

Thesis

Sophia Yang

Introduction

In the wake of the COVID-19 pandemic, the Trinity College of Arts and Sciences at Duke University introduced academic policy changes. During the Spring 2020 semester, all courses were switched to Satisfactory/Unsatisfactory (S/U) grading as opposed to the traditional graded (A- F) scale. While this was the only semester where all courses were allowed to be S/U and count for all requirements, other policy changes during the pandemic regarding S/U still remain.

Referencing Duke University Undergraduate Bulletins, prior to Spring 2020, S/U grading was only permitted to count towards total graduation credits. After Spring 2020, S/U grading has additionally been permitted to count towards general education requirements (commonly known as “T-reqs”). Some faculty members have noted an increase in student usage of S/U and departments are moving towards assigning courses to be graded (A-F) only. The most recent of many changes has been a switch in defaults so instructors have to opt-in for allowing S/U rather than having to specifically opt-out of allowing S/U¹. With the rollback of most pandemic-era policies such as masking and social distancing long gone, this raises the question of whether the S/U policy should return to its pre-pandemic form.

A growing body of literature has examined the impact of the pandemic on student learning and a plethora of studies have been done on pass/fail grading in post-secondary education (although primarily in medical school). Rather than exploring impacts on student outcomes, this thesis seeks to determine how the changes in flexible grading policy have impacted student behaviors with the goal of better informing educators and administrators.

¹<https://dukechronicle.com/article/duke-university-arts-and-sciences-council-changes-satisfactory-unsatisfactory-grading-system-duke-athletics-changing-landscape-20250503>

Background and Context

Birth of Traditional A-F Grading

Education is notably missing from the US Constitution, but in the period between 1852 to 1918 all states had passed legislation requiring compulsory education (Diorio 2023)². As a result, K-12 enrollments nearly tripled from 1870 to 1910 (Goldin 2006)³. Simultaneously, the Morrill Acts of 1862 and 1890 provided federal land for the establishment of public colleges, opening the doors for access to higher education. Historically, students were evaluated by descriptions and evaluations given by individual teachers. Even entrance exams for college were made by individuals and occurred at individual schools, making results unreliable (Schudson 1972; Wechsler 1977)⁴⁵. The massive increase in the number of students required a revolutionary new approach to education that could scale with the increasing demand for education. A standard grading system was needed.

The movement towards the usage of report cards and grades as a success indicator became widespread. Instead, the rise of national examinations such as those by the College Board arose (Valentine 1987)⁶. Now, the shift was towards more consistent report cards across a school or district as a success indicator. Additionally, universities began using academic “credits” to quantify the amount of work and subsequently the amount of effort a student took on during a given period. Combined with grades, one could compare a student not just within their class but also with students in other classes in the same school. However, grading systems remained variable across the country in regards to what to measure and their frequency (Ashbaugh and Chapman 1925)⁷. Many debates arose ranging from the potential misinterpretation of grades (Bixler 1936)⁸ to the discrepancy between standardized tests and teacher’s grades (Hadley 1954)⁹ to the balance of extrinsic and intrinsic motivation (Sumner 1935)¹⁰. Yet despite

²<https://www.ebsco.com/research-starters/history/history-public-education-us>

³Goldin, Claudia, “School enrollment and pupil-teacher ratios, by grades K-8 and 9-12 and by public-private control: 1869-1996.” Table Bc7-18 in *Historical Statistics of the United States, Earliest Times to the Present: Millennial Edition*, edited by Susan B. Carter, Scott Sigmund Gartner, Michael R. Haines, Alan L. Olmstead, Richard Sutch, and Gavin Wright. New York: Cambridge University Press, 2006. <http://dx.doi.org/10.1017/ISBN-9780511132971.Bc1-509>

⁴Schudson, M. S. (1972) Organizing the meritocracy. *Harvard Educational Review*, 42(1), 34-69. <https://eric.ed.gov/?id=EJ056906>

⁵Wechsler, H. S. (1977) *The Qualified Student: A History of Selective College Admission in America, 1870-1970* (New York, NY: Wiley-Interscience). <https://eric.ed.gov/?id=ED150933>

⁶The College Board and the School Curriculum. *A History of the College Board’s Influence on the Substance and Standards of American Education, 1900-1980*. <https://eric.ed.gov/?id=ED285443>

⁷Ashbaugh, E. J. and Chapman, H. B. (1925) Report cards in American cities. *Educational Research Bulletin*, 4(14), 289-293.

⁸Bixler, H. H. (1936) School marks. *Review of Educational Research*, 6(2), 169-173.

⁹Hadley, S. Trevor (1954) A school mark-fact or fancy. *Educational Administration and supervision*, 40, 305-312.

¹⁰Sumner, R. G. (1935) What price marks?. *Junior-Senior High School Clearing House*, 9(6), 340-344.

the multitude of flaws grading had, the rapidly growing demand of education necessitated a solution.

During the early 1900s, research was conducted to determine the best way to assign grades. Studies such as that by (Starch 1913)¹¹ found that using a 100 percent scale was highly inconsistent across teachers. In response, some suggested a categorical system of “diagnostic letters” to reduce the impact of inconsistency on reported grades (Finkelstein 1913)¹². By the 1940s, the A-F grading system was adopted by over 80% of U.S. schools, rising in popularity along with the 4.0 scale (Schneider and Hutt 2013)¹³.

Rise of Pass/Fail

While A-F grading rose to prominence, its core problems remained. Grades were still inaccurate, being assigned differently across instructors, departments, and institutions. Additional studies found poor correlations between college grades and post-educational success (Hoyt 1966)¹⁴. Concerns about student learning under A-F grading were also raised. As put by Stallings and Leslie,

The undergraduate perceives grades as that proverbial sword hanging over his head which forces him to study content he otherwise might not study. The power of ‘the grade’ is strong enough to restrict his studying to material which he anticipates will be on tests... (Stallings and Leslie 1970)¹⁵.

Criticism of traditional A-F grading led to an era of educational innovation. Many schools began experimenting with alternate forms of grading, the most prominent of which was the pass/fail system. Pass fail grading was not a new idea, with records in American higher education from as early as 1851 (Smallwood 1935)¹⁶. However, it remained obscure until the 1960s and 1970s. Proponents of pass fail grading argued that it would foster an intrinsic interest in learning and greater exploration of academic courses. As put by Weller, it was hoped to “free the instructor and the student to communicate on a colleague to colleague

<https://www.jstor.org/stable/30174436>

¹¹Starch D (1913). Reliability and distribution of grades. *Science* 38, 630-636. Medline, Google Scholar
<https://doi.org/10.1126/science.38.983.630>

¹²Finkelstein IE (1913). *The Marking System in Theory and Practice*, Baltimore: Warwick & York.
https://books.google.com/books?id=4wsRAAAAMAAJ&lr=&source=gbs_navlinks_s

¹³Schneider, J., & Hutt, E. (2013). Making the grade: a history of the A–F marking scheme. *Journal of Curriculum Studies*, 46(2), 201–224. <https://doi.org/10.1080/00220272.2013.790480>

¹⁴Hoyt, D. P. (1966). College grades and adult accomplishment: A review of research. *The Educational Record*, 47, 70-75.

¹⁵Stallings, W. M., & Leslie, E. K. (1970). Student attitudes towards grades and grading. *Improving College and University Teaching*, 18(1), 66-68
<https://files.eric.ed.gov/fulltext/ED060054.pdf>

¹⁶https://books.google.com/books/about/An_Historical_Study_of_Examinations_and.html?id=OMgjAAAAMAAJ

basis” (1983)¹⁷. By the early 1970s, no penalty grading was present in some capacity in over two-thirds of a sampled 2500 American colleges and universities (Elsner and Brydon 1974)¹⁸.

Pass fail grading was not without faults. Multiple studies of the time found that pass/fail grading was often used to concentrate more effort in graded classes to boost or maintain grade point averages (Quann 1971; Collins and Nickel 1975)¹⁹²⁰. Whether this encouraged exploration outside the major is unclear with both positive (Sgan 1969)²¹ and negative (Johnson 1970; Weems et al. 1971)²²²³ reports. What was apparent was that students using pass/fail were less engaged in course material than their graded counterparts. In a study by Karlins et al, traditionally graded students reported completion of 80% of readings and 85% attendance as opposed to pass/fail students completion of 61% of readings and 74% attendance (Karlins et al. 1969)²⁴. Critics also argued against the binary extremes of pass/fail as well as highlighting administrative challenges regarding dean’s list, calculation of grade point averages, and transfer students. Schools thought that they needed traditional grades to motivate students and that grades convey important information about a student to future employers or higher level educational institutions. Regarding the intent to create bonds between instructors and students, Weller (1983)²⁵ found that pass/fail grading did not increase faculty evaluation time and institutions were divided on if it had a positive impact on faculty evaluation of students. Nearly 2 to 1 of the pass/fail institutions surveyed believed pass/fail grading did not result in a more positive student perception of grading.

Decline of Pass/Fail

While research had been conducted on education, the issue of education had largely remained out of the public eye until the 1980s. It was common belief that schools did not matter, and this was given scientific backing by the 1966 Coleman report which found that family background was more influential to student achievement than schools themselves (Coleman et al. 1966)²⁶. As a result, a relaxed attitude towards academics was commonplace and had provided the backdrop for introducing pass/fail.

¹⁷<https://eric.ed.gov/?id=EJ288937>

¹⁸<https://eric.ed.gov/?id=ED088549>

¹⁹<https://files.eric.ed.gov/fulltext/ED051737.pdf>

²⁰Collins, J. R., & Nickel, K. N. (1975). Grading policies in higher education: the Kansas Study/ the National survey. *University Studies*, 103, Wichita.

²¹<https://doi.org/10.2307/1979547>

²²Johnson, Jack T. Evaluate program, not grading. *College and University Business*, 49(3), 1970, 77-78.

²³Weems, John E., Clements, William H., Quann, Charles J., Smith, K., and Schefelbein, Barbara E. Pass -fail; Were the hypotheses valid? *College and University*, 46, 1971, 535 -556.

²⁴Karlins, Marvin, Kaplan, Martin, and Stuart, William. Academic attitudes and performance as a function of differential grading systems: An evaluation of Princeton’s pass-fail system. *Journal of Experimental Education*, 37(3), 1969, 33-50.

Available online at <https://www.tandfonline.com/doi/permissions/10.1080/00220973.1969.11011129?scroll=top>

²⁵<https://eric.ed.gov/?id=EJ288937>

²⁶<https://eric.ed.gov/?id=ED012275>

However, newly emerging research was beginning to suggest that schools did matter. In response to Coleman, the effective schools movement sought to analyze characteristics of schools that correlate with higher academic achievement. Edmonds (1979) expanded upon prior studies such as that of Weber (1971)²⁷ to analyze practices used by schools with high performing students and outlined characteristics of effective schools²⁸. As Edmonds put it, “We can, whenever and wherever we choose, successfully teach all children whose schooling is of interest to us. We already know more than we need to do that. Whether or not we do it must finally depend on how we feel about the fact that we haven’t so far”²⁹. His work was later expanded upon with additional and refined correlates of effective schools by others including Lezotte (1991)³⁰. Independent research from the UK by Rutter et al. (1982)³¹ further strengthened the case for better schools.

In 1983, the National Commission on Excellence in Education published *A Nation at Risk*. In this monumental report, researchers found consistent declines in high school and college student achievement scores and recommended high school graduation requirements (Gardner et al. 1983)³². The report describes the state of American education as “unilateral educational disarmament” and warns of a “rising tide of mediocrity”, capturing media attention across the country. Overnight, education became a nonpartisan issue. Pamphlets from the Department of Education made research more accessible to the public (1986)³³, the National Board for Professional Teaching Standards was established, exams began to shift away from multiple choice questions, and the first education summit of the nation’s governors was held (Ravitch 1990)³⁴. As a result of the growing importance of education to the general public, schools largely returned to a system of traditional A-F grading. The binary nature of pass/fail grading obscured the student data necessary to measure student achievement and improvement of the education system.

Rising Educational Attainment

Educational attainment in America rose sharply in the mid to late 1980s as college degrees became increasingly important. Papers from economists found that education was a way

²⁷<https://eric.ed.gov/?id=ED057125>

²⁸https://files.ascd.org/staticfiles/ascd/pdf/journals/ed_lead/el_197910_edmonds.pdf

²⁹https://files.ascd.org/staticfiles/ascd/pdf/journals/ed_lead/el_197910_edmonds.pdf

³⁰Lezotte, L. W., *Correlates of Effective Schools: The First and Second Generation*
<https://www.effectiveschools.com/Correlates.pdf>

³¹*Fifteen Thousand Hours*; Rutter, et al, Harvard University Press, Cambridge, MA
ISBN 9780674300262 <https://www.hup.harvard.edu/books/9780674300262>

³²<https://eric.ed.gov/?id=ED226006>

³³What Works. Research about Teaching and Learning. Department of Education, Washington, DC.; Office of Educational Research and Improvement (ED)
<https://files.eric.ed.gov/fulltext/ED263299.pdf>

³⁴*Education in the 1980’s: A Concern for ‘Quality’*

Diane Ravitch — January 10, 1990

<https://www.edweek.org/policy-politics/opinion-education-in-the-1980s-a-concern-for-quality/1990/01>

to signal and screen for high-ability workers (Stiglitz 1975)³⁵. Enticed by the promise of employment, more Americans obtained post-secondary degrees. According to data from the US Department of Education National Center for Education Statistics, the percentage of American adults aged 25 or older who held at least a Bachelor’s degree continuously rose from 6.2% in 1950 to 25.6% in 2000 to 37.5% in 2020 (2025)³⁶. A significant source of this increase was the rise of for-profit “diploma mills”. Fueled by the exploitation of financial aid and the political climate of deregulation and privatization, education became an industry. In 1990, there was only a single publicly traded for-profit university but by 2000 there were 40 publicly traded for-profit universities (Beaver 2017)³⁷. The Senate Committee on Health Education, Labor and Pensions found that “Between 1998 and 2008, enrollment at for-profit colleges increased 225 percent, compared to 31 percent growth in higher education” (2012)³⁸.

The goal of these for-profit institutions is to cut costs and grow profits. To do so, these institutions prioritized enrollment over teaching. They employed 10 times the recruiters for every career-service employee and hired mostly part-time staff (Senate Committee on Health, Education, Labor, and Pensions 2011)³⁹. As a result, the outcomes of students at these institutions has been subpar. Data from the Department of Education suggests that most for-profit career programs fail to benefit students with 72% of programs having graduates earning less than high school dropouts, compared to 32% at public institutions (2012)⁴⁰.

The Price of Education

In the emerging “credential society” social classes were distinguished by the degree which one held and the prestige associated with that school (Collins 1979)⁴¹. Degrees from elite schools acted as insurance for the future against rising the “fear of falling” of the middle class as household wealth inequalities rose (Ehrenreich 1989)⁴². Backed by impressive scholars and research contributions, admission into these elite colleges has always been challenging. Now, with the over saturation of degree holders, many would pay whatever price necessary for prestige to ensure financial stability.

At first, institutional rankings originated from athletic college affiliation (Ivy League) and regional primacy (e.g. Duke in the South, USC and Stanford in the West). By the end of the

³⁵Stiglitz, J. E. (1975). The Theory of “Screening,” Education, and the Distribution of Income. *The American Economic Review*, 65(3), 283–300. <http://www.jstor.org/stable/1804834>

³⁶Hanson, Melanie. “Education Attainment Statistics” EducationData.org, 2025-01-14, <https://educationdata.org/education-attainment-statistics>

³⁷Academe, Vol. 103, Number 1, in the January-February 2017 <https://www.aaup.org/academe/issues/103-0/rise-and-fall-profit-higher-education>

³⁸https://www.help.senate.gov/imo/media/for_profit_report/PartI.pdf

³⁹https://www.help.senate.gov/imo/media/for_profit_report/ExecutiveSummary.pdf

⁴⁰<https://www.ed.gov/sites/ed/files/policy/highered/reg/hearulemaking/2012/notice-proposed-rulemaking-march-14-2014.pdf>

⁴¹<https://archive.org/details/credentialssociet0000coll/page/n5/mode/2up>

⁴²<https://archive.org/details/fearoffalling00barb>

20th century, third-party ranking systems like that from U.S. News beginning in 1983⁴³. A dominant emergent strategy in the battle for the best students has been raising tuition and offering lucrative scholarships to high-achievers.

States have cut funding for public universities since the 1980s. This only accelerated in the 21st century. In response to the Great Recession (2007-2009) and mandatory spending programs like Medicaid, between 2008 and 2013 appropriation for the median public university declined by over 20% per full-time student⁴⁴. As a result, many public institutions were forced to increase tuition. A decade after the recession, state funding for higher education has not rebounded in most states⁴⁵.

According to analyses by Banks et al. (2024)⁴⁶, annual tuition and fees at private 4-year institutions during the 1979-1980 academic year was \$11,357 (adjusted for inflation), compared to the \$2,599 (adjusted for inflation) at public institutions. Over time this gap has widened to a difference of over \$20,000 by 2019-2020. By the 2019-2020 academic year, average annual tuition and fees at both public and private 4-year institutions had risen nearly 3 times the cost in 1979-1980, adjusted for inflation. Without adjusting for inflation, the cost of higher education has jumped 10-fold.

Grade Inflation

As the institutions changed, so did the students. Seeking to distinguish themselves from the increasing number of degree holders and limited by rising tuition, students sought to maximize their grade point averages (GPA) for future profit rather than of pure educational interest. The “entrepreneurial student” shopped “for bargain courses, encouraged by a faculty whose jobs are defined by “course load”, administrators who deal in credit hours as if they were coin, [and] institutions whose corpus evolves steadily into the corporate” (Haswell 1999)⁴⁷. Yet, not all gains in GPA necessarily match skill.

During the Vietnam War (1955–1975), college enrollment was used to avoid the draft. As a result, failing a student could directly result in their conscription. Evidence has shown an increase in grading leniency due to this policy (Bejar and Blew 1981; Birnbaum 1977)^{48,49}. In

⁴³<https://www.usnews.com/rankings>

⁴⁴<https://www.amacad.org/sites/default/files/academy/multimedia/pdfs/publications/researchpapersmonographs/PublicResearch>

⁴⁵<https://www.cbpp.org/research/state-budget-and-tax/unkept-promises-state-cuts-to-higher-education-threaten-access-and>

⁴⁶<https://law.stanford.edu/stanford-center-for-racial-justice/projects/private-universities-in-the-public-interest/private-universities-in-the-public-interest-white-paper/>

⁴⁷<https://www.jstor.org/stable/359043>

⁴⁸Bejar, I. L. and Blew, E. O. (1981) Grade inflation and validity of scholastic aptitude test. *American Educational Research Journal*, 18, 143–156.
<https://doi.org/10.3102/00028312018002143>

⁴⁹Birnbaum, R. (1977) Factors related to university grade inflation. *Journal of Higher Education*, 5, 519–538.
<https://doi.org/10.1080/00221546.1977.11776572>

a study by Rojstaczer and Healy (2012)⁵⁰ they found, “in 1960, as in the 1940s and 1950s, C was the most common grade nationwide; D’s and F’s accounted for more grades combined than did A”. By the end of the Vietnam War, As and Bs made up half to two thirds of grades in American colleges (Davidson 1975)⁵¹. After the conclusion of the Vietnam War, grades remained a measure of more than a student’s academics. Grades were affected by all manner of things from a teacher’s concern about student self-esteem, departmental policy to attract students, and the impact of grades during job search (Schneider and Hutt 2013)⁵².

One of the most predominant reasons for grade inflation was the rise of student evaluation of teaching (SET). SET first began to rise in popularity alongside the Civil Rights movement as a way for students to voice their complaints (Valsan and Sproule 2008)⁵³. Under the belief that student evaluations measure teaching effectiveness, administrators realized the opportunity evaluations presented to advertise their programs with some universities going as far as using evaluations as a component of consideration for promotion, tenure, and resource allocation. By 1980s, SET became commonplace in American higher education (Centra 1993; Wachtel 1998)⁵⁴⁵⁵. Yet, “the typical SET questionnaire treats the student as a customer and measures the satisfaction of the student with his or her professor, and not learning” (Crumbley 2010)⁵⁶. Multiple studies have found significant, positive correlations between student evaluations and student grades (Langbein 2008; Ellis 2003)⁵⁷⁵⁸. On the other hand, SET rankings are not significantly correlated with actual student learning (Uttl 2017)⁵⁹. An article by Neath titled “How to Improve Your Teaching Evaluations Without Improving Your Teaching” even goes so far as to suggest multiple methods such as getting evaluated before exams and grading leniently (Neath 1996)⁶⁰. Simultaneously, efforts to cut costs and increase profit margins resulted in a rise of nontenured, adjunct faculty (Bettinger and Long 2010)⁶¹. These educators’ careers depended significantly on SET rankings. As such, over time, professors became increasingly aware of the implications SET rankings could have on their careers.

Despite the increase in proportion of A’s, this does not seem to reflect an increasing caliber

⁵⁰<https://journals.sagepub.com/doi/abs/10.1177/016146811211400707>

⁵¹Davidson, J. F. (1975) Academic interest rates and grade inflation. *Educational Record*, 56, 122–125. <https://eric.ed.gov/?id=EJ122844>

⁵²Schneider, J., & Hutt, E. (2013). Making the grade: a history of the A–F marking scheme. *Journal of Curriculum Studies*, 46(2), 201–224. <https://doi.org/10.1080/00220272.2013.790480>

⁵³Valsan, C., & Sproule, R. (2008). The Invisible Hands behind the Student Evaluation of Teaching: The Rise of the New Managerial Elite in the Governance of Higher Education. *Journal of Economic Issues*, 42(4), 939–958. <https://doi.org/10.1080/00213624.2008.11507197>

⁵⁴<https://eric.ed.gov/?id=ED363233>

⁵⁵<https://doi.org/10.1080/0260293980230207>

⁵⁶<https://link.springer.com/article/10.1007/s10805-010-9117-9>

⁵⁷<https://doi.org/10.1016/j.econedurev.2006.12.003>

⁵⁸<https://doi.org/10.1080/00220670309596626>

⁵⁹<https://doi.org/10.1016/j.stueduc.2016.08.007>

⁶⁰Neath, I. (1996). How to improve your teaching evaluations without improving your teaching. *Psychological Reports*, 78, 1363–1372. <https://doi.org/10.2466/pr0.1996.78.3c.1363>

⁶¹https://scholar.harvard.edu/files/btl/files/bettinger_long_2010_does_cheaper_mean_better_-_impact_of_using_adjuncts_-_restat.pdf

of student. According to the 2019 National Assessment of Educational Progress High School Transcript Study, the average GPA has increased from 3.00 in 2009 to 3.11 in 2019 while over the same time period Grade 12 assessment scores decreased in mathematics and did not significantly change in the sciences⁶². Additionally, it is still debated as to how good of a predictor high school GPA is compared to standardized tests like the ACT and SAT. In a study comparing the predictive power of high school GPA against composite ACT scores, Noble and Sawyer (2004) found that across all levels of achievement, ACT scores provide greater differentiation than high school GPAs on success in the first year in college⁶³. In particular, “at 93 percent of the institutions, a student with a 4.00 high school GPA had less than a 0.50 probability of earning a 3.75 or higher first-year GPA” in higher education and “in some cases, HSGPA values less than 3.00 provided little differentiation in terms of students’ chances of achieving different first-year GPAs” (Noble and Sawyer 2004)⁶⁴. On the other hand, some studies show that ACT scores and high school GPA are both valid predictors of first year performance (Westrick et al. 2015)⁶⁵. Overall, there is insufficient evidence to suggest that the quality of students has risen, despite significant increases in GPA. GPA is no longer an effective tool for differentiating skill.

Falling Confidence in Higher Education

Grade inflation devalues education. The weight of a 4.0 GPA no longer carries the weight it once did. For colleges, participation in grade inflation lessens rigor, lowers quality of education, and degrades reputation (Chan, Hao, & Suen, 2007; Ehlers & Schwager, 2016)⁶⁶⁶⁷. For students, they are deprived of feedback and left unmotivated by lack of recognition of exceptional effort (O’Halloran & Gordon, 2014)⁶⁸. For society, grade inflation means graduates are unprepared for the workforce without the skills, dedication, knowledge, and work ethic desired by employers (Franz, 2010; Love & Kotchen, 2010; Yang & Yip, 2003)⁶⁹⁷⁰⁷¹.

With more degree holders and higher GPAs, degrees are no longer a sufficient edge over other job seekers to obtain employment. Simultaneously, the rate of tuition increase has outpaced wages. This discrepancy has not gone unnoticed. Multiple news sources have published articles

⁶²<https://www.nationsreportcard.gov/hstsreport/>

⁶³<https://www.proquest.com/docview/225613390?sourcetype=Scholarly%20Journals>

⁶⁴<https://www.proquest.com/scholarly-journals/is-high-school-gpa-better-than-admission-test/docview/225613390/se-2>

⁶⁵Westrick, P. A., Le, H., Robbins, S. B., Radunzel, J. M. R., & Schmidt, F. L. (2015). College Performance and Retention: A Meta-Analysis of the Predictive Validities of ACT® Scores, High School Grades, and SES. *Educational Assessment*, 20(1), 23–45. <https://doi.org/10.1080/10627197.2015.997614>

⁶⁶<https://doi.org/10.1111/j.1468-2354.2007.00454.x>

⁶⁷<https://econpapers.repec.org/RePEc:oup:cesifo:v:62:y:2016:i:3:p:506-521>

⁶⁸<https://link.springer.com/article/10.1007/s10734-014-9758-5>

⁶⁹<https://doi.org/10.1016/j.econedurev.2009.10.013>

⁷⁰<https://ideas.repec.org/a/pal/easeco/v36y2010i2p151-163.html>

⁷¹<https://www.asc.ohio-state.edu/yang.1041/grade-inflation.pdf>

with headlines such as “Price Of College Increasing Almost 8 Times Faster Than Wages”⁷². This crisis has been expedited by rising housing costs and other costs of living. In order to afford degrees, federal student loan debt increased by over seven-fold between 1995 and 2017 (Durk and Perry 2020)⁷³.

Increasingly, the American public has been losing trust in higher education. According to a 2024 Gallup poll, reported confidence in higher education has fallen since 2015 from over 65% down to 36% in 2024⁷⁴. Meanwhile the percentage of people reporting very little/no confidence has tripled from approximately 10% to 32%. Additionally, the gap in unemployment rates of Americans aged 25 and up by educational attainment has shrunk. What used to be a 5% difference in unemployment rates between those without a high school diploma and bachelor degree holders in 2005 is nearly halved in 2025⁷⁵. However, the relative difference in median income of high school graduates and bachelor degree holders has stayed roughly the same since 2004, adjusted for inflation⁷⁶.

Flexible Grading in the 21st Century

In the early 21st century, research and usage of pass fail grading remained largely obscure. The most prominent use of pass fail grading was found in medical schools. Doctors are expected to be lifelong learners, staying up to date with the newest techniques, treatments, and health problems. To do so, the character of a doctor must be taken into account, particularly their intrinsic desire to learn. Additionally, higher than average rates of stress and burnout had been reported in medical students and the negative effects of distress have been well studied⁷⁷. As a result, a wave of medical reforms were made including resident duty restrictions, self-development groups, and pass/fail grading systems.

In a 2011 review of pass/fail and well-being literature (1980-2010) in medical schools, it was found that all (four) of the papers reported improvement in some measure of well-being (stress, anxiety, depression, self-control, good health, level of satisfaction, group cohesion, and amount of free time) (Spring et. al)⁷⁹. Student satisfaction was measured and found to have increased in two of the papers⁸⁰. However, there were discrepancies in the long term effect of pass/fail

⁷²https://www.forbes.com/sites/camilomaldonado/2018/07/24/price-of-college-increasing-almost-8-times-faster-than-wages/?utm_source=chatgpt.com

⁷³<https://eric.ed.gov/?id=ED610721>

⁷⁴<https://news.gallup.com/poll/646880/confidence-higher-education-closely-divided.aspx>

⁷⁵<https://www.bls.gov/charts/employment-situation/unemployment-rates-for-persons-25-years-and-older-by-educational-attainment.htm>

⁷⁶<https://www.census.gov/library/stories/2025/09/education-and-income.html>

⁷⁷<https://pubmed.ncbi.nlm.nih.gov/10926029/>

⁷⁸<https://www.sciencedirect.com/science/article/pii/S0025619611610574>

⁷⁹<https://doi.org/10.1111/j.1365-2923.2011.03989.x>

⁸⁰https://journals.lww.com/academicmedicine/fulltext/2009/05000/a_change_to_pass_fail_grading_in_the_first_two.28.aspx

⁸¹https://journals.lww.com/academicmedicine/abstract/1995/04000/the_effect_of_pass_fail_grading_and_weekly_quizzes.19.aspx

with some claiming continued effect after the first semester while others found a return to typical levels in later semesters.

Spring et. al also reviewed an additional five papers (9 total) on pass/fail and academic outcomes (GPA, scores, residency attainment and performance). Grades were not found to be significantly different between pass/fail and tiered grading cohorts. Pass/fail cohort average significantly higher than the pass/fail cut-off in⁸². The pass/fail system was not found to adversely affect academic performance. However, acceptance into desired residency programs may be negatively impacted by pass/fail. In terms of residency attainment, roughly 73% of directors claimed they did not give preference to tier-grading schools and 33% of programs who filled all spots preferred students from tier-graded schools⁸³. Surveys of students and directors also showed majority belief that pass/fail evaluation hindered the ability to compete for residency (29, 14). Although, other studies have shown that some students believe grades are already arbitrary while others prefer grading for motivation⁸⁴. However, it appears that the actual impact of grading tiers may depend on the institution giving the grades as in practice when program directors were asked to compare Stanford's pass/fail classes with their own classes, only 3% of the Stanford graduates were judged as "poor" compared to their peer group⁸⁵. There was also evidence that pass/fail grading reduced competition and external motivation for grades without decreasing the amount of time students spent studying⁸⁶, defying expectations of increased laziness

Research supporting pass/fail and flexible grading continued to be published in the late 2010s. For instance, a decade long longitudinal study of medical student well-being found an 85% decrease in depression rate and a 75% decrease in anxiety in first-years when switching from a four-tier to two-tier grading system and restructuring their curriculum, among other changes (Slavin 2019)⁸⁷. Interestingly, some have also suggested the usage of pass/fail as a method of combating grade inflation (Blum 2017)⁸⁸ while others argue pass/fail is a cause of grade inflation⁸⁹.

In summary, there has been an increased interest in pass/fail arising from a prevalent attitude that educational reforms were necessary. Researchers have found evidence that flexible grading systems reduce student distress, support collaboration, and encourage intrinsic learning without having a substantial impact on academic performance and test scores (White and Fantone 2009)⁹⁰. However, there is also evidence that suggests that a two-tiered system makes it near impossible to distinguish between satisfactory and truly exceptional students, harming

⁸²<https://pubmed.ncbi.nlm.nih.gov/7718068/>

⁸³<https://pubmed.ncbi.nlm.nih.gov/7401142/>

⁸⁴<https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0240587>

⁸⁵<https://pubmed.ncbi.nlm.nih.gov/10353289/>

⁸⁶<https://pubmed.ncbi.nlm.nih.gov/5548587/>

⁸⁷https://journals.lww.com/academicmedicine/fulltext/2019/06000/reflections_on_a_decade_leading_a_medical_student.27.aspx

⁸⁸<https://core.ac.uk/download/pdf/132324597.pdf>

⁸⁹<https://www.jstor.org/stable/27797025>

⁹⁰<https://pubmed.ncbi.nlm.nih.gov/20012686/>

future career prospects as well as disincentivizing some students from putting forth their best effort and persevering through challenges. There is no definitive consensus in the literature about whether or not pass/fail is a better system than traditional A-F tiered grading.

Higher Education and COVID-19