Teaching Students to Feel Pleasure and Pain at the Wrong Thing: The History of Grades and Grading

Brian A. Williams

ABSTRACT: Despite their ubiquity and widespread acceptance in contemporary education, formal grading systems are relatively recent innovations in the history and philosophy of education. Far from innocuous tools which aid the student's academic development, grades and grading systems developed as *ad hoc* tools for ranking students against one another in academic competitions. This article examines the history of assessment, grades, and grading in light of the longer tradition of education and suggests alternative practices that could better orient students toward the true, good, beautiful, holy, healthy, and beneficial. By understanding how and why contemporary approaches to grades developed, classical educators will be equipped to mitigate the unintended and often unseen adverse consequences grades have on their students. Ultimately, this article seeks to liberate teachers and students to pursue the intrinsic goods of learning over against the fleeting and extrinsic rewards of making the grade.

THE PROBLEMS WITH GRADES

Academic institutions are sites of profound human formation in which a student who journeys from kindergarten through college will spend seventeen of the most formative years of her life. The normative practices that characterize these places will help her gain knowledge, learn skills, and prepare for gainful employment, but they will also form her affections and condition how she relates to herself, her peers, and the world outside her mind. In the words of sociologist Christian Smith, academic institutions are examples of "social worlds . . . thickly webbed with moral assumptions, beliefs, commitments, and obligations" that become embodied in rituals, practices, and policies, and which inform a

Brian A. Williams is Dean of the Templeton Honors College and the College of Arts and Humanities at Eastern University. He is the co-editor of *Everyday Ethics: Moral Theology and the Practices of Ordinary Life* (Georgetown University Press, 2019).

92–113
doi: 10.5840/principia20229276

student's "assumptions, expectations, beliefs, aspirations, thoughts, judgments, and feelings."[1]

One of the normative practices in contemporary educational institutions, classical or otherwise, and one that embodies an implicit commitment to competition and rank, is the attempt to quantify or translate learning and academic achievement into a single number or letter. Because this practice was unknown to educators before the nineteenth century, contemporary educators cannot consult the long tradition of liberal arts education for insight into the nature and use of grades nor to understand whether academic competition and ranking helps students achieve the ends of liberal arts education, namely, their integrated intellectual, moral, aesthetic, spiritual, physical, practical, and social formation.[2] The longer tradition did not use grades, nor were grades designed to help students achieve these ends. This ought to give classical educators pause and, given their relative institutional freedom, provoke them to evaluate the ongoing utility and necessity of the grading system, especially during the years before a student's grade point average (GPA) begins accruing in secondary school.

In what follows I first raise some concerns with grades and grading, narrate a brief history of the practice of grading and grading systems, and then briefly identify practices that could mitigate the adverse effects of grades on students. I devote space to the historical development of grading systems in order to explore how and why our current practices emerged, what they were designed to accomplish, and how consistent they are with the longer tradition of education. Contrary to many assumptions, our system of grades and grading are not the result of sustained pedagogical prudence or careful deliberation. They are late-modern *ad hoc* tools designed to efficiently rank students against one another in competitions for scarce prizes.

I am not the first or only person to question the value of grades and grading. Isadore E. Finkelstein and Mary Smallwood[3] wrote booklets in 1913 and 1935, respectively, questioning, in Finkelstein's words, the astonishingly "blind faith that has been felt in the reliability of the marking systems."[4] Similarly, in 1930, the great Black classicist and educator Anna Julia Cooper concluded,

[1]Christian Smith, *Moral, Believing Animals: Human Personhood and Culture* (New York: Oxford University Press, 2003), 8. See also James Davison Hunter, *Death Of Character: Moral Education In An Age Without Good Or Evil* (New York: Basic Books, 2000), 16–26; Charles Taylor, *Modern Social Imaginaries* (Durham, NC: Duke University Press, 2004), 23–30.
[2]These seven areas of formation orient students, in turn, to the true, good, beautiful, holy, healthy, beneficial, and neighborly.
[3]Mary Lovett Smallwood, *An Historical Study of Examinations and Grading Systems in Early American Universities* (Cambridge, MA: Harvard University Press, 1935).
[4]I. E. Finkelstein, *The Marking System in Theory and Practice* (Baltimore: Warwick & York, 1913), 1.

> we have been so ridden with tests and measurements, so leashed and spurred for percentages . . . that the machinery has run away with the mass production and quite a way back bumped off the drivers. I wonder that a robot has not been invented to make the assignments, give the objective tests, mark the scores and—chloroform all teachers who dared bring original thought to the specific . . . needs of their pupils.[5]

Similarly, the educational theorist Alfie Kohn has questioned the usefulness of grades in multiple articles, editorials, and books, arguing that grades and external rewards, contrary to expectations, actually demotivate student learning.[6] Other studies question how grades might influence a child's "academic self-esteem," which refers to how readily and deeply a child identifies as a "student," thinks herself "smart," or regards school as a place she belongs or enjoys. This identification seems to start uniformly high but diminishes in and after grade 3—often when grades are introduced—and diminishes most sharply in ethnic minorities, for whom lower grades more quickly and strongly reinforce negative cultural stereotypes about academic ability, and whose lower grades have in some studies been correlated with implicit and explicit teacher biases.[7]

Grades are not the only feature of a school that can negatively impact students, but the following list identifies several ways that grades often do. Grades:

- train students to love an extrinsic end ("good grades") rather than the intrinsic goods of learning;
- increase likelihood students will restrict their learning to what can and will be graded;
- decrease wonder, delight, and interest in what is being learned;
- increase competition and envy;
- decrease motivation to enroll in difficult courses or pursue projects that might harm one's GPA;
- increase temptation to cheat, because cheating make little sense without grades;

[5]Anna Julia Cooper, "The Humor of Teaching," in *The Voice of Anna Julia Cooper*, ed. Charles Lemert and Esme Bhan (New York: Rowman & Littlefield, 1998), 35.
[6]See Alfie Kohn, "The Case Against Grades," *Educational Leadership* (November 2011); Alfie Kohn, "From Degrading to De-Grading," in *What Does it Mean to be Well Educated?* (Boston: Beacon Press, 2004).
[7]See, for example, Dario Cvencek et al., "Self-Concepts, Self-Esteem, and Academic Achievement of Minority and Majority North American Elementary School Children," in *Child Development* 89, no. 4 (July/August 2018): 1105: "For example, compared to majority students, minority students are more likely to encounter negative stereotypes about their ability and intelligence, a scarcity of positive academic representations or role models, and teacher bias regarding perceptions of their classroom behavior."

- decrease student motivation to challenge oneself or pursue excellence if satisfied with a mediocre grade;
- increase anxiety and depression;
- decrease likelihood students will read a teacher's meaningful assessment of their work if both are delivered at the same time;
- increase the temptation for teachers to replace meaningful assessment with a less meaningful letter or number;
- decrease the granularity and usefulness of a teacher's assessment;
- increase the pretense of objectivity;
- decrease the extent and depth of useful communication between school and parents;[8]
- force the teacher to adopt the role of a judge delivering a verdict on past action, rather than a coach or master craftsman offering assessment and instruction useful for future improvement;
- decrease the chance students will pursue serious learning after graduation when that learning will no longer be exchanged for a scarce prize like a good grade;
- increase the tendency for older students to become rational consumers bargain hunting for the highest grade at the lowest cost.

These last two dispositions easily lead students to regard knowledge and understanding as the *means* for acquiring grades, which thus come to be seen as the end or *telos* of education. This happens when we treat grades and GPAs as valuable commodities or academic currency that students bank to trade in and trade up for other things like self-esteem, parental approval, honor from peers, college placement, scholarships, and eventually a degree. If most students are asked what they want to "get out of" a class, they will answer "a good grade." Parents and teachers reinforce this attitude every time they ask a student "how are your grades?" instead of more meaningful questions like "can you tell me about what you are learning?"; "what are you struggling to understand?"; "what do you think about that book?"; "what are you glad to have learned?"; or "what do you still have questions about?" Instead, students are asked, "how are your grades?" By doing so, teachers and parents communicate that grades are what they care about, what students should care about, and what the school is designed to produce. Likewise, teachers, parents, students, and schools treat grades like currency the school pays students in exchange for their academic labors, which students bank for the sake of future withdrawals and purchases. And though

[8]I am grateful for this insight to Robyn Burlew of Veritas School, Richmond, VA.

grades lose their value and become worthless once a degree is handed over, the damage endures in students who were led to value the accumulation of "grade currency" over everything else.

So it appears that grades, grading, and GPAs can order a student's affection toward the wrong things. In the *Nicomachean Ethics*, Aristotle rightly asserts that it is important to learn virtue from an early age and that virtue is bound up with learning to feel pleasure and pain toward the right things in the right way at the right time and for the right reasons. Conversely, vice is bound up feeling pleasure and pain at the wrong things in the wrong way for the wrong reasons.[9] Therefore, he claims, we should "educate the young by steering them with pleasure and pain" because "what is most conducive to virtue of character is to enjoy what one ought" (*EN* 1172a19–25). If educators habituate students to feel more pleasure from getting good grades and accruing grade currency than wonder, learning, and wisdom, and if students come to attach their self-esteem and shame to their grades and GPA, then the system of grades and grading potentially nurtures students toward vice, inhibits their flourishing, and thus undermines the express purpose of a classical liberal arts education. That this can happen implicitly is the point of Christian Smith's contention that people internalize the moral assumptions and beliefs embodied in the practices of thickly webbed social worlds like schools. If grades carry these negative effects, educators are right to ask about their origin, purpose, and utility.

MEDIEVAL TEACHING LICENSES AND LEAVING EXAMINATIONS

As we consider the historical development of grading systems—including the awkwardly overlapping four-point grading scale, A–F letter grades, and 100 percent scale—one question we need to ask is *cui bono*: whose good are grades designed to serve? [10] We can identify at least three agents whose ends grades could serve. First, there are intrinsically educational ends like student learning,

[9]Aristotle, *Nicomachean Ethics*, trans. Joe Sachs (Newburyport, MA: Focus Publishing, 2002). See 1099a7–22; 1104b5–1105a17: "Hence it is necessary to be brought up in some way straight from childhood, as Plato says, so as to take delight and feel pain in those things in which one ought, for this is the right education." On the importance of forming youthful habits of virtue in Plato's educational writing, see Mark E. Jonas and Yoshiaki Nakazawa, *A Platonic Theory of Moral Education* (New York: Routledge, 2021), especially chapters two and three on the importance of habits and avoiding youthful "encrustations."

[10]For additional aspects of the history of grades and grading systems, see Christopher Healy and Stuart Rojstaczer, "Where A Is Ordinary: The Evolution of American College and University Grading, 1940–2009," *Teachers College Record* 114 (July 2012):1–23; Jack Schneider and Ethan Hutt, "Making the Grade: A History of the A–F Marking Scheme," *Journal of Curriculum Studies* (May, 2013): 201–24. For a good overview of various contemporary critiques of grades and grading, see Lorin W. Anderson, "A Critique of Grading: Policies, Practices, and Technical Matters," *Education Policy Analysis Archives* 26, no. 49 (April 2018): 1–31.

intellectual formation, character development, and the nurturing of wonder and the intellectual appetite. Second, there are institutional ends like accurately ranking students against one another, enlarging class sizes, efficient "feedback," and performance tracking. Third, there are the transactional ends of the multi-tiered educational system, aimed at efficient communication, coordination, and movement of students within and through an interlocking national and international network of primary, secondary, and tertiary schools.

We will begin our story in the thirteenth century but move quickly into the eighteenth and nineteenth centuries. The importance of the eighteenth and nineteenth centuries in particular becomes clear in light of descriptions of grades as "a crucial expression of the modernist impulse"[11] and a representative of "the eclipse of traditional authority by bureaucratic rationalization."[12] When historian Christopher Stray summarizes how ranked grading enabled one famous examination to accurately identify, and publish the name of, the lowest scoring student, he describes it as a celebration of "the competitive system itself, dominated by a ranking procedure of unparalleled intensity and precision."[13] None of these descriptions of grades and grading should sound particularly promising to classical educators wary of industrialized bureaucracy, economic instrumentality, and toxic competitiveness.

As we will see, the history of grades is bound up with the history of universities, the history of examinations, the granting of diplomas, and social changes like compulsory schooling, child labor laws, immigration, school access for women and minorities, mobility, the G. I. Bill, the Vietnam War, and the recent consumerist approach to education. The following short section can do no more than sketch the basic evolution of grades and grading both because of space limitations but also because grades and grading did not develop in a linear way nor did they emerge from a single institution or educator. Instead, their emergence and adoption was haphazard, experimental, and diffuse, all of which reveal their *ad hoc* and improvisational nature.

[11]Schneider and Hutt, "Making the Grade," 202. This article is a good overview of the topic and contains a useful collection of early twentieth-century concerns about the inadequacy of grading systems and their adverse effects. However, the article's reference to the "European model" of competitive grading that was "largely used for pedagogical purposes" overstates the situation since there seem to be little to no "pedagogical purposes" served by the competitive grading system either in the article or in the historical record. The article references "motivating students," but it seems clear that grading systems were not designed for this purpose nor that this intent was broadly "European."

[12]William Clark, *Academic Charisma and the Origins of the Research University* (Chicago: University of Chicago Press, 2008): 14.

[13]Christopher Stray, "The Shift from Oral to Written Examination: Cambridge and Oxford 1700–1900," *Assessment in Education Principles Policy and Practice* 8, no. 1: 39.

The story could begin in multiple places, but one possibility is the University of Toulouse in the thirteenth century.[14] The history of medieval universities like Toulouse began with cathedral schools like the one that educated Alcuin in York, palace schools like the one Alcuin took over at Aachen under Charlemagne, and abbey schools like Hugh's school at St. Victor outside Paris. Cathedral and palace schools were designed to train boys to be literate priests for the church and civil servants for the king's court. Students were taught a version of the trivium and, if promising, the quadrivium. In 1079, Pope Gregory VII issued a decree requiring the proliferation and regulation of cathedral schools. By the next generation, cathedral schools were producing independent and itinerant scholars, most of whom would migrate toward royal or cathedral towns where there was a demand for literate persons to serve in the cathedral and the king's court. This is the story of the University of Oxford, which grew where it did in part because Henry II, whose father had been educated at the cathedral school of Laon, built Beaumont Palace in the cathedral town of Oxford, bringing church and court together, both of which required educated personnel.

These independent scholars eventually formed guilds or corporations like other trades, partly to regulate prices on tuition, rents, and vellum. The young "scholar," the teaching "bachelor," and the "master" or "doctor" roughly correspond to the "apprentice," "companion," and "master" of the trade guilds, respectively. Because *universitas* was a fairly common term for a corporation or body, universities began as the *universitas magistrorum et scholarium*, the corporation—or guild union—of teachers and scholars. And from these guilds in Bologna, Paris, and Oxford the world's first three universities, arguably, began.

After a "town vs. gown" riot in Paris, Pope Gregory IX issued the papal bull *Parens scientiarum* (1231), which conferred an independent status on the University of Paris and removed it from the supervision of local civil and religious authorities. Among other things, it meant the university could issue "teaching licenses" to graduating students without those graduates being examined for theological orthodoxy by non-university agents like cathedral school chancellors or bishops. This was followed in 1233 by another papal bull insisting that anyone

[14]The standard source for information on the medieval universities, subsequently corrected in some particulars, is Rashdall Hastings, *The Universities of Europe in the Middle Ages* (Oxford: Clarendon Press, 1895). See especially "The Organization of the Studium" and "The Universities of Medicine, Arts, and Theology" in vol. 1: 206–53. See also Stephen C. Ferruolo, *The Origins of the University: The Schools of Paris and Their Critics, 1100–1215* (Stanford, CA: Stanford University Press, 1985); Peter Jaeger, *The Envy of Angels: Cathedral Schools and Social Ideals in Medieval Europe, 950–1200* (Philadelphia: University of Pennsylvania Press, 1994); Olaf Pedersen, *The First Universities: Studium Generale and the Origins of University Education in Europe* (Cambridge: Cambridge University Press, 1997); Lynn Thorndyke, trans., *University Records and Life in the Middle Ages* (New York: The Norton Library, 1944).

formally allowed to teach in the University of Toulouse (which Pope Gregory IX had founded) must be allowed to teach at any other university (*jus ubicunque docendi*) without having to undergo further examinations. Other universities soon sought the same privileges for their teaching licenses.[15]

The teaching license was essentially a shorthand form of communication validating the scholar's achievement and abilities, confirming that he had been assessed and found capable of practicing the scholarly trade. In other words, his apprenticeship was complete and he should be accepted as a full member of the scholar's guild. The license was necessary because, unlike a furniture maker whose cabinets and chairs might validate his abilities, the scholar could not produce anything to verify his abilities other than a teaching license. These official teaching licenses were eventually called "diplomas," meaning "folded piece of paper," because traveling scholars would fold and carry their licenses from one university to another.[16]

As teaching licenses became increasingly important, so did the "leaving examinations" that students had to endure before receiving a license. This final examination included a series of oral disputations on set questions with other students and professors.[17] It was the most significant of a student's three assessments: 1) the daily oral examination and assessment by one's teachers; 2) the private examination and assessment of a graduating student by representatives of the faculty; 3) the final public lecture and disputation with other students and university scholars.[18]

We should note that though students were tested and assessed during the "leaving examination," they were not graded or ranked against one another. This is a crucial distinction: assessment is fundamental to education, but ranked grades are not. The former implies "sitting alongside" (*assidere*), helping students learn, while the latter implies "ranking against," establishing a hierarchy among them. Like the apprentice carpenter or stonecutter, the medieval student was released to practice his craft when he had mastered it to a level comparable to and determined by other masters. This absence of grading and ranking seems to have endured until at least the eighteenth century. There was no "grade point average" or class rank, neither were there "valedictorians" or "salutatorians" at commencement ceremonies.[19] And, as is still the case in many European schools,

[15]Many students attended university to secure valuable political and commercial contacts rather than a teaching license or degree.

[16]Similarly, a "diplomat" is one who also carries a "folded piece of paper."

[17]Humorously, the regulations at some universities required students to swear not to seek physical revenge if the leaving exam went poorly, and professors to swear not to summon students for exams in the middle of the night.

[18]See Clark, *Academic Charisma*, 97.

[19]Despite their Latinate names, these competitive commencement honors are relatively recent—and decidedly American—inventions. While "firsts" are notoriously elusive, it seems Jonathan Edwards may have been the first undergraduate to have offered a valedictory

colleges, and universities, early commencement ceremonies emphasized the university, the course of study, and the graduates' responsibilities and privileges, but not individual students or academic ranking. We should not miss the significant point: no educator at any level from the ancient Greeks through the Medieval era through the Renaissance thought that ranking students against one another or translating their intellectual achievements into numbers or letters served a pedagogical purpose or educational end.

THE BIRTH AND GROWTH OF GRADES

Starting in the eighteenth and nineteenth centuries, several developments contributed to the rise of the now-ubiquitous grading systems. First, examinations moved from oral to written. Second, schools began to competitively rank students in order to award scarce prizes, which encouraged more precise forms of assessment. Third, exams move from "leaving exams" taken just before graduation to end-of-year exams, to end-of-semester exams, and then to the innovative decision to let individual teachers determine when exams would be given in each class. Fourth, increased enrollment in schools and the desire for an efficient means of communication between schools led to the development of new techniques for quantifying assessment with simple numbers, letters, and words.[20]

One of the first steps seem to have been taken at the University of Cambridge in its famous Senate House Exams for honors B.A. candidates. During the eighteenth century, partly under the influence of Sir Isaac Newton (Professor of Mathematics at Cambridge until 1701), and partly due to the increasing availability of cheap pen and paper, this "leaving exam" developed into what became known as the Mathematical Tripos examination.[21] The advent of inexpensive writing materials enabled the exam to introduce more complex mathematical questions that could be written down and computed on paper.[22] Prior to this, the examinations had been conducted exclusively through oral disputations, small group questions, and individual interviews, and included questions in moral philosophy, natural philosophy ("natural science"), as well as mathematics. After this, the oral components largely became the means by which students were

address in 1720 at Yale. Though typically reserved for one of the masters of the college, Edwards was selected for this honor. Even so, it appears the term "Valedictorian" was first used at the College of William and Mary, when the top Latin student was chosen to give a "valediction," that is, literally, the "goodbye" speech, at the commencement ceremonies of 1772. Likewise, a "salutatorian" gives a "welcome" speech at commencement, from *salutatorius*, "pertaining to a greeting." Harvard (beginning in the 1760s) and Yale (beginning in 1815) also invited students to speak at commencement, but these students were selected by faculty and graduates rather than automatically by cumulative grade point average.

[20]See, among many sources, Clark, *Academic Charisma*, 112–17; 132–34.

[21]"Tripos" for the three-legged stool on which students sat.

[22]See John Gascoigne, "Mathematics and Meritocracy: The Emergence of the Cambridge Mathematical Tripos," *Social Studies of Science* 14, no. 4 (November 1984): 547–84.

seeded for the main event, the written exams, which became almost exclusively mathematical and increasingly competitive. Eventually, the Tripos developed into a multi-round tournament, in which losers dropped out and winners progressed through increasingly challenging problems. [23]

By 1753 at least, examinees were ranked into four groups: the top performers called "wranglers," hearkening back to the verbal wrangling of oral disputations; the second class was "Senior Optimes"; the third class, "Junior Optimes"; and everyone else, the *hoi polloi*, who became known as "polly-men," and, finally, 'Pollmen'.[24] Around this time, the Tripos exams also began individually ranking students in the top two categories, and publishing the names of at least the top ten. In addition, many of the top "wranglers" were offered positions within the university, and the top "wrangler" was apparently awarded a lifetime stipend from the University's endowment.[25] This increased the stakes because students now competed for fame and a scarce commodity. Furthermore, examiners suffered an increased temptation to show partiality toward students from their own colleges within the university. All these in turn increased the need for more precise and objective means for scoring and ranking.

Though sources are sparse, one suggests that these more precise means may have been first introduced by Professor William Farish (1759–1837). He was the Tripos "senior wrangler" and recipient of the Smith's Prize in Mathematics in 1778, Mathematics tutor from 1792, and Professor of Chemistry and Natural Philosophy from 1794, all at the University of Cambridge.[26] His 1837 obituary in the *Christian Observer* notes the telling detail: "He was the means of introducing into the University of Cambridge the system of classifying the candidates for a degree according to the number of marks obtained at their examination."[27]

[23]The rise of academic competition and "meritocracy" also seems to have emerged from a desire to challenge the traditional significance, even within universities, of the aristocratic oligarchy. However, aristocratic students were exempt from the Tripos.

[24]W. W. Rouse Ball, *A History of the Study of Mathematics at Cambridge* (Cambridge: Cambridge University Press, 1889), 170; Clark, *Academic Charisma*, 110; Stray, "Written Examination," 86.

[25]Clark, *Academic Charisma*, 110; Stray, "Written Examination," 86; Gascoigne, "Mathematics and Meritocracy," 553, 561.

[26]See Keith Hoskin, "The Examination, Disciplinary Power and Rational Schooling," *History of Education* 8, no. 2 (1979): 135–46.

[27]Anonymous, "Obituary of Rev. William Farish," *Christian Observer* 429 (1837): 611–13. Professor William Farish's interests resembled his father's, the Rev. James Farish, a vicar and natural philosopher who, among other things, is the source for our knowledge of Benjamin Franklin's experiments to calm turbulent water with oil. Farish and Franklin's mutual friend, Dr. William Brownrigg, forwarded Franklin a letter from Farish in which he describes his own attempt to replicate Franklin's experiment. Brownrigg describes Rev. Farish as an "old friend, a worthy clergyman at Carlisle, whose great learning and extensive knowledge in most sciences would have more distinguished him had he been placed in a more conspicuous point of view." In the letter, Rev. Farish expresses doubt about the

Though not explicitly named, this almost certainly refers to the Tripos given its prominence in the degree examinations, and it would be unsurprising given that Farish's academic work, despite being Professor of Chemistry, was almost exclusively in theoretical and applied mathematics.

However, several urban myths exist about Professor Farish and the invention of grades. One of these claims that he invented grades in order to increase his own financial compensation. This widespread story claims that Farish adopted letter grades (A, B, C, D, E, F) from a local shoe company that used them to rank the relative quality of its shoes. Because using simple letter grades decreased the time Farish had to spend on any one essay and exam, he could increase the number of students in his classes, and because he was paid per student, so the story goes, he could increase his salary. Several websites refer to him as the "world's laziest teacher," and several books and articles repeat these claims, each of which reference the others, all without a historical source.[28] Fortunately for Professor Farish, this story is almost entirely fabricated.

First, during Farish's time at Cambridge, students did not write essays and the only exams given were leaving exams (including the Tripos). Second, attendance at Farish's classes was optional, and the number of students who attended never influenced his pay. Third, as for his supposed lazy and acquisitive character, bent on rationally maximizing his profits, Farish was also an active abolitionist, organizer of the Cambridge Auxiliary Bible Society and the Missionary Society, and vicar of two of the poorest parishes in Cambridge—hardly the lazy, money-grubbing teacher of the urban myth.

The only truth to the story seems to be that Farish may have begun assigning numerical scores to written Tripos Examination questions, possibly in order to overcome the potential bias mentioned above and to provide a more accurate

reports of Franklin's experiments: "I suspect all of a little exaggeration." Another acquaintance of Rev. James Farish wrote that he "possessed more knowledge in several parts of learning than the generality of scholars possess in any one." See "To Benjamin Franklin from William Brownrigg, 27 January 1773," https://founders.archives.gov/documents/Franklin/01-20-02-0021. See also Joost Mertens, "The honour of Dutch seamen: Benjamin Franklin's theory of oil on troubled waters and its epistemological aftermath," *Physics Today* 59 (January 2006): 36–41.

[28]These stories may simply be embellishments of Neil Postman's passing reference to Professor Farish in *Technopoly* as the inventor of quantitative grading. Postman relies on Keith Hoskin's "Examination," whereas Hoskin's source is William Farish's obituary in the *Christian Century* mentioned above. The websites that repeat this erroneous story are numerous, but it can also be found in Kay Cheng Soh, "Grade point average: what's wrong and what's the alternative?" *Journal of Higher Education Policy and Management* 33, no. 1 (February 2011): 27–28; Thom Hartmann, *Thom Hartmann's Complete Guide to ADHD: Help for Your Family at Home, School and Work* (Nevada City, CA: Underwood Books, 2000); and Susan P. Giancola, *Program Evaluation: Embedding Evaluation Into Program Design and Development* (Thousand Oaks, CA: Sage, 2020), 32.

means for scoring and ranking competitors. However, many subsequent teachers do resemble the fictitious Farish, assigning grades because number or letter grades are easier and faster than other feedback, and because they have too many students to mentor, too many assignments to assess meaningfully.

Before leaving the Tripos, we should note that the system of scoring and averaging individual questions in order to produce precise final scores—regardless of whether it came through Farish—was 1) designed to *rank* students who had 2) freely entered a *competition* against one another, in which 3) everyone competed by taking the *same* exam that 4) centered on *mathematics*. Each of these points is significant for the accurate ranking that quantified grading was designed to facilitate. However, the system designed within these extremely narrow parameters has obviously come to be used indiscriminately for all manner of disparate examinations, subjects, and contexts that have nothing to do with competitive ranking. Furthermore, they have been imposed upon students of all ages, whether competitive ranking is helpful or harmful.

The next step came when universities adopted this basic method of scoring individual questions and exams and began examining students at the end of every year in order to determine who would receive scarce prizes, fellowships, and scholarships in the next. Harvard had adopted this practice at least by 1865. Then came examinations at the end of each course, retaining the graded ranking system even when it was disconnected from prizes. From there, professors were allowed to give examinations whenever they saw fit—a new practice allowed at Harvard at least by 1883. So the method designed to rank and compare students endured, even when accurate ranking and comparing became irrelevant and impossible.

Shortly before these developments in the United States, eighteenth-century Prussian grammar schools had developed a new technique that reinforced reductive assessment. These are the same schools that impressed the American educational reformer and Massachusetts Secretary of Education Horace Mann. The Prussians had made schooling compulsory in 1763, and under Horace Mann's influence, Massachusetts followed suit in 1852. Compulsory schooling significantly increased the numbers of students in school and pressured schools and teachers to develop new techniques for tracking their academic development. One of these new techniques was the frequent evaluation of student progress and performance, which was reduced to a descriptive word or two, entered on a grid, and sent home to parents. This was the precursor of the "report card," originally called a "Censur-Tabell" or "Schul Tabell." Students were not yet receiving letter grades, but they were receiving one-word evaluations for each academic subject and expected behavior.[29] For example, academic performances could be described as "good," "fine," and "decent," or "stupid," "mediocre," and "small." Behavior was "pious" or "disobedient." And "general abilities" were described in simple

[29]Clark, *Academic Charisma*, 119–22.

terms like "more memory than judgment," "slow in ability," "simpleminded," "speculative," and so forth.[30]

One implication is that communication from the school to parents about the progressive academic and moral formation of their children becomes generic and limited to shorts words that could fit within the confines of a small grid. It is not entirely clear that in the case of grade card grids "something is better than nothing." The individual report card may have been an improvement over another model, however, in which students of various ages and abilities were subject to daily and weekly examinations that resulted in reorganizing the seating chart according to student performances. The top performers literally moved their desks to "the head of the class" and the low performers moved to the back. The idea was that daily public competition and ranking—two recurring themes of this modern story—would extrinsically motivate students to perform their best. One can imagine that it did motivate the few top students who were capable of making it to "the head of the class," while also, of course, turning learning into a competitive sport that encouraged the vices of pride among some and despair among others. Instead, Mann, following the Prussian model, advocated for written examinations and monthly report cards that would accrue over time like "a merchant's ledgers" for his accounts, similar to the later "grade point average."[31] So we inherited individually scored examination questions and competitive academic ranking from Cambridge and the report card from the Prussians.

The history of grades eventually passes through the private journal of Ezra Stiles, the seventh President of Yale University (1778 to 1795). In his journals, President Stiles records the details of his private and professional life, including which year of students he examined in which subjects on which days. In a journal entry from December 1, 1782, he notes that at the request of a number of students, he began teaching Hebrew, and that he ranked the students into three divisions, presumably on the basis of their abilities.[32] This seems to have been rather standard pedagogical practice. However, in his journal from April 5, 1785, Stiles recorded that he examined fifty-eight seniors in Latin and Greek, and in his "college memoranda" about that examination, he notes that he ranked them

[30]Clark, *Academic Charisma*, 120. It is unclear when and where grades began being used in grammar schools generally, but Schneider and Hutt cite two examples from the U. K. that suggest ranked grading may have been used there haphazardly in the early nineteenth century ("Making the Grade," 204–5).

[31]Schneider and Hutt, "Making the Grade," 206.

[32]Franklin Bowditch Dexter, ed., *The Literary Diary of Ezra Stiles*, vol. 3, *January 1, 1782–May 6, 1795* (New York: Charles Scribner's Sons, 1901), 48. He also notes that on this day, there were 218 students present for some event: "the greatest number ever together at once in an American College."

according to results into one of four named categories: "Of these 58, 20 Optimi, 16 2nd Opt[imi], 12 Inferiores (Boni), 10 Pejores."[33] Here begins the four-point grading scale. By 1813, Yale had apparently translated Stiles' four categories into cardinal numbers, 1–4, and used them to represent how well graduating students did on their final examinations. By 1819, Yale was using quarter points, so 3.25 and 2.5 and so forth, and extended this system to track underclassmen. These marks were recorded in the faculty's "Book of Averages"—the beginning of the cumulative GPA.

It is worth noting that the cumulative average was kept secret from students precisely in order to avoid the competitiveness the university expected would occur if students knew their grades. In fact, as Lyman Bagg explains in his 1871 reflections on his time at Yale, the "merit marks" between 1–4 that were given for each recitation, were recorded in code using "a peculiar system of notation, known only to the officer, so that if by chance a student should get hold of the score-book of his division he would not be able to make out very closely the significance of the hieroglyphics contained therein."[34] If a student fell below a certain average, his "division master" would warn him that he needed to improve, or if he applied to know his grade point average, Bagg reports that he would be given "some such general information as that he is doing well, or very well, or improving, or falling off a little, or doing poorly."[35] Though secret throughout a student's career, the ranking was used at graduation to determine commencement honors, when it was revealed to each student individually. Thus Yale's Book of Averages functioned like Horace Mann's academic "merchant's ledger" and began the four-point scale and the cumulative GPA.

By the second half of the nineteenth century grades had become increasingly normative, even though there was no uniform system of symbols, numbers, letters, or words for efficiently translating and communicating academic accomplishment. For example, Yale moved from a four-point scale to a nine-point scale to a scale between 200 and 400 points, before returning to the original four-point scale derived from President Stiles. In the 1830s, when Henry David Thoreau was at Harvard, a "Scale of Merit" established by President Josiah Quincy scored everything in multiples of eight, which would accrue daily over a student's entire four years. In his 1888 *Harvard Reminiscences*, Andrew P. Peabody explains that Quincy instituted the scale in order to "reform the unmethodical way in which college rank had been determined."[36] In this system, "a student's daily record" would "constitute his due and fitly earned place in the scale of rank or merit."[37] Marks could also be lost for moral or behavioral infractions (and in 1820s and

[33]Ibid., 154.
[34]Lyman Hotchkiss Bagg, *Four Years at Yale* (New Haven: Charles C. Chatfield, 1871), 584.
[35]Ibid., 578–9.
[36]Andrew P. Peabody, *Harvard Reminiscences* (Boston: Ticknor and Company, 1888), 29–30.
[37]Ibid., 30.

'30s Harvard there were many, including drunken riots, brawls, cannon fire, brothel-visitations, and, of course, skipping chapel). According to Peabody, "this blended ratio of scholarship and character" determined a student's rank and his eligibility for endowments and honors. Even though Peabody was charged with keeping the records, President Quincy would examine them on a weekly basis "as if the most momentous interests were at stake."[38] According to Peabody, this system remained relatively unchanged for at least three presidents after Quincy. The maximum that could be earned over a student's four years at Harvard was somewhere between seventeen and twenty-nine thousand points or more, depending on how various essays and examinations were scored. Thoreau, who had a troubled relationship with Harvard, apparently left with a little over twelve thousand points.

These examples reveal that the systems of quantified grading were *ad hoc* improvisations that neither naturally emerged from nor were intrinsic to education. They were not designed to motivate students or to nurture their love of learning or intellectual formation. Instead, they were designed to serve the administrative ends of efficient tracking, simple communication, and competitive ranking.

By 1890, Harvard had abandoned Quincy's Scale of Merit, but began tracking student performance in each class by slotting them into one of five groups labeled A to E—the beginning of letter grades. Students were not ranked individually, but were ranked by group within each class.[39] By 1896, Harvard had an elaborate system for calculating the ratio of students' letter grades in order to determine who would graduate *cum laude*, *magna cum laude*, and *summa cum laude*—honors Harvard invented in 1869 and 1880. However, accurate ranking became impossible after Harvard's President Charles Eliot introduced the college credit-system in the 1860s and '70s. This system, which replaced the core curriculum model and allowed students to earn individual units of "college-credit" for each course they elected to take, was endorsed by the National Education Association in 1894 and the Carnegie Foundation for the Advancement of Teaching around 1910.[40] Because students no longer took the same sequence of courses, which meant that one student's program of study could be more or less difficult than another's, Eliot recognized that university-wide commencement honors would be impossible. He acknowledged that the university can only "provide academic honors at graduation for distinguished attainment in single subjects."[41] However, even though it became impossible to rank students' aca-

[38]Ibid., 31.
[39]Smallwood, *Early American Universities*, 51.
[40]The high school equivalent of the college-credit hour is the "Carnegie Unit." See Ellsworth Tompkins and Walter H. Gaumnitz, *The Carnegie Unit, Its Origin, Status, and Trends* (U.S. Department of Health, Education, and Welfare, Bulletin No. 7, 1954).
[41]Charles William Eliot, "Liberty in Education," in *Essays and Addresses* (New York: The Century Co., 1898), 145.

demic achievements against one another across the same university, the system designed to do so endured and now ranks one student's GPA against students from other academic institutions, courses of study, majors, and professors.

However, Harvard may not have been the first to translate academic performance into letters. The 1778 "moderator's book" from Cambridge's Tripos Examinations recorded the results of the oral disputations in mathematics and moral philosophy that were used to seed students for the final competition. The book from this particular year uses the notations "A+, A, and A-" to denote excellent performances, "E+, E, and E-" for good ones, "a+, a, and a-" for fair ones, and "e+, e" for mediocre ones.[42] It is not clear how these were used or combined for determining where students would be seeded in the final tournament, and it may be that they were simply shorthand notations to remind the moderator of a student's performance when he was involved in the seeding process. Neither is there any evidence that this use of letters influenced subsequent practices like those at Harvard a century later. Though the presence of letter grades in the 1778 "moderator's book" is interesting, it is Harvard's practice that becomes influential.

In sum, we inherited scored exam questions from Cambridge, the four-point scale and GPAs from Yale, and letter grades and graduation honors from Harvard. As an aside, this allows us to recognize the outsized influence possessed by institutions like Yale, Harvard, Oxford, and Cambridge, whose practices influence the culture and experience of nearly every other academic institution.

For all the relative significance of Cambridge, Harvard, and Yale, the system of correlating letter grades with individual assignments, rather than to groups of students, and to aligning them with the 100% scale seems to have begun with Mount Holyoke College. In 1896, again in order to provide a useful system for ranking students, Mount Holyoke combined letters, percentages, and adjectives that resemble the Prussian "Schul Tabell":

A = 95–100% Excellent

B = 85–94 Good

C = 76–84 Fair

D = 75 Passed

E = below 75 Failed

A year later Mount Holyoke modified the numbers slightly and added "F" for anything below 75%. After 1896, letter grades become increasingly ubiquitous and entrenched, and by the early 1900s, the letter grade system had spread into primary and secondary schools, as evidenced on millions of Prussian-inspired

[42]Ball, *Mathematics at Cambridge*, 171–72.

grade cards from the twentieth century.[43] Even so, the 1971 report of the National Educational Association records that as late as 1971, only sixty-seven percent of primary and secondary schools nation-wide used letter grades.[44] It is hard to put a date on the end of the "E," but it seems to have started disappearing around 1930, with the dreaded "F" much more easily standing for "Failure."

SOCIAL PRESSURES

The urgency to assess students with simplistic but efficient letters and numbers was solidified in the late nineteenth and early twentieth centuries when school populations boomed. For example, whereas in 1870, 7.6 million students were enrolled in public elementary and secondary schools in the United States, representing 57 percent of five-to-seventeen-year-olds, by 1930 there were 25.4 million, representing 81.7 percent of the same age group.[45] Causes for the dramatic increase include compulsory schooling, child labor laws, girls and minorities coming to school, massive European immigration, and the increasing economic advantages of education. After World War II, the G.I. Bill dramatically swelled numbers in colleges and made high school and college attendance newly normative. For example, in 1939 approximately 1.5 million students were enrolled in post-secondary education; by 1969 that number had jumped to 8 million.[46] Increased class sizes increased the need for efficient techniques to organize, assess, and track achievement, and to communicate that achievement between schools laterally as increasingly mobile students changed schools, and vertically as more and more students moved up into high school and college.

The influence of social contexts on the use of grades can also be seen in the fluctuation of grade distribution over decades. According to the analysis of Christopher Healy and Stuart Rojstaczer, "in 1960, as in the 1940s and 1950s, C was the most common grade nationwide" and "D's and F's accounted for more grades combined than did A's," but by 2009 "A" was "by far the most common grade awarded on American four-year campuses . . . even on campuses with students of modest academic caliber."[47] They also demonstrate that during the Vietnam War, grades were disproportionately awarded at the upper end of the scale, and suggest that professors may have inflated grades to ensure students did

[43]In 1902, Professor Herbert Mumford of the University of Illinois seems to have borrowed this grading system for standardizing the market classes and grades of cattle and beef, resulting in "Grade A" sirloin, and so forth. See Herbert Mumford, Bulletin 78, *Illinois Agricultural Experiment Station*: "Market Classes and Grades of Cattle with Suggestions for Interpreting Market Quotations."

[44]National Education Association, "Reporting pupil progress to parents," *Res Bulletin,* October, 1971, vol. 49:81–83.

[45]Thomas D. Snyder, ed., *120 Years of American Education: A Statistical Portrait* (National Centre for Education Statistics, 1993), "Table 8," 34.

[46]Ibid, "Table 23," 75.

[47]Healy and Rojstaczer, "Where A Is Ordinary," 5–6.

not drop out and become eligible for the draft. Similarly, though grade averages fell after the Vietnam War, the proliferation of top grades exploded in the 1980s and '90s without any improvement in academic attainment. This occurred as colleges began treating students as "consumers" and education as a "product" sold by the college. Like any business, colleges want to attract and retain happy consumers, and high grades keep students happy. Similarly, colleges began using student evaluations in these years and tying them to the retention, salary, tenure, and promotion of faculty. Higher grades produce happy consumers, who write favorable teaching evaluations, which increase a professor's chance for tenure and promotion. Unfortunately, recent research suggests that student evaluation scores are rife with bias and have little correlation with the quality of teaching and learning in a classroom.[48]

CONCLUSION

In conclusion, we return to the three parties who could be served by the *techne* or modality of quantified grades and grading: students, individual institutions, and the vertically tiered education system.

First, the student. No one, it seems, argues that letter or number grades serve the pedagogical function of forming students intellectually, morally, aesthetically, spiritually, physically, practically, or socially. Nor does anyone suggest grades are intrinsic to or especially useful for helping students nurture a posture of wonder, a creative imagination, intellectual appetite, depth of inquiry, verbal eloquence, intellectual honesty and humility, moral and spiritual seriousness, physical health, sensitivity toward beauty, concern for truth, or love of God, country, and neighbor.

At best, grades give the students an overly simple way to understand their proximity to or distance from the aggregate of knowledge, skills, or dispositions that their teachers expect them to develop. But it is not clear that quantifying this or reducing it to a number or letter is useful. If it were, we might expect coaches, for example, to grade their players at the end of every drill or practice, employers to grade employees at the end of the day, or master craftsmen to assign a number grade to each part of a piece of furniture made by their apprentices.

[48]See Robin Wilson, "New Research Casts Doubt on Value of Student Evaluations of Professors," *Chronicle of Higher Education* (January 16, 1998); P. B. Stark and R. Freishtat, "An Evaluation of Course Evaluations," *ScienceOpen Research* (September 29, 2014); John W. Lawrence, "Student Evaluations of Teaching are Not Valid," *American Association of University Professors* (May–June 2018); Justin Esarey and Natalie Valdes, "Unbiased, reliable, and valid student evaluations can still be unfair," *Assessment & Evaluation in Higher Education* 45 (2020): 8, 1106–20. Political scientists Mirya Hollman, Ellen Key and Rebecca Kreitzer maintain a list of similar studies at https://docs.google.com/document/d/14JiF-fT—F3Qa-efjv2jMRFRWUS8TaaT9JjbYke1fgxE/edit.

They do not, because doing so would not help players, employees, or apprentices know where or how to improve.

Similarly, using a single letter or number to communicate formative and summative assessment to a student's parents or guardians significantly diminishes the granularity of that communication, leaving parents and guardians with little to no knowledge of how to help their student, especially when that letter grade has been negatively affected by non-academic behavioral factors like poor attendance, tardiness, or turning in work late. For example, a "C" in English might mean that a student understands some things quite well and other things quite poorly, but the letter grade itself does not help the parent or guardian know where the student's strengths and deficiencies lie. In addition, the student might have consistently produced outstanding work in English, but also consistently turned her work in late—a character or behavioral fault not made apparent by an adverse academic grade.

One could argue that grades increase student motivation by dangling extrinsic rewards, like gold stars, though one entirely detached from the learning itself. It is a reasonable conjecture, and one suggested by some educators, but the persistent findings of Alfie Kohn and other educational researchers not only suggest that intrinsic motivation is longer lasting, but that grades demotivate students from pursuing what is most important, namely learning.[49] Instead, grades motivate good students to pursue "good grades" and demotivate others who either think "good grades" are unattainable or who are content with their "average grades." Grades were not developed to motivate students, and even when they do, they often motivate students simply to outperform their classmates. This risks capitalizing on either their pride or their insecurities, or both, and tempts them to predicate their well-being on being "better than" their fellow classmates. Finally, if "because it will be graded" is the only incentive a teacher can give student for doing an assignment, then it is likely a poor assignment or the teacher a poor pedagogue. And if modern institutions use grades to motivate students, it is instructive to note that many of the institutions that originally designed them intentionally kept them secret from students.

Second, the records indicate that grades were primarily developed to help institutions easily, efficiently, and accurately rank students against one another in a context of competitive scarcity. However, that this is neither intrinsic to education nor academic institutions is clear from the fact that universities existed for seven hundred years, examining and assessing students, but not ranking them with grades. This distinction between assessment and grading is essential for

[49]See Kohn, "Degrading"; "Against Grades"; and "Rewards Are Still Bad News (25 Years Later)," *New York Times*, October 28, 2018. See also Anderson, "Critique of Grading."

contemporary classical educators. To question the utility of grading and ranking is not to question the validity of assessment.

Third, grades could be seen as a simplistic *lingua franca*, designed for the most simplistic and efficient communication from a school to an admissions team at a selective university or program, a hiring committee, or a Human Resources Department. However, it is impossible to know how a 3.86 GPA from one institution compares with a 3.32 or 4.1 from other schools with different teachers, curricula, assignments, examinations, course requirements, and so forth. Therefore, it is worth considering whether the medium of letters and numbers is adequate to carry the meaning that the systems assumes and many claim. Morse code and Twitter may be fine for some things, but they are entirely inadequate for communicating Dante's *Commedia*, a love letter, a political debate, or Bach's Cello Suites.

In sum, grades were designed so that a school could competitively rank its own students against one another, which can be done accurately to a limited degree when students study the same curriculum and sit for the same examinations. When schools began grading students in this way, they used grades as a type of academic currency that could be traded in and traded up for scarce resources or prizes. Grades did not naturally emerge from the learning experience nor because they nurture student learning nor because they positively effect intellectual, moral, affective, or spiritual formation. Instead, their cumulative effect is a problematic disordering of students' loves, to use Augustinian language, or a training in learning to feel pleasure and pain at the wrong things, to use Aristotelian language. Both of these undermine the formation of virtue and lead to the problems identified in the article's opening. But when there is nothing to purchase with the academic commodity, as in kindergarten through most middle schools, or when receiving one's doctorate, then grades serve little purpose.

However, students are in an inter-locking system that trades on grades, at least from high school through graduate school. As they move up, resources like acceptance letters and scholarships become increasingly scarce, and so classical educators wrestle with how to use grades in order to enable students to move up through the system without allowing grades to undermine the actual intellectual, moral, spiritual, and practical goods and ends that educators think important.

However, given their relative independence, classical educators and institutions have more freedom than their public-school counterparts to develop practices that diminish the significance of grades and ranking in the souls of their students and parents, even if they still pay out grades to students who trade them in for seats at selective institutions of higher learning or increased financial scholarships. Several classical schools have decided to remove grades from their grammar schools altogether, and to minimize their impact on middle and high school students. Below I enumerate several practices schools have adopted to

obviate the adverse effects of grades on both students and parents, even if these do not fully resolve the difficulties:

- explain to parents how and why grades and ranking do not foster students' intellectual, moral, aesthetic, spiritual, physical, practical, and social formation;
- equip parents and teachers to ask students meaningful questions about their learning, rather than merely inquiring about grades;
- instruct students not to discuss their grades with other students, and ask parents to reinforce this;
- consider whether reducing formative and summative assessments to a single letter or number is conducive to the education of students in K–8 and, if not, consider discontinuing their use in these years, since their college-relevant GPA will not be tracked until high school;
- with younger students, replace simplistic grades with narrative subject area reports that describe aspects like the student's focus, work ethic, general character, and specific knowledge or skill that needs improvement;
- assign ungraded, and therefore low-stress, essays, quizzes, and exams, especially early in a semester, solely to help students learn and not to generate grades;
- separate assessment from grades by returning narrative assessment with a student's work, but waiting several days to deliver the letter or number grade on that work;
- require upper school students to petition the faculty or an administrator in order to gain access to their grades, and to see them only in consultation with the faculty, or perhaps their parent or guardian.
- replace the GPA with "mastery transcripts" that report on students' progress in specific areas of knowledge, skills, and character;[50]
- approach assignments and assessments the way coaches approach practice, treating assignments like forward-looking drills designed to help students improve;
- if a letter or number is placed on an assignment accompanied by useful assessment, place the letter or number inconspicuously at the bottom of the last page;
- if a letter or number does accompany an assignment, consider using only letters without "+" or "–" or a few select and carefully explained numbers,

[50]See the Mastery Transcript Consortium at http://www.mastery.org/.

e.g., 100, 94, 84, 74, 0, given that all grades, especially numerical scores, are approximations;[51]

- deliver assessments of major work orally to individual students;
- allow older students to have a voice in their grade for major assignments, or even the course, by supplying rubric for self-evaluation, discussed in personal consultation with the teacher;
- desist using class rank and GPA-based honors like "valedictorian" and "salutatorian," or replace these with faculty-determined honors for graduates who best embody, e.g., the true, good, beautiful, holy, healthy, beneficial, and neighborly, or some other set of goods or virtues.

These are just a few of the many ways classical schools and educators are diminishing the significance of grades and graded ranking in the educational experience of students and parents, recognizing that grades and ranking are practices that form the school's thickly webbed social world within which students are formed. Freedom from grades allows teachers and professors to help students look to the future rather than to the past, becoming like coaches or master craftsmen guiding young apprentices as they learn the craft or trade. This allows assessment to contribute to the student's holistic well-being, helps them come to love learning as a basic human good, and avoids harming them in order to serve the convenience of teachers and efficiency of institutions.

[51]These suggested numbers come from Joshua Gibbs, "A Medieval Catechism," *Circe Institute*, https://www.circeinstitute.org/blog/medieval-catechism.