CrossMark

Grade Inflation in Higher Education: Is the End in Sight?

Michael J. Carter · Patricia Y. Lara

Over the past few decades, claims of grade inflation in American higher education have been ubiquitous, with ample evidence documenting its prevalence and severity.[1] Many have condemned the trend toward grade inflation, noting that students spend less time studying in courses that inflate grades,[2] and that students who receive inflated grades in introductory or preliminary courses often do poorly in advanced courses.[3] In this article we present summary findings of a study we conducted that examines grading trends across a recent five-year span in two of the largest higher education systems in the United States: the University of California (UC) and California State University (CSU) systems.[4] We show that changes in grade distributions in many campuses have begun to plateau, but note that it may be premature to claim that grade inflation is an issue of the past. We also cite and discuss a potential correlate of grade inflation others have ignored: the relationship between grade point averages and semantic definitions of grade categories.

[1] Roger A. Arnold, "Way That Grades Are Set Is a Mark against Professors," *Los Angeles Times*, April 22, 2004, http://articles.latimes.com/2004/apr/22/opinion/oe-arnold22.

[2] Philip Babcock, "Real Costs of Nominal Grade Inflation? New Evidence from Student Course Evaluations," *Economic Inquiry* 48, no. 4 (October 2010): 983–96.

[3] Valen E. Johnson, *Grade Inflation: A Crisis in College Education* (New York: Springer-Verlag, 2003).

[4] The complete study findings are available on the NAS website at http://www.nas.org/articles/grade_inflation_in_higher_education_is_the_end_in_sight.

Michael J. Carter is associate professor of sociology at California State University, Northridge, Northridge, CA 91330-8318; michael.carter@csun.edu. His main research interests are in social psychology, specifically the area of self and identity. His research has appeared in the *American Sociological Review* and *Social Psychology Quarterly*.

Patricia Y. Lara is an MA candidate in sociology at California State University, Northridge, Northridge, CA 91330-8318; patricia.lara.707@my.csun.edu. Her interests include college student development, qualitative research methodology, and criminology. Her research has appeared in the *International Journal of Research & Method in Education*.

Springer

Still an Issue?

There are two issues concerning the phenomenon of grade inflation in contemporary higher education. The first regards the fact that over time distributions for letter grades have gradually increased (become inflated), so that As and Bs are now more commonly assigned than Cs, Ds, or Fs. The second issue is more of a question: Are grade inflation trends that have been documented in higher education over the past decades *continuing* to occur? In other words, are the *inflated* grade distributions that now define higher education still *inflating*? We are interested in the latter question.

There is no consensus on the causes of grade inflation. Regardless of *why*, the fact that grades have increased begs a question: Considering that the common grading scale has lower and upper limits—i.e., the scale ranges from zero to 4.0—how far toward the high end of the scale will grades increase before the limit is reached? After all, inflation connotes *change*—by definition inflation represents a *sustained increase*. Considering the common grading scale's upper limit, grades cannot inflate forever; at some point inflation must—theoretically—stop. A look at recent data shows that such an end may be in sight.

Grade Inflation Trends: 2009–2013

We examined current grading trends in the UC and CSU systems to discover the prevalence of grade inflation.[5] Between 2009 and 2013 the average grade point average (GPA) across all UC campuses (nine total) was 3.03. Across CSU campuses (twenty-three total), the average GPA was 2.93. Among UC campuses, the highest GPAs were found at UC Berkeley (3.29), the lowest at UC Riverside (2.77). Among CSU campuses, the highest GPAs were found at San Francisco State and Sonoma State (both = 3.11); the lowest at CSU Bakersfield (2.74). One can debate whether these GPAs are too high (indeed, many would claim it so). A separate question is whether these averages have *changed* over the past half-decade. To answer this question we performed a series of statistical (regression) analyses on UC/CSU grade distribution data, the results of which allow us to determine if GPAs

---

[5]Grade data for the UC and CSU campuses was the most recent available at the time the study was conducted. Data for CSU Los Angeles represents GPA trends between 2008 and 2011.

have changed significantly in recent years. Figures 1 and 2 present results of our analysis for each UC and CSU campus, respectively.[6]

At first glance, figures 1 and 2 seem to reveal (visually) that grade distributions at many UC and CSU campuses showed an upward trend between 2009 and 2013. However, when analyzing the data for statistical significance, only half (five out of nine: Berkeley, Riverside, Santa Barbara, San Diego, and UCLA) of the UC campuses showed a significant increase in GPA over that time. Grade inflation trends were even less noticeable among CSU campuses; only one-third had significantly higher GPAs across the years observed (Dominguez Hills, East Bay, Fullerton, Northridge, Pomona, San Diego, San Jose, and San Luis Obispo). Across both university systems, *nineteen* campuses we examined had stable GPAs, and in the CSU, Humboldt State's GPA actually *decreased* significantly. So, the notion that grade inflation (i.e., an ongoing rise in grades) continues to plague all areas of higher education is not altogether true: thirteen of the thirty-two UC/CSU campuses (41 percent) inflated grades (meaning that they showed a steady increase in grade point average between 2009 and 2013).

These findings offer some optimism to those who fear that grade inflation will continue unabated, though the fact that two-fifths of the universities examined in this study still inflate grades may quickly quell that optimism (and it is possible that the universities that did not inflate grades could return to the practice in the future). Moreover, it should be reiterated that regardless of which universities in our sample inflated (i.e., steadily increased) grades between 2009 and 2013, *all* have average grade distributions that are high overall. Let us now discuss *why* grades might be higher at some university campuses than others. Here we cite a correlate of GPA that others have not generally considered.

Letter Grades as Semantic Categories

Research on grade inflation usually treats changes in GPA over time as the sole indicator of inflation. For example, if today a university's combined student GPA is 3.12 and it was 2.98 five years ago, most would cite this as indicating grade inflation. This is sensible, but often in these studies the semantic meaning of "grade point average," i.e., what GPA actually represents, is glossed over or ignored. After all, a student's performance in a class is not usually assessed by a

---

[6]Figures 1 and 2 present the study results as linear fitted values; regression coefficients for the analysis performed is presented in the complete report, available on the NAS website.

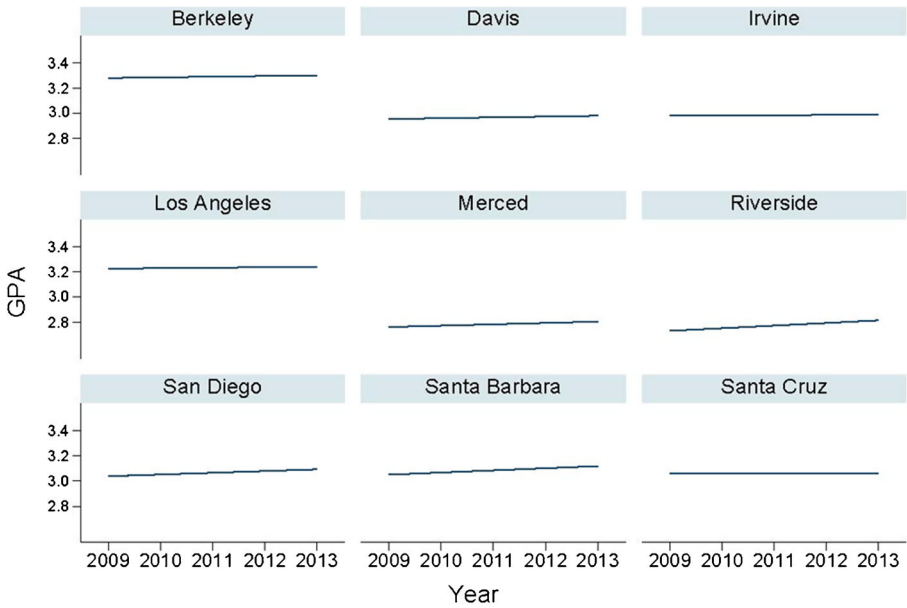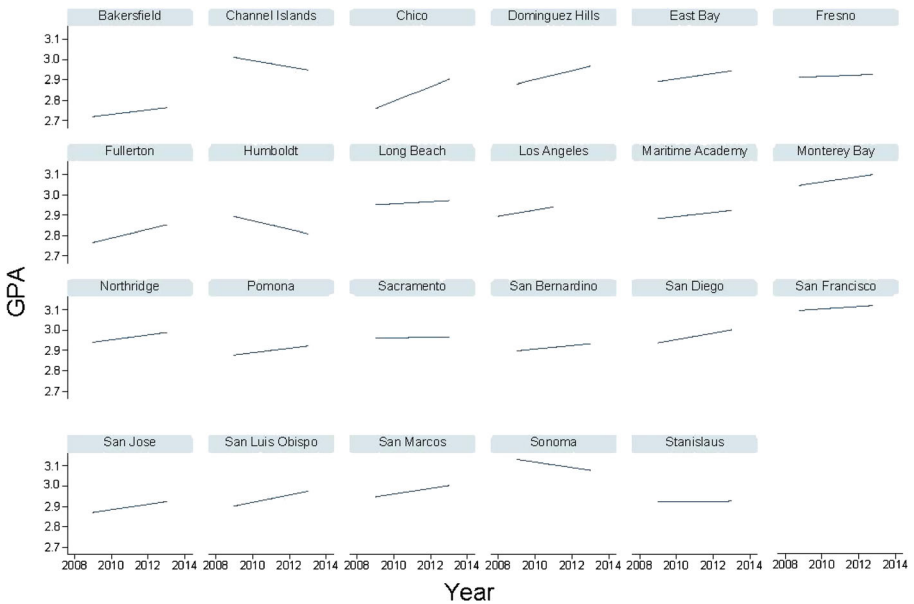**Figure 1** Grade Trends in UC Campuses: 2009–2013 (Fitted Values)

Figure 2 Grade Trends in California State University Campuses: 2009–2013

grade point, but with a *letter grade*. There is a potential problem in treating GPA as an exact proxy of a combination of letter grades. GPAs can be reified measures of competence since they represent different units of analysis (i.e., grouped data consisting of combined grade points) from what was measured in the unit of observation (i.e., a student's performance in a particular class). Letter grades are *translated* into numeric scores to determine grade point averages, at which point any semantic meaning of what a letter grade represents washes out. GPAs are thus—to some degree—numeric aggregates that are *less* than the sum of their parts. Let's examine this in more detail.

When GPAs are calculated it is under the assumption that 3.0 equates to a B, and that a B equates to work that is above average. However, in practice it does not follow that the semantic meaning of a student's competence equates to a numeric value. This becomes clear when examining grade definitions across universities: quite a bit of variance exists regarding definitions used to represent letter grades. Sometimes a B means "above average," but other times it means something else, such as "high pass." Again, these differences wash out when letter grades are translated into grade points. A 3.0 from one university is virtually the same in value and meaning as it is at another, but perhaps we should not be so quick to ignore the semantic meanings of letter grades and what they represent. Assuming that in every university "3.0" = "B" = "above average performance" misses crucial detail, and can even lead to a faulty interpretation of student performance. Let us use grading schemas across UC and CSU campuses to illustrate this more clearly, focusing on the relationship between *assessment scales* and *semantic grade categories*.

Assessment Scales and Semantic Grade Categories

It may be assumed that, for the most part, student assessment strategies are standardized and similar across college campuses. By standardized and similar we mean that the *assessment scale* (i.e., using the 0–4.0 grade point scale) and *semantic grade categories* (i.e., labeling an A as "excellent," a B as "above average," a C as "passing," etc.) used to represent student competence are the same across universities. A closer look reveals that this is only partly so. For example, while most universities use the traditional A through F assessment scale, not all employ a plus/minus system for each letter grade. Using our data to illustrate this, within the UC and CSU systems, UC Santa Cruz uses plusses and minuses for *some* grades (A+, A-, B+, B-, C+), but not all grades (anything below a C is not modified by a plus or minus). CSU Dominguez Hills has no D- grade in its A to F scale. These examples reveal that there is more

diversity in the assessment scales universities employ—and thus more diversity regarding how students are assessed—than has (perhaps) been taken into account.

More variability also exists across universities regarding *semantic grade categories* (or grade definitions). For example, among the twenty-three CSU campuses, *nine* different definitions are used for the grade A, *twelve* for the grade B, *eight* for the grade C, *eleven* for the grade D, and *ten* for the grade F. [7]

To illustrate, seven CSU campuses define a C as "satisfactory," two as "satisfactory achievement," and six as "average." That some campuses define a C as "average" where others do not is noteworthy. *Average* connotes that most students should fit this definition. Indeed, San Diego State's *General Catalog* states that a C is "the most common undergraduate grade," adding that a B is "definitely above average." [8] This implies that more Cs should be awarded than Bs. Defining grades as average or otherwise may seem irrelevant (after all, few administrators, professors, instructors, or students would challenge the notion that a "C" equates to "2.0" when determining one's GPA), but it raises a question: Do semantic grade distinctions affect how professors grade students? If a university defines a C as "average," which implies a set distribution and by definition connotes that most students should be assessed that grade in any given course, do those universities have GPAs that are closer to the numeric value associated with a C (i.e., 2.0) than universities who do *not* define a C as average? Our findings reveal that this is so.

Using our sample, we conducted a t-test comparing GPAs of campuses that define a C as "average" to campuses using some other definition for C. A t-test is a statistical procedure that compares the means of different distributions to reveal whether they significantly differ from one another. Our results showed that universities that define a C as "average" have significantly *lower* GPAs (and GPAs closer to 2.0) compared to universities that use some other definition for the C grade. One potential implication of this—though it cannot be determined given the nature of the data examined—is that when a C is defined as "average" professors seem to be more likely to grade as if a C is the most common grade given.

Again using our sample, we also found a significant difference in GPA across universities depending on how the F grade is defined. Some campuses label an F as "failing" or "failure"; others define an F as "unacceptable performance," "unacceptable work," "performance has been such that minimal course

---

[7] Differences among definitions are both substantial and subtle; for the purpose of this article we consider a substantial difference as defining the grade A as "Excellent" versus "Highest Level Performance"; we consider a subtle difference as defining an A as "Outstanding" versus "Outstanding Achievement."

[8] San Diego State University *2014–2015 General Catalog*, 468, http://arweb.sdsu.edu/es/catalog/2014-15/GeneralCatalog/!GeneralCatalog.pdf.

requirements have not been met," "non-attainment," "poor performance," "unsatisfactory achievement," or "performance of the student has been such that course requirements have not been met." Our results show that for our sample, universities that define an F as "failing" or "failure" have significantly lower GPAs than universities who use some other definition for the F grade.

Study Implications

The findings we present here help answer two nagging questions that continue to surface in higher education: Is grade inflation still occurring? If so, what contributes to its sustenance? Our answer to the first question is that grade inflation is still occurring, at least in some universities. However, we provide evidence that shows that grade inflation may be plateauing more than is commonly believed. Secondly, in examining the relationship between variability in semantic grade definitions and GPA, we hope that we have added to the literature that addresses the causes and correlates of grade inflation. To date, no studies on grade inflation have examined whether instructors assign grades differently depending on how letter grades are defined. Our research provides some evidence that variability in how grades are *defined* correlates with grade *outcomes*.

Many have offered strategies to curb the practices that lead to grade inflation, including implementing pass/fail grading systems, better articulating grading expectations, focusing on earned grades versus entitlement, and reporting median grades on student transcripts.[9] Our findings regarding the structural relationship between grade category definitions and grade distributions contribute to the understanding of why GPAs have increased over time. If universities carefully conceive the definitions they use for letter grades they may be able to reduce and even reverse grade inflation. Perhaps defining a C as average across universities—and *emphasizing* that it is the most common grade given—would lower inflated distributions. Other factors may explain more of the variance in grade distributions, but our results are compelling moving forward. Since our findings are preliminary, more research is needed.

It should be mentioned that the analysis (and data) in this study has limitations. The sample used in our study only addresses grading trends in universities on the West Coast (and only two of the university systems in this geographic region). Additional research is needed to determine if grade inflation in universities across the nation is beginning to plateau or decline. Furthermore,

---

[9]Jan Tucker and Bari Courts, "Grade Inflation in the College Classroom," *Foresight* 12, no. 1 (2010): 45–53.

respective grade distributions for each university in our sample are likely influenced by many variables beyond those mentioned, including students' socioeconomic status and level of preparedness before entering college. Even considering these limitations, our analysis provides clarity about assessment practices and the state of contemporary higher education. The grade inflation issue that many believe plagues American higher education might have an end in sight.