance.

Conclusion

Low grading standards in university education departments are part of a larger culture of low standards for educators, and they precede the low evaluation standards by which teachers are judged in K–12 schools. The culture of low standards for educators is problematic because it creates a disconnect between teachers' perceptions of acceptable performance and the perceptions of everyone else.

Society resists change, and resistance to change is particularly acute in education. But there is no rational

reason for the low grading standards in education departments. Rather than asking why these grading standards should be changed, perhaps the more reasonable question is *why they shouldn't be changed*. Put differently, if we were to start over with university education and could choose the grade distributions in each discipline, would we choose the currently observed discrepancy between education departments and all other academic departments?

University administrators are in the best position to intervene and modify the low grading standards in education departments. In the absence of administrative action, external accountability in K–12 schools will also likely lead to higher standards in education departments over time, although the pace of change will be slower.

Notes

1. Daniel Weisberg, Susan Sexton, Jennifer Mulhern, and David Keeling, *The Widget Effect: Our National Failure to Acknowledge and Act on Differences in Teacher Effectiveness* (New Teacher Project, Brooklyn, NY, 2009).

2. See Robert M. Weiss and Glen R. Rasmussen, "Grading Practices in Undergraduate Education Courses: Are the Standards Too Low?" *The Journal of Higher Education* 31, no. 3 (March 1960): 143–49.

3. The other departments in the figures are, in alphabetical order: biology, chemistry, computer science, economics, English, history, math, philosophy, physics, political science, psychology, and sociology. Tables that identify the departments in the figures can be found in Cory Koedel, "Grading Standards in Education Departments at Universities," *Education Policy Analysis Archives* 19, no. 23 (August 2011): 8, 18.

4. Across all academic departments, grading standards at universities appear to have declined over time. Education departments are at least keeping pace with the overall decline. See Philip Babcock and Mindy Mark, "Leisure College, USA: The Decline in Student Study Time," *AEI Education Outlook* (August 2010), www.aei.org/outlook/100980.

5. See Cory Koedel, "Grading Standards in Education Departments at Universities."

6. See, for example, Peter Arcidiacono, "Ability Sorting and the Returns to College Major," *Journal of Econometrics* 121, no. 1–2 (2004): 343–75; and College Board, *2010 College-Bound Seniors: Total Group Profile Report* (New York, NY, 2010). Arcidiacono reports average math and verbal SAT scores for science, business, social science/humanities, and education majors upon college entry. For math, the average scores

by major are 566, 498, 500, and 458, respectively; the average verbal scores are 499, 444, 481, and 431. The College Board provides an even more detailed comparison using more recent data. The SAT score gaps reported by the College Board are similar to those reported by Arcidiacono.

7. See Cory Koedel, "Grading Standards in Education Departments at Universities."

8. See Philip Babcock, "Real Costs of Nominal Grade Inflation? New Evidence from Student Course Evaluations," *Economic Inquiry* 48, no. 4 (October 2010).

9. Babcock takes a conservative estimation approach in his analysis. There are reasons to expect his estimate to understate the effect of grade inflation.

10. This anecdote is based on an interaction between a school principal and the superintendent in a large, urban school district, as conveyed to the author by the superintendent.

11. See Brian Jacob and Lars Lefgren, "Principals as Agents:

Subjective Performance Assessment in Education," *Journal of Labor Economics* 26, no. 1 (January 2008). In the figure, all ratings were rounded to the nearest integer (for approximately 7 percent of the ratings, a teacher was given a score that involved a fraction of a point—for example, 8.5). I thank Brian Jacob and Lars Lefgren for providing the complete ratings distribution from their study.

12. See Douglas N. Harris and Tim R. Sass, "What Makes for a Good Teacher and Who Can Tell?" (unpublished manuscript, Florida State University, Tallahassee, FL, 2010).

13. For a review of the recent literature, see Eric A. Hanushek and Steven G. Rivkin, "Generalizations about Using Value-Added Measures of Teacher Quality," *American Economic Review* 100, no. 2 (May 2010).

14. Federal No Child Left Behind legislation has attempted to introduce market-like incentives of this type into schools. Still, most schools are unaffected by these sanctions.