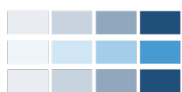


Where Have All the Zeros Gone? The Landscape of No Zeros Grading Policies

Dan Goldhaber
Maia Goodman Young
Luke Caldwell

August 2026

WORKING PAPER No. 345-0826



CALDER

National Center for Analysis of
Longitudinal Data in Education Research



Where Have All the Zeros Gone? The Landscape of No Zeros Grading Policies

Dan Goldhaber

American Institutes for Research / CALDER

University of Washington / CEDR

Maia Goodman Young

University of Washington

Luke Caldwell

University of Washington

Contents

Contents.....	i
Acknowledgments	ii
Abstract.....	iii
1. Introduction	1
2. Background on NZPs	3
3. Data Collection and Analysis	11
4. Findings	23
5. Discussion and Conclusions	31
References	38
Tables and Figures.....	45
APPENDIX A. Handbook Data Collection.....	51
APPENDIX B. Technical Appendix and AI Queries	53
APPENDIX C. Additional Tables and Figures	57

Acknowledgments

We would like to thank the Office of the Superintendent for Public Instruction (OSPI) for collecting and providing the publicly available administrative data we use. We also wish to acknowledge useful comments from Michael DeArmond, Joe Feldman, Thomas Guskey, Sarah Morris, Douglas Reeves, and Adam Tyner. This work is supported through a grant from a foundation that wishes to remain anonymous. Correspondence regarding this article should be addressed Dan Goldhaber, dgoldhaber@air.org.

CALDER working papers have not undergone final formal review and should be cited as working papers. They are intended to encourage discussion and suggestions for revision before final publication. Any errors are attributable to the authors.

CALDER • American Institutes for Research
1400 Crystal Drive 10th Floor, Arlington, VA 22202
202-403-5796 • www.caldercenter.org

Where Have All the Zeros Gone? The Landscape of No Zeros Grading Policies

Dan Goldhaber, Maia Goodman Young, & Luke Caldwell

CALDER Working Paper No. 345-0826

August 2026

Abstract

Debates about grading reform are intensifying, yet little is known about the formal policies that govern how teachers assign grades. This paper examines one prominent example – no-zeros policies (NZPs), which prohibit assigning zeros for assignments or course grades – by analyzing policy documents from school districts in Washington State and linking them to administrative data on district and school characteristics. We find that NZP adoption has more than doubled over the past decade, from roughly 30 districts in 2015-16 to nearly 70 in 2024-25, with the sharpest growth at the high school level. We develop a typology revealing substantial variation in how NZPs are structured, where they appear, and how they are communicated to families. Perhaps most striking, despite being framed as equity-promoting reforms, NZPs are less common in districts serving higher proportions of low-income students. We discuss implications for research and policy and call for greater attention to the design and distribution of grading policies as a foundation for evaluating their impact on students.

1. Introduction

Debates about grading and grade inflation are often in the headlines. Recent opinion pieces argue that “If Everyone Gets an A, No One Gets an A,” and propose “Yes, Grade Inflation Is Real. But is it a Real Problem?” (Alex, 2022; Flannery, 2026). This furor over grades coincides with concerns about weakening of grading standards and rising grade inflation, and the increased popularity of reforms aiming to make grading more equitable (Griffith & Tyner, 2025; Horowitch, 2025; Nierenberg, 2023; Piper, 2025; Sanchez, 2023).

There are meaningful disagreements about the purpose and mechanics of grading, and the degree to which we ought to be concerned about grade inflation (Brookhart et al., 2016; Denning et al., 2025; Flannery, 2026), but the intensity of the discourse around grades stands in stark contrast to the scarcity of evidence underlying it. Teacher grading *standards* matter for later student outcomes (Betts & Grogger, 2003; Denning et al., 2025; Figlio & Lucas, 2004; Gershenson et al., 2024), but most of the conversation in the media focuses on anecdotes or a general sense that the meaning and measures of grades are shifting. There is little documentation of the prevalence or the nuances in the *policies* that help shape grading standards: documentation that is an essential foundation to understanding the extent to which the landscape of grading is shifting. Without knowing the kinds of policies in place and their prevalence, we are hamstrung in evaluating their impact on students.

And yet there is some evidence that grading policies are shifting, giving rise to multiple alternatives to the grading scheme most adults are familiar with, which aggregates assignments into a 0-100 percent grading scale that is then translated into an A through F grade.¹ However, in

¹ While familiar, this grading system is far from universally used historically. Indeed, there is a long history of different forms of grading across the history of schooling in the United States (e.g., Kirschenbaum et al., 2021; Schneider & Hutt, 2014). Nor is this grading system simple or uniform. Depending on the teacher, course, school, or district, students may have all points for all assignments averaged to create a final percentage score, or certain kinds

the last several decades a range of alternative grading systems have emerged, adopted at the teacher, school, or district level – and in some cases formalized into policy. These systems often consist of a constellation of grading philosophies, practices, and policies that educators believe will make grading more motivational, accurate, and clear to students and parents. One issue that these systems often try to address is the issue of zeros in students’ grades.² Often this results in policies where teachers are prohibited from assigning zeros in some capacity for class assignments or course grades. As we describe in more detail below, these so called “no-zeros policies” (“NZPs”) vary widely across districts, which may affect how grades function as tools for communication and motivation.³

Recent survey evidence finds that around a quarter of teachers report some form of no zeros policy in their schools or districts (Griffith & Tyner, 2025) – a striking figure that suggests NZPs are far from a fringe phenomenon. However, these teacher reports, while informative, capture only whether an NZP exists, not what form it takes, whether teachers are responding to district versus school policy mandates, and when these policies began. To address these issues, we analyze documents from districts and schools in Washington State to identify the characteristics and variation of NZPs adopted, as well as when they were adopted. We then link these policy data to publicly available school system data to understand what kinds of districts and schools enact NZPs. Specifically, we address the following questions:

of assignments may be given more weight (e.g., final tests or papers may weigh more than homework). Additionally, the conversion from a percentage grade to a letter grade may vary by district or state. For instance, while often percentage grades are converted to letter grades on a ten point scale (90-100% equates to an A, for instance), until 2015, North Carolina had a seven point statewide grading scale where an A was 93-100%, a B was 85-92%, and so on.

² As with many elements of grading, the role of zeros is complicated. We expand on this further in Section 2.

³ We note here that we have a relatively expansive definition of NZPs that includes minimum grading where students cannot receive below some non-zero number on a 0-100 percentage scale, replacing zeros for missing or plagiarized work, and grading scales for assignments or courses that eliminate a zero or zero-equivalent (such as an F).

1. How has no zeros policy adoption in districts and schools changed over the last decade?
2. What is the nature and variation of NZPs adopted by districts and schools?
3. What district and school characteristics are associated with having an NZP?

We see a significant rise in NZPs in Washington State at all grade levels over the past decade: school districts with an NZP rose by over 110%, from just over 30 school districts in the 2015-16 school year to nearly 70 districts in the 2024-25 school year. We find that NZPs are much more common in elementary than secondary grades, and they are largely a district- rather than a school-level phenomenon. We also observe a wide range of policies that qualify as NZPs, but most districts use a grading scale that eliminates a zero, and a smaller subset mandates that teachers may not assign zeros when they don't have evidence of student work. Finally, we find that, all else equal, students in districts with higher proportions of low-income students are less likely to experience an NZP when accounting for year and district fixed effects.

2. Background on NZPs

Grades serve multiple purposes, so it is useful to begin by focusing on how these purposes are connected to grading standards and reforms.⁴ Following that, we discuss the limited empirical evidence on the effects of grading standards on student outcomes in general. We conclude with a discussion of the implementation of NZPs, how they are perceived, and a limited body of evidence on their prevalence.

⁴ It is important here to note that not all grading reforms include NZPs. Some scholars suggest that the problems caused by zeros can be eliminated by eliminating the averaging of grades and moving toward proportional grading scales (Brookhart & Guskey, 2020; Feldman, 2018, 2023; Guskey, 1994; O'Connor & Wormeli, 2011). However, eliminating zeros is a grading reform we see often in media discourse and in schools and districts, and, as we discuss earlier, it is relatively easy to enforce.

2.1 *No Zeros Policies and the Purposes and (Sometimes) Contradictions of Grading*

Broadly speaking, at the K-12 level various kinds of NZPs are often associated with the movements for equitable grading (Feldman, 2018, 2023) and standards-based grading (O'Connor & Wormeli, 2011; Reeves et al., 2017). Advocates of these reforms argue that – while grades *can* be useful for motivating students, communicating with parents, and signaling who needs help and who is ready for advancement – traditional grading undercuts these key purposes by demotivating students, imprecisely mixing cognitive and non-cognitive measures in ways that obscure information about student abilities, and varying teacher-to-teacher, thus sending unreliable signals. In response, they call on schools to adopt a constellation of reforms that aim to make grading more accurate, resistant to bias, and focused on academic mastery rather than behavior (Brookhart & Guskey, 2020; Feldman, 2018, 2023; Guskey & Jung, 2013; Link & Guskey, 2019). In reality, this can mean adopting a wide variety of policies and practices (Link & Guskey, 2022), but some form of NZPs are often (although not always) present in these reforms.⁵ Unlike other shared practices – such as encouraging teachers to use descriptive feedback or tying student performance to specific standards – the prohibition against zeros is concrete enough to be documented through sources such as board policies, district documents, and student handbooks, making it a particularly useful entry point for studying the prevalence of grading reform more broadly. Additionally, NZPs can often be enforced mechanically through grading software, meaning that they may be less likely to be mediated through principals' or teachers' interpretations of policy.

⁵ Not only do grading reforms vary in their components, but they vary in their fidelity to the pure application in the literature. Feldman, the author of *Grading for Equity*, notes that teachers may take elements of equitable grading and graft it onto traditional grading creating a “Frankenstein grading system” (Hess, 2024). NZPs in particular may be low-hanging fruit (Kunnath, 2022), which – contrary to the intent of grading reformers – may lead districts or schools to adopt these policies independently from other interconnected grading practices such as allowing retakes, weighting more recent work more heavily than less recent work, and reporting grades for behavior separately from grades for academic performance.

As we describe further in Section 3.2, NZPs vary widely in their construction and purpose. For instance, one purpose of grades is to motivate students, incentivizing learning that can be challenging and even unpleasant. Supporters of grading reforms that include NZPs argue that low grades demotivate students, and receiving zeros for assignments in a traditional 0-100 percentage-based grading system pulls a student's course grade down, making it mathematically hard to pass the class and demotivating students to continue trying.⁶ Instead, proponents suggest reconfiguring the grading scale to create equal-sized or proportionate grading categories (usually either by using a 0 or 1-4 scale or replacing zeros with a minimum grade of 50%, thus making each letter grade cover a range of 10 percentage points). This, they argue, makes it easier for students to recover from setbacks, thus motivating them to try even after initial failure.⁷ There is, of course, disagreement about this point. While some argue that there is no evidence that students are motivated by low scores (Guskey & Jung, 2013), others argue that the standard percentage grading system – complete with zeros – motivates students to try harder and learn more (Buck, 2022; Tyner, 2024). Following this logic, the way to encourage students to attempt

⁶ This is partly a mathematical problem; if a student must reach an average score of 60% to pass a class, a small number of zeros can have an outsized effect on that average grade, particularly if the student isn't doing well to begin with.

⁷ Not only do proponents of this sort of NZP argue that it increases student motivation, but they also argue that these policies create more accurate measurements of what students know and are able to do because they are more mathematically sound (Canady & Hotchkiss, 1989; Feldman, 2018, 2023; Reeves, 2004). As with many issues around grading, this is a longstanding (see Canady & Hotchkiss, 1989; Reeves, 2004) and contentious claim. Feldman and others argue that traditional 0-100 grading scales distort student achievement by creating a larger range for failure (usually 0-59%) which makes it harder for students to recover their grades after struggling for a portion of the year – even if they eventually achieve the standard being taught. Thus, they argue, the final grade a student achieves better reflects their mastery of a standard if it isn't dragged down by previous failures. However, others have argued that these grading practices are linked to grade inflation – a phenomenon that indicates these practices are *less* mathematically accurate (Tyner, 2024). Others argue that holding the minimum passing grade at a 60% encourages student mastery because students must achieve above a reasonable threshold to pass a course.

difficult learning is not to create proportional grading scales but rather to keep them, challenging students to work harder to receive higher grades.⁸

Another purpose of grades is to measure student learning and communicate those measurements accurately to stakeholders. This includes providing essential feedback to teachers and students to improve performance. Some proponents of NZPs argue that student scores should represent data on what students know and are able to do in relation to set standards. Thus, including zeros for work that is missing or plagiarized introduces inaccurate data into grading calculations in that this work is actually an absence of evidence rather than a measure of student ability (Feldman, 2018, 2023; Kunnath, 2022; Townsend, 2022). The solution, they argue, is to replace zeros with placeholders and institute non-grade related consequences.⁹ This, proponents argue, makes grades a more accurate measure of students' knowledge and abilities, rather than capturing soft skills like time management and teamwork, parental interventions, perceived student effort, or teacher biases. However, critics of such policies note that removing zeros from missing evidence of student work potentially eliminates important data about work habits and skills that enable students to succeed. They also argue that zeros can be a motivator for students to complete work that they had not previously turned in.

If grades are to accurately communicate measures of student achievement, they also must be standardized across teachers and subjects, so they are interpretable to stakeholders. However,

⁸ We note that supporters of grading reforms do not disagree with the idea that students should be challenged through rigorous grading standards. Indeed, many proponents argue that the reforms they propose *do* increase grading standards when implemented with fidelity (Feldman, 2018, 2023)

⁹ This is part of a larger issue of grades as a “hodgepodge” measure determined by school or teacher norms and comprised of both cognitive and noncognitive measures (Bowers, 2011; Brookhart et al., 2016). Indeed, Guskey and Link (2019) identify noncognitive elements such as work habits and student progress as being as much as 10% of student grades: the equivalent of a difference between an A and a B. Some scholars have argued that, rather than eliminate zeros for missing work, teachers should break the grade up and report different elements separately (Guskey, 2020). However, still others have argued that hodgepodge grades are useful as a composite measure of student readiness for post-secondary opportunities (Galla et al., 2019; Geiser & Santelices, 2007).

research documents clear inconsistencies in grading practices and standards across teachers (Brookhart, 1993; Guskey & Link, 2019).¹⁰ Proponents of NZPs also argue that eliminating zeros as a matter of policy serves as a mechanical way to standardize what teachers may put into the gradebook; thus hopefully creating more consistent measures of student achievement. For instance, one way to reduce inconsistencies in grades is to narrow the range of potential scores. NZPs that use proportional grading scales that go from one to four or from fifty to one hundred accomplish this. Another source of variation within grades is teachers making individual decisions about how to handle missing or plagiarized assignments. A district or school policy eliminating zeros for these sorts of student work will, by nature, standardize grades.¹¹

However, the motivational and communicative purposes of grading are complicated and may at times be at odds with one another (Hutt, 2025). While supporters of grading reforms argue that introducing practices such as eliminating zeros when assignments may be proxies for students' effort rather than achievement of academic standards (e.g., homework completion, participation in activities, or on-time arrival to class) provides a more accurate measure of student mastery of the material,¹² such changes also alter the incentive structure for completing

¹⁰ Feldman (2018) described how teachers interpreted their own grading policies as ideal: "Teachers with higher failure rates argued proudly that their grading reflected higher standards, that they were the 'real teachers.' A teacher with low failure rates explained that he was the only teacher who cared enough to give students retakes and second chances" (Feldman, 2018, p. xix). There is also quantitative evidence of this variation; for instance, Figlio and Lucas (2004) provide evidence that within-school teacher-level grading standards vary almost as much as the total variation in grading standards in the school district under study.

¹¹ This is far from the only way to gain consistency in grades across classes, teachers, and schools (Guskey, 2024). However, as we discuss earlier, NZPs are often enforced mechanically through gradebook software, meaning that they are relatively resistant to individual interpretation. While it may be preferable to have teachers engage in professional development or collaborative conversations about the purpose of grades and how best to standardize assessment practices, adopting grading software may present as easier for consistency to districts.

¹² These measures of effort, some argue, actually capture data about students' culture, and family situations rather than academic achievement (Calarco, 2020; Feldman, 2018, 2023). What counts as participation depends on cultural norms such as asking questions and taking turns talking. A student who is perpetually late for a morning class may have family obligations that prevent on-time arrival. Homework completion may indicate involved parents who provide directive support (sometimes even completing homework for their students) rather than students' own achievement.

work that leads to student learning. For instance, eliminating zeros for missing or incomplete homework will likely decrease the amount of homework students do, thus removing opportunities for students to practice and refine their skills.¹³ Because the purposes of grading can at times be at odds with one another, one cannot a priori conclude that one method of grading is better than another when it comes to motivating students or communicating student progress.

2.2 *Limited Research on Grading Standards and No Zeros Policies*

The empirical evidence on the effect of grading on students is limited, and there are very few studies that focus on grading *policies* as levers for improving student short- or long-term outcomes.¹⁴ A small empirical literature finds consistent evidence that higher teacher grading *standards* lead to greater student achievement later in a student's schooling (Betts & Grogger, 2003; Figlio & Lucas, 2004; Gershenson et al., 2024; Johnson, 2025; Mozzenter, 2019).^{15,16} Recent evidence based on following students in Maryland and Los Angeles (Denning et al., 2025), for instance, finds that students who are exposed to teachers who, on average, assign higher grades than might otherwise be expected based on other measures of student learning subsequently experience lower test scores, reduced chances of high school graduation, and lower

¹³ Proponents of alternative grading methods such as equitable grading argue that implementing grading reforms actually increases student homework and classwork completion. Feldman asserts that when teachers took up grading for equity, "Students completed assignments because they found that doing so improved their performance on assessments, not because of the homework or classwork points they could earn or lose" (Feldman, 2023, p. xliii)

¹⁴ There is quite a bit of research on the impact of feedback and assessments on students' motivation and learning (Black & Wiliam, 1998; R. Butler & Nisan, 1986; Hattie & Timperley, 2007; Shute, 2008). Additionally, there is some literature on the effect of grades and grading on student outcomes at the post-secondary level (e.g., M. Butler, 2025; Fallan & Opstad, 2012; Tan, 2023)

¹⁵ Grading standards are measured in different ways across studies, but generally grading standards are calculated by comparing student grades and test scores. When student receive the same test score but lower grades than similar peers, this indicates that they have been exposed to harder grading standards at the teacher or school level.

¹⁶ There is relatively little quantitative evidence on the mechanisms for how grading standards might affect achievement, but research at the college level finds a connection between higher anticipated grades and less time spent studying (Babcock, 2010). Other research finds that students on the margin between two letter grades who receive the lower of the two grades (for instance, a student who has an 89% and receives a B) are more likely to take easier courses and receive worse grades in subsequent semesters (Tan, 2023).

wages.¹⁷ Specifically, Denning et al. (2025) estimate that students who have a teacher who inflates grades one standard deviation above the average experience earnings reductions from \$42 to \$133 in each of the six years following high school graduation. While these effects may be seen as small, they are arguably consequential since teachers influence many students over their careers, and the effects may accumulate.¹⁸

There is limited evidence on the effects of grading policies on grading standards and student outcomes. The only large-scale study we found on the impact of grading policies on grading standards at the K-12 level in a U.S. context examines North Carolina’s 2015 shift from a seven-point grading scale to a ten-point scale. This change effectively lowers state grading standards, and researchers find that this more lenient grading policy leads to increased absenteeism and lower ACT scores among students at the bottom of the testing distribution, thereby widening gaps between student subgroups rather than narrowing them (Bowden et al., 2023).

Evidence on the adoption and impact of grading reforms in general and NZPs in particular remains limited.¹⁹ There is anecdotal evidence of increased adoption of NZPs.

¹⁷ Denning et al. (2025) distinguish this kind of grade inflation, which they term “mean grade inflation,” from grade inflation where teachers inflate grades at the margin of passing, which they call “passing grade inflation.” This latter form of grade inflation is associated with increased high school graduation and increased likelihood of enrolling in a two-year college.

¹⁸ The authors estimate that students in the classroom of a teacher who inflates grades by 0.2 points or more on a 4-point scale collectively lose about \$160,000 in lifetime earnings (Barshay, 2026). And this is for a single year’s worth of students – not the total population of students taught over the teacher’s tenure.

¹⁹ Research on standards-based grading and equitable grading show mixed results on the effects of these reforms. For instance, some researchers find that implementing standards-based grading lowers student performance and work completion (Huey et al., 2022; Townsley & Varga, 2018), while others find that standards-based grading is associated with *increased* achievement (Kramer et al., 2023; Pollio & Hochbein, 2015) and reduced test anxiety (D. Lewis, 2022). Research from The Equitable Grading Project (2024) asserts that students of teachers who received training in equitable grading practices (which includes NZPs) had higher rates of alignment between grades and test scores than students of teachers who did not receive the training. Note, however, that the Equitable Grading Project was established by Joe Feldman, the author of *Grading for Equity*, and is promoted by the Crescendo Group, a for-profit consulting organization also founded by Joe Feldman, which provides professional development for schools and districts to implement equitable grading.

Districts from Sacramento, California to Jackson, Mississippi to Portland, Oregon to Washington, D.C. have adopted NZPs over the last two decades, and survey data suggests the phenomenon extends well beyond these specific examples. In a Fordham Institute survey of nearly one thousand teachers, respondents were asked to identify so-called equitable grading policies in their schools including 1) prohibiting grade penalties for late work, 2) allowing unlimited retakes and revisions, 3) excluding homework and/or participation from grades, and 4) a mandated minimum grading policy (e.g., no assignment can get below a 50%) or equivalent no-zeros policy (Griffith & Tyner, 2025). About half of respondents report that their school or district has adopted some form of equitable grading, and twenty-seven percent of respondents say that their school or district has adopted some form of NZP.²⁰

Research on teacher and student perceptions of equitable grading policies and NZPs is also in a nascent stage. In Griffith and Tyner's (2025) study, over three quarters of teachers indicated that the practice of giving students partial credit for missing work (e.g., minimum grading) was either harmful or extremely harmful. However, a survey of 506 teachers in Arkansas identifies support for equitable grading among 51% of survey respondents (although, notably, this support varies across teachers' political ideology) (Morris et al., 2023). There is additional recent research suggesting that students in general prefer NZP or minimum grading policies—although preferences differ across student groups, with pre-AP students resisting NZPs and students experiencing high levels of stress preferring them (Morris et al., 2025).

But the empirical literature on the impact of NZPs specifically is quite small, with existing evidence drawn from single districts or small numbers of schools. For instance, Carifio and Carey (2012; 2015) study an NZP in a Massachusetts high school where students who

²⁰ The survey does not differentiate between the different kinds of NZPs that schools and districts may have implemented. We discuss these possible types of NZPs in Section 3.

received below a 50% on a course grade in a given quarter were instead given a 50%. This grade was then averaged with other quarter grades to calculate a final course grade. The researchers (2012; 2015) find that students who received a minimum grade in at least one quarter received higher state test scores than their peers with similar GPAs who never received a minimum grade. Carifio and Carey thus argue that minimum grading policies do not inflate students grades relative to test scores.²¹ The researchers also identify cost savings associated with students not needing to make up courses they had failed and find that the NZP has no impact on grade inflation or social promotion.²² These studies, however, are limited to a single type of NZP in a single school and do not, thus, capture the variety in structure and scope of NZPs described in the literature and enacted in districts and schools. In the following section, we develop a typology that captures this variation systematically.

3. Data Collection and Analysis

While research on the impact of NZPs is still emerging, the policies themselves, as noted above, are concrete enough to be recorded in formal policy documents such as school board policies, district documents, and student handbooks.²³ We collect these documents and employ natural language processing to query them about whether this reform exists in districts and schools.

²¹ In the NZP studied, students who received below 50% at the quarter had their grade raised to 50%. These quarter grades were then averaged across four quarters for a year-long class and two quarters for a semester-long class. This policy can be thought of as a “macro” NZP that makes changes at the term level, whereas the policies advocated by Reeves (2004), Feldman (2018, 2023) and other are often micro-NZPs, which occur at the assignment level. We discuss these variations in NZPs further in Section 3.

²² Morris and Johnson (2024) simulate the implementation of an NZP in Arkansas to create cost savings by reducing credit recovery costs and increasing graduation rates. They estimate an increase in cost-effectiveness from 6% to 23%. However, these calculations do not account for the costs of passing students who do indeed need remediation.

²³ As we discuss later, these *policies* are distinct from *practices* in that they are codified and identifiable in documents available to parents and other stakeholders. Practices, on the other hand, may be published internally, agreed upon informally, or take up by only a subset of teachers. We cannot currently determine much about grading practices as the district, school, or teacher level with publicly available documents.

3.1 Data Collection

Data for this study come from two main sources: the first is district board policies, student handbooks, and other district documentation that describe district and school grading policies. The second is publicly available administrative data from Washington State. We describe each data source in more detail below.

District Documents

For every district in Washington State, we gather three major types of documents: board policies at the district level, district documents such as webpages and parent resources, and student facing handbooks at the district and school levels. Because of differences in update cadence (e.g., handbooks update annually, board policies do not) and state mandates for reporting district information such as board policies, we follow different procedures for collecting each type of document.

Board Policies: District policy documents vary by state, but in Washington, districts are highly encouraged to have a board policy that describes a district’s rules for grades, progress reports, and transcripts.²⁴ These policies often include a statement as to the board’s understanding of the purpose of grading before detailing specifics including how grades and learning standards will be communicated with families and when grades may be withheld. Some board policies also include a grading scale, the minimum passing grade, a description of what is to be included on a high school transcript, and how GPAs will be calculated. Many board policies about grading conform to the sample language from the Washington State School Directors’ Association, although some districts’ policies are more specific, with details about the

²⁴ In Washington, these policies are often titled “Grading and Progress Reports,” although they may also be named such things as “High School Grade and Credit Marking,” or “Measuring Student Progress.” They can be found on district websites or on BoardDocs, which collect and store board policies, meeting recordings, agendas, and notes.

meaning of different levels in a chosen grading scale, rationale for taking up district-wide standards-based grading, or references to the theoretical grounds on which the grading policy is based.²⁵ Additionally, several board policies explicitly address concerns around equity. One district has a provision that “grades will be communicated in a manner that: 1. Accurately documents and communicates student progress toward meeting standards. 2. Provides timely non-biased feedback to facilitate reflection, self-evaluation, and growth. 3. Are inclusive and equitable,” and another district notes “Grading practices should be equitable and homework should meaningfully contribute to overall student learning.”

Because board policies record the dates on which policies were changed and adopted, we can track changes in board policy over time using board minutes and older versions of board policy. This allows us to identify starting dates for NZPs codified in board policy.^{26, 27}

We collect most board policies from district websites. When these policies are not available, we email superintendents to acquire the policies or to confirm that the district does not have a board policy regarding grading.²⁸

²⁵ For instance, most districts begin their board policy with a broad statement on the communicative purpose of grading like, “Our district believes that the cooperation of school and home is a vital ingredient in the growth and education of the student and recognizes the district’s responsibility to keep families informed of student welfare and progress in school.” Some district policies explicitly refer to other purposes of grading. One district lists several purposes of grading including providing feedback students can use for self-evaluation, communicating student progress toward standards, evaluating curriculum and modifying instruction, and identifying students who need extra support.

²⁶ In cases where we cannot identify definitively when a board policy changed using board documents alone, we use The Wayback Machine and web searches to confirm dates. When we cannot identify changes in policy using these sources, we contact district leaders.

²⁷ Because board policies can be changed in the middle of a school year, we determine that board policies passed in the second half of the school year – from January to August – apply to the following year. When we observe board policies that were adopted between September and December, we corroborate policy start dates using handbooks, news articles, taskforce documents, and conversations with district leaders. There are only two districts where the school board approves the policy in the first half of the school year. We do not see substantive changes in our regression results if we assume the policy was enacted in that school year or the following year.

²⁸ We initially see missing board policies from 54 districts. We email superintendents up to three times over the course of two months. We received responses from 24 districts and confirmed that 14 of these districts do not have board policies related to grading. We did not receive responses from 31 districts. During this process, as districts updated their websites, we find additional districts’ board policies. At the end of this process, we have 26 districts, 8% of the sample serving 1.8% of students, without board document related to grading. Twelve of these districts

Student Handbooks and Other District Documents: Because district policy language tends to be quite broad in many districts, and sometimes allows for discretion at the school level, it is not always possible to identify NZPs from board policy alone. Thus, another place where we look for information about grading is in student handbooks, which we collect at the district and school levels (when they exist), and other parent- and student-facing district documents about grading.²⁹

We collect these documents from district and school websites, but unlike board policies, handbooks are not listed in the same place across schools or districts.³⁰ Because of this inconsistency, we use a systematic search process where we first review navigation menus of school and district websites to see if there is a link to a student handbook. If we do not find a link, we use search tools embedded on the school or district website to search for a student handbook using key terms such as “handbook” and “grading.” If we still do not find a handbook for that school, we use a search engine to search for the school’s name, district name, and the term “handbook” (see **Appendix A** for details about the timing of these searches and the logic and assumptions behind our search strategy).³¹ With these documents, we use the school year in which they are published to identify the relevant dates the policy was in place. For instance, if we observe an NZP for the first time in a district-level handbook from the 2022-23 school year, we assume the policy was adopted then.

also do not have high schools. Given that these board documents often deal with high school credits and GPA, it makes sense that these districts might not have these policies and procedures.

²⁹ These documents include PDF “grading handbooks,” “frequently asked questions” pages, sample report cards, newsletters that go home to parents, and PowerPoint presentations from family information nights.

³⁰ Nor are handbooks always in an expected format. While many districts have PDF copies of their handbooks, other handbooks are in the form of Google Docs or dynamic webpages. We even find one handbook published through a newsletter platform used to communicate with families.

³¹ In our systematic search of district websites for handbooks, we also sometimes find data from grading taskforces that consist of administrators, teachers, and sometimes parents and students. These data include meeting notes, reports, and presentations. We include these data in the overall corpus because they are helpful in dating changes to grading policy.

In total, we collect handbooks and/or district grading documents for all but 5 of the districts in the state. We reach out to leaders at these districts and receive responses from two, confirming that there is no school or district handbook available. The remaining three districts serve 0.1% of Washington students. We discuss the ways in which we handle these different types of documents and the coding assumptions we make in Section 4.2.³²

Administrative Data

We merge information about grading policies from district and school documents with publicly available district and school administrative data from Washington State and from the National Center for Education Statistics Common Core of Data. These data include district and school size and urbanicity, student demographics including the number of students receiving special education services, the number of students receiving English language services, the number of students from underrepresented minority groups (including students identified as being American Indian/Alaska Native, Black/African American, Hispanic, Hawaiian or Other Pacific Islander, or two or more races), and the number of low-income students. We also include the numbers of students receiving a score of proficient or above on state standardized tests, and the highest grade-level offered in each district.³³

3.2 Document Coding and Typology Development Using Natural Language Processing

To identify districts with NZPs, we upload all the documents we collect into a database and use a process known as retrieval-augmented generation to identify and analyze passages

³² We intentionally exclude from our analysis alternative learning environments, online or virtual schools, and choice schools, which serve about 7% of Washington students.

³³ We do not have test score proficiency data from 2020, and the data we have for 2021 is from a fall administration of the state tests in the beginning of the 2021-22 school year.

relating to grading policies.³⁴ We describe this process in more detail in Appendix B. We then develop a typology of NZPs and apply that typology to the documents in the corpus.

Typology Development

We review a sample of policy documents and related literature to identify key dimensions of NZP policies.³⁵ This initial scan informed the development of a typology that we then used to categorize the policies. Our eventual typology has four dimensions: 1) the source of the policy, 2) the application of the policy 3) the grading scale used in the policy, and 4) the policy's treatment of evidence of student work.

Source. NZPs can be found in three different sources: board policies, district documents (including district-level handbooks), and school handbooks. In some cases, we find these policies in multiple places, usually in board documents as well as non-board related district documents. And at times, we find contradictory policies across different kinds of documents, which we discuss further in Section 4.

Application. Districts and schools can also implement NZPs for student assignments and/or student course grades. Assignment NZPs occur within a teacher's gradebook where an assignment that may have otherwise received a zero receives some other mark instead. This has the effect of raising a student's final grade, by increasing the average scores of the assignments

³⁴ This approach was selected due to its ability to utilize large language models for classification tasks while also providing suitable context to allow the LLM to make policy determinations grounded in primary source documentation.

³⁵ The literature on grading reform includes several variations on how districts and schools might eliminate zeros. These include policies that change the grades students can receive on assignments such as replacing a zero with a non-zero mark such as a 50% (also called minimum grading) and recalibrated grading scales such as a 1-4 scale or a scale that describes progress toward standards (e.g., "approaching standards" or "meeting standards") rather than using numbers. Similar policies occur at the course level such as those studied by Carifio and Carey (2012; 2015) where students receive a minimum grade of 50% at the quarter, which would then be averaged with their next quarter's grade. The grading reform literature also includes recommendations against assigning zeros to certain kinds of work, such as work that is missing, incomplete, or that violates academic integrity policies (Feldman, 2018, 2023; Guskey & Brookhart, 2019; Reeves, 2004) .

teachers aggregate to produce a final letter grade for the term. Course NZPs occur once a teacher has compiled all student scores into a summative grade. In these policies, the F students might have received for a course, or the zero that might have been included in their GPAs is replaced by some other mark. This effectively increases students' GPAs by omitting zeros from the overall average. In the case of the studies from Carifio and Carey (2012; 2015) described above, this could look like course grades at the quarter or semester being raised to 50% and then averaged with the grades students earn in the course during other terms. However, we also observe news articles about districts that eliminate all grades below a C or D for middle and high school students (Esquivel, 2021; Gonzales, 2024; Jones & Fensterwald, 2021). While it might seem that disallowing zeros for assignments necessarily leads to eliminating zeros for course grades, we find that this is not the case. As we describe in Section 5, we find cases where students *cannot* receive zeros on assignments, but once all assignments are tabulated, they *can* still receive an F – which equates to a zero in GPA calculation – for a course.

Grading scale. Policies can also vary in the grading scales they use. In some cases, students receive a minimum grade, such as 50%. In other cases, the 0-100 grading scale is replaced with a 1-4 grading scale or a scale that avoids numbers all together, replacing them with a range of marks that directly reference standards (e.g., “BE” for “below expectations” to “AE” for “above expectations”).³⁶ This also includes policies where a district only offers *course grades* of A through D, or that a district does not use zeros in student GPAs, instead giving students an

³⁶ We do find some districts with a *zero* to four grading scale. While some argue that the move from a 0-100 grading scale to a 0-4 grading scale effectively lowers expectations for students (Griffith & Tyner, 2025), we exclude these districts from our main analysis because they do allow for students to receive zeros. We note, however, that a 0-4 grading scale solves for a similar problem as a minimum grading policy that raises the lowest possible grade to a 50%. And we find that results are substantively similar when we include these 0-4 grading scale districts in our models, suggesting that there may be commonalities across districts that choose to adopt any form of alternative grading.

“NC” or “no credit” mark and excluding the course from GPA calculations. Thus, grading scale NZPs can be applied to assignments or courses.

Evidence of student work. The final dimension of this typology concerns whether teachers can assign a zero when student work is missing, incomplete, or academically dishonest. In some cases, this is part of a larger system of standards-based practices where teachers look at a collection of a student’s work toward a given standard and make an assessment that ignores missing evidence. In other cases, teachers enter a non-zero null score into the gradebook for any assignment where they cannot determine a student’s level of proficiency.³⁷ The assignment does not count for or against the student’s grade and it may or may not (depending on the district or school) also receive a flag or placeholder indicating that the work is missing, incomplete, plagiarized, or otherwise unacceptable as evidence of learning.³⁸

Document Coding

Using this typology and an LLM,³⁹ we initially identify 139 districts as possibly having an NZP or flagged as ambiguous, meaning either that the model had information about grading from the documents but could not make a clear determination or that there was insufficient data in the corpus. We then look to see if the *district* has one of several NZP types at the elementary, middle, and high school level by looking only at the district-level handbooks, district documents,

³⁷ These kinds of policies can be hard to categorize. For instance, what are we to make of a policy in a school handbook that reads, “Cheating will be cause for disciplinary action and a requirement to redo the work”? It is unclear from this language whether the student receives some amount of credit for the redone assignment. In cases like this, we infer that language such as “required to redo” or “must redo” as an indication that some credit is offered in exchange for the revised work; however, we note that this ambiguity may be confusing for families. We contrast this with districts that have policies mandating that students be *allowed* to make up missing work but do not preclude students from receiving a zero if they do not complete the work.

³⁸ While it is theoretically possible to imagine an *evidence of student work policy* at the *course level*, (e.g., a policy where a student’s final grade is withheld until they complete enough work to give a teacher an adequate picture of what the student knows and is able to do) we did not observe any examples of this in our sample.

³⁹ We run variations on this initial prompt across five LLM models; however, we find that only two – Claude Sonnet 4.6 and Claude Opus 4.7 – provide reliable outputs. Hand checks of discrepancies between Opus, Sonnet, and the other three models (Claude Haiku 4.5, GPT 5.4, and GPT5.4 Mini) reveal that the latter three models consistently return false positives, so we use only Sonnet and Opus to identify the 139 districts for further screening.

and board policies. If we determine that a district has a policy of a certain type at a given grade level, we assume that the policy applies to all schools at that level. However, we also observe NZPs in school-level documents such as handbooks. Thus, we also run the same queries for each school in each of the 139 districts, using these school-level documents.⁴⁰

Using the results of these queries, we classify a district as having a *district wide* NZP at a given school level if: 1) district documents such as board policies or district handbooks include an NZP for that school level; 2) school handbooks at a given level reference a district-level NZP with language like “per district leadership” or “as is district policy,” or 3) *all* school handbooks at a given school level reference some level of NZP. We also find cases where over fifty percent of schools at a given school level in a district reference an NZP. If the remaining handbooks neither confirm the presence of an NZP nor have a policy where students explicitly receive a zero, we make an arbitrary distinction that there is a district-level NZP.⁴¹ We provide more details about our logic for determining district versus school-level policies in Appendix A.

To ensure reliability and accuracy of our results, we run queries through two different LLM models – Claude Opus and Claude Sonnet. These two models have a 91% agreement rate for district verdicts and 95% agreement rate for school verdicts. When we identify disagreements in verdict, we hand-check each LLM’s reasoning and the underlying policies and decide which verdict to keep.⁴² When we find contradictions across document types, we default first to board

⁴⁰ Specific queries available by request. These prompts are designed to extract the following pieces of information for each query: (1) a categorical judgment representing an answer to the query (yes/no/unknown), (2) a confidence level (high/medium/low), and (3) explanatory notes documenting the model's reasoning process.

⁴¹ To illustrate this, imagine in a district with six middle schools, if we do not have any evidence of an NZP in the district’s board policy or district-wide handbook, we look at school handbooks. Imagine that in four of the six middle school handbooks, we see a description of a grading scale from one to four, with no zeros. If the remaining three handbooks do not include a grading scale, we assume that the district has a non-zero grading scale for middle school. However, if any of the remaining handbooks include a zero to four or a zero to 100 grading scale, we determine that this is a school-level NZP rather than a district-level NZP.

⁴² For the most part, we can make clean distinctions as to where each policy aligns with the dimensions of our typology. However, because of the nature of grading in the early grades, it can be difficult to assess from policy

policies, then other district documents, and finally school handbooks. We discuss the frequencies of these contradictions further in Section 4.

Once we identify the type and school level of NZPs in each district, we run historical queries through our database to identify when districts or schools adopted an NZP. In places where we do not have historic documents or where there are gaps between years, we contact district leaders to ascertain when the NZP began. In some cases, leaders are only able to give us a rough start date, often placing the start of the grading policy before their tenure at the district. In these cases, we indicate that the date is an estimate; however, for over 90% of the districts, we can identify that either an NZP is over a decade old or confirm the correct start date of the NZP. We do not identify any districts or schools in the state that adopted and abandoned an NZP during the period between the 2015-16 and 2024-25 school years; thus, the sample of districts with NZPs in the 2024-25 school year included the total sample that we observe having a no zeros policy.⁴³

3.3 *Analytic Sample*

To check our assumption that districts publish NZPs in board policies and other district documents including student handbooks, we email the superintendents of a randomly selected 20% of districts that we identify as not having an NZP. We received responses from 55% (26) of

documents the unit of analysis that elementary NZPs apply to. That is, we cannot always definitively tell if an NZP applies to assignments and courses or just courses in elementary schools. Across all elementary grading documents in our corpus, we observe that most discuss using a non-zero scale on progress reports and report cards, and none of them discuss using such a scale on assignments. Thus, we code elementary school non-zero scale policies as occurring at the course level, rather than the assignment level.

⁴³ The exception to this is COVID. Like many states, Washington issued specific guidance around grades for the end of the 2019-20 school year as schools transitioned to remote instruction (OSPI, 2020). These rules eliminated the possibility of failing a class, and some districts adopted a grading policy where only A's were permitted (Bazzaz & Long, 2020). Thus, every district in the state had an NZP at the end of the 2019-20 school year. However, there was no such prohibition against failing grades from the state in the 2020-21 school year. We do observe a small number of districts that adopted temporary NZPs for part of the 2020-21 school year while school occurred remotely but then returned to policies that allowed zeros when school restarted in person.

the districts we contacted.^{44, 45} Two of the 26 responding districts provided us with internal documentation of grading *practices* and *philosophies* that are consistent with no zeros *policies*.⁴⁶ An additional nine districts described some form of no zeros grading *practices* with no policy documentation.⁴⁷ This suggests that publicly available documents do not fully capture the prevalence of no zeros *practices*, even if they provide good information about formal *policies*.

The first two columns of **Table 1** describe the characteristics of the sample of districts with an NZP (Column 1) and without an NZP (Column 2) at any school level, weighted by the number of students enrolled in the district.⁴⁸ In subsequent columns we also report the weighted means of districts with and without NZPs at each school level separately. While all 295 traditional school districts in Washington have elementary schools, not all have middle and high schools, so the sample of schools varies by school level.⁴⁹

We identify 68 districts – just under a quarter of all districts in Washington – with NZPs. About a quarter of these (18) have NZPs at multiple levels. Across all school levels, students

⁴⁴ We send up to three emails and conduct up to three phone calls to obtain information about district grading policies. In our communications we explain the different kinds of policies we were looking at and ask superintendents or designees if their district utilized any of these policies at any level of schooling.

⁴⁵ Of these 26 responding districts, four indicated that they did not want to participate.

⁴⁶ The division between practices, philosophies, and policies is somewhat troublesome here. For our purposes, we assume that – regardless of what they are called – rules about grading that are publicly available constitute *policy*. If documents are *not* publicly available, we use the language used by the district (in this case, “grading practices” and “grading philosophy”)

⁴⁷ All nine of these districts reported some form of elementary no zeros grading practice, two reported middle school practices, and one reported high school practices.

⁴⁸ Because of the wide variety of school district sizes in Washington, weighting by student enrollment provides an indication of the grading policies that students in Washington tend to experience. Unweighted means are reported in Appendix C. These are mostly qualitatively similar, although they differ in magnitude and there are several results that are not statistically significant in the unweighted t-tests that are significant in the weighted ones and vice versa. The most notable difference is the comparison between district urbanities in the weighted and unweighted statistics. Specifically, city districts take up a larger portion of the districts that have NZPs; however, students in city districts who experience an NZP make up a smaller proportion of students in districts with NZPs.

⁴⁹ Cut offs for which grades count for elementary, middle, and high school vary by district. We consider a district serving students through high school if the highest grade it serves is 12th. We count a district as serving students through middle school if its highest grade is 7th, 8th, or 9th. Districts serving students only through 5th or 6th grade are counted as serving students only through elementary school. There are 295 districts with elementary schools, 283 with middle schools, and 252 districts with high schools.

who attend a district with an NZP are more likely to attend a suburban district and less likely to attend a district in a town or rural area. They are also more likely to attend larger districts with more students scoring proficient or above on state tests and smaller populations of low-income students.⁵⁰ The districts these students attend also, on average, have larger populations of students who are part of an underrepresented minority group and smaller populations of special education students. However, as we describe in our regression analysis, some of these correlations are not identifiable at statistically significant levels when accounting for other district variables.

While in most cases we observe NZPs *as a district phenomenon*, we do find 10 districts (7% of all districts in the state, which serve 6% of students in 2025) with within-district heterogeneity in NZP adoption. In other words, we see a small number of districts where there are *school specific NZPs* even when there is no district-wide policy adoption.⁵¹ We include these summary statistics in **Table 2** and note several similarities between the district and school NZP samples. Students exposed to NZPs at the school level are more likely to attend larger schools at the high school level, but smaller schools at the middle and high school levels. We do not observe statistically significant differences in student demographic characteristics.

3.4 *Analytic Methods*

To describe the associations between district characteristics and the adoption of NZPs, we estimate regressions on the relationship between traits and the probability of district d adopting NZP_{dt} in year t . We model the probability that $NZP_{dt} = 1$ using both logistic regressions

⁵⁰ Students who experience NZPs at the middle school level disrupt this trend in that they attend districts with *smaller* shares of students testing proficient or above on state tests and *larger* populations of low-income students.

⁵¹ Some of these districts *do* have district-level NZPs at one school level, but we only have evidence of a *school-level* NZP at another. For instance, one district has a district-wide minimum grading policy for high school but has a wide range of grading policies at the middle school level, with some schools employing standards-based grading with no zeros, and other schools using an F equivalent for work that falls below standard.

and linear probability models.⁵² The predicted probability of a district having adopted an NZP at a given level is the outcome variable in Equation 1:

$$p(NZP_{dt} = 1) = \alpha_0 + \alpha_1 X_{dt} + \alpha_2 Y_t + \alpha_3 Z_d + \varepsilon_{dt} \quad (1)$$

where NZP_{dt} is a binary indicator equal to 1 if district d has adopted any form of NZP in year t , and equal to 0 otherwise. In this model, X_{dt} includes district-level variables such as urbanicity, district size, and proficiency on state tests as well as student variables including the share of students receiving special education or English language services, underrepresented minority students, and students from low-income families.⁵³ The variable Y_t indicates school year fixed effects and the variable Z_d includes district fixed effects, which are included in some specifications. We run these regressions for all districts at any school level (e.g., elementary, middle, and high school) and separately by school level. We cluster standard errors at the district level as this is primarily a district-level intervention and weight models by district size.⁵⁴ We also evaluate the probability of a *school* having an NZP when the district does not have an overarching policy. The model parallels Equation 1, but it includes school characteristics rather than district characteristics, standard errors are clustered at the school level, and regressions are weighted by school size.

4. Findings

We begin by describing the change in grading policies over time. We then discuss the prevalence of each type of NZP and examples of the variety in policies we observe. We conclude

⁵² We use both logistic and linear models because, while logistic models make sense for a binary variable, it is more straightforward to include fixed effects in linear models, as logistic models suffer from the incidental parameters problem when including fixed effects (Chamberlain, 1980). Given that NZPs occur in under 25% of districts in the state, OLS results with fixed effects are also likely to be more accurate than logistic marginal effects (Timoneda, 2021).

⁵³ In some specifications we also include a binary indicator representing if a neighboring district had adopted an NZP in a prior year to identify potential policy contagion.

⁵⁴ Unweighted results are similar with a few notable differences, which we discuss in Section 4.

by describing the relationship between NZP status and district characteristics. Because NZPs are primarily a district-level phenomenon, we mainly focus our analysis on district NZPs, although we also discuss variation in school-level NZPs when relevant.

4.1 *The Rise of No Zeros Policies Over Time and Variation Across and Within Districts*

Figure 1 depicts the proportion of school districts in Washington with NZPs from 2015 until the 2025 school year, including both the percentage of NZPs overall and the percentages disaggregated by school level.⁵⁵ In every year NZPs are most prevalent at the elementary level, then at the middle school level, and then at the high school level. We also find a rapid increase over time in the proportion of districts with NZPs at all levels of schooling: since 2016, the number of districts with NZPs has more than doubled, from 32 to 68, reaching nearly a quarter of all school districts in Washington by 2025.⁵⁶ The increase in NZP adoption is most rapid at the high school level (from 1.6% of districts in 2016 to 7.1% in 2025), followed by middle schools (3.5% to 10.6%), and elementary schools (10.8% to 18.0%).⁵⁷

While NZP adoption is mostly a district phenomenon, we do see heterogeneous adoption of NZPs across schools within eleven districts. We see six districts where only one elementary school has an NZP, five districts where only one middle school has an NZP, and one district where only one high school has an NZP. Because of the limited size, it is difficult to draw conclusions about this sample. We include regression results for the full, elementary, and middle

⁵⁵ While we only report the proportions over the last decade, NZPs have been in existence in many districts prior to the 2015-16 school year. For instance, we find one case where a district has an NZP policy dating back to 1989. About half of districts with NZPs at the elementary level have NZP policies that predate the 2015-16 school year. A smaller percentage of districts with middle school NZPs or high school NZPs – 30% and 22% respectively – have had these policies in place for over a decade.

⁵⁶ These 68 districts serve nearly half of the students in the state; however, because districts implement NZPs at specific school levels, a far smaller proportion of students experience NZPs in any given year.

⁵⁷ As we describe previously, this is likely a lower-bound proportion of districts with no zeros *practices* at the elementary level because elementary schools tend to have less explicit policy information available. For instance, it is likely that many more than 32 districts use an elementary report card that doesn't include an F or a zero, but they do not have publicly available documentation of this practice.

school samples in Appendix C (**Table C3**). In general, students who experience these heterogeneous policies tend to attend schools with smaller shares of underrepresented minority students, and students who attend a middle school with an NZP are more likely to attend smaller schools.

Figure 2 displays the results of our typology, with each panel depicting a different dimension and each pie chart representing all 68 districts with NZPs. Each slice is labeled with the percentage of districts at the elementary, middle, and high school levels, as well as the percentage with NZPs at multiple levels.

Where We Find NZPs

A substantial portion of districts (38%, $n = 26$) only have evidence of an NZP in school-level handbooks, while the remaining 62% ($n = 42$) have evidence of NZPs in some kind of district-level documentation, be that a district documents (26%, $n = 18$), board policy (12%, $n = 8$), or both (24%, $n = 16$) (Figure 2 Panel A).⁵⁸ While there is not much research about how parents get information about school and district policies,⁵⁹ we can hypothesize that some sources such as school handbooks are more likely to be pushed directly to parents through email or physical copies sent home with students. We suspect that district policy documents, including district-level handbooks, are probably less accessible, but are still easily searchable on district websites, and we further speculate that board policies are likely the least easily accessed given that they often reside outside of a district's webpage and require some knowledge of how to navigate policy documents.

⁵⁸ Only two of the 16 districts with NZPs in multiple sources have evidence of these policies in school handbooks as well as district handbooks and board policies.

⁵⁹ While there is considerable research about communication between schools and parents, there is limited research about how parents get information about school and district policies. We find some literature about parent responses to policies about drug use, school issued personal devices, and bullying (Evans-Whipp et al., 2007; Lindstrom Johnson et al., 2019; Utter et al., 2022). We find no research studies describing how parents access or understand grading policies.

In addition to how policies might be accessed, we find issues around the consistency of the information within and across sources. For example, one district's board policy reads, "When there is insufficient evidence of learning to report a semester grade, an 'F' will be used for reporting student achievement," but five paragraphs beneath that is a grading scale that doesn't include an F, or any letter grade that equates to a zero in a student's GPA. We also observe eleven districts (16% of districts with NZPs) where there is explicit inconsistency between board policies and school handbooks, with the board policies describing NZPs and handbooks explicitly referencing the use of zeros. In one district, the board policy shows that an F grade is from 50%-59%, suggesting that there is some sort of minimum grading policy, but all school handbooks show that the lowest score for an F is 0%. And when a district has multiple schools at a given level (and thus multiple school handbooks), we find 15 districts with district-level NZPs in which grading policy is not uniformly shared across all school handbooks. In one district we find three high school handbooks with almost identical language explaining that, while district board policy has not changed, district leadership requires teachers to use minimum grading, saying teachers "will not give any grades lower than 50% on any assignment or assessment even for a missed assignment." This language is missing from the handbooks of the remaining high schools in the district, suggesting that there is perhaps an uneven distribution of policy information.

NZP Units of Measurement: Course and Assignment Level NZPs

Most districts with NZPs have them at the course level (88%, $n = 60$) (Figure 2 Panel B). This is largely driven by districts with elementary school NZPs that use a non-zero grading scale

for report cards.⁶⁰ While elementary course NZPs are most common, we see variation in course level NZPs adopted at the middle and high school level. To offer an illustrative example: one district doesn't assign Fs for course level grades until the semester report card while another district omits F grades entirely, instead only assigning students a mark indicating they have not received credit. These marks are not counted in the student's GPA.

Just under half of districts have NZPs at the assignment level ($n = 35$), with almost 40% of districts having NZPs at both the assignment and course level ($n = 25$). About half of the districts that have both a course and assignment NZP have both a non-zero grading scale for progress reports and report cards – a course level NZP – as well as some form of evidence of student work NZP, e.g., an assignment level NZP. For instance, one district has an A through D grading scale for high schoolers as well as a provision in their board policy that asserts that entering a zero for missing work creates inaccurate data and should be avoided. Another district uses a one through four standards-based grading scale and requires students to redo any work that violates the academic integrity policy rather than receiving a zero into the grade book. This amalgamation of policies has the potential to be more confusing and less transparent for parents and students who must track multiple rules that influence student grades.

Non-Zero Grading Scales

Eighty-four percent ($n = 57$) of districts with NZPs have a policy that eliminates zeros from either the assignment or course grading scale (Figure 2 Panel C). As described in Section 3, we observe considerable variation across policies. Some districts have policies that eliminate zeros from high school GPA calculations. In a district with this kind of policy, students can get

⁶⁰ Recall, as we discuss above, that it can be challenging to determine if a non-zero grading scale is used at the assignment level in elementary school, so we make the analytic decision based on the evidence we do have that these non-zero grading scales apply to course grades rather than assignment grades.

an incomplete or a no credit mark if they fail a class, but these marks are not included in GPA calculations. Other districts use a 1-4 standards-based grading scale for assignments. For high school credit-bearing courses, these numbers are usually converted into letter grades, which then may be translated back into a 0-4 GPA scale. In one district, for example, students in middle and high school are assessed on a scale of 0.5 to 4 for each standard. The scores for each standard are then averaged and converted to letter grades where average scores below 1.99 equate to an F, which then becomes a zero for the purposes of GPA calculations.

However, we have found several instances where this conversion is not completely clear. For instance, one district uses a *zero* to four grading scale for courses in middle and high school that aligns to letter grades A-D and “No Credit,” as opposed to an F. Unlike many board policies that include the conversion between letter grades and grade points, this district leaves it ambiguous as to whether the No Credit grade is factored into students’ GPA, simply stating that students do not receive credit for the course. In other cases, the conversion to letter grades may not be standardized. One district’s board policy states that “teachers may establish their own conversion scale from 1, 2, 3, 4 grades to letter grades,” which allows teachers to individually determine how to map a non-zero grading scale onto an A-F system.

Evidence of Student Work Policies

A much smaller percentage of districts with NZPs (35%, n = 24) have a policy that eliminates zeros when teachers are missing evidence of student work (Figure 2 Panel D). As with the non-zero grading scale dimension of the typology, within this broader dimension of our typology, these policies are quite varied. In some districts, students are required to make up missing work or work that violates academic integrity policies through a series of mandated support activities that persist until the work is completed satisfactorily. One district has a middle

school-level policy where any academic integrity violation, including cheating and the use of AI, leads to a required redo of the assignment rather than a zero. This policy is enforced by mandating that students attend support sessions before school, during lunch, or after school until the work is completed. Another district explains in its handbook, “The use of ‘zeros’ is an outdated practice and does very little in the way of teaching students actual responsibility. Therefore, students will no longer be assigned a zero for missing work.” The handbook then details the procedures that will be used to ensure that students complete their work to an acceptable standard, including mandatory study halls, phone calls and texts home, and even in-school suspensions.

However, in other districts, missing work is simply excluded from the grade book. In one district, students don’t receive any score at all for missing or incomplete assignments. Rather, these assignments receive a comment in the gradebook and are not counted toward students’ final grades. Another district assigns missing or incomplete work a score of two out of a potential four points and marks the assignment as “missing.” And still other districts vary these policies by the type of assignments. One district board policy reads, “All formative work is considered practice of content and skills and is not calculated in student achievement grades,” meaning that missing formative work – such as homework or classwork – would not be entered into the gradebook and therefore would never receive a zero. But another district has the opposite policy: zeros are disallowed for missing or incomplete *summative* work such as tests, essays, and final projects.

4.2 *Characteristics of Districts with NZPs*

Table 3 reports coefficients from linear probability models estimating the likelihood of a district having an NZP (overall and at the elementary, middle, and high school levels).⁶¹

Columns 1 and 2 describe the district and student characteristics associated with a district adopting NZPs at any schooling level, columns 3 and 4 describes the same for elementary school NZPs, columns 5 and 6 middle school NZPs, and columns 7 and 8 high school NZPs.

Across the different specifications, we find that few characteristics of districts – including the types of students they serve – are associated with NZP use.⁶² However, we do find a few exceptions. In looking at overall NZP adoption at any school level, the percentage of low-income students in a district is negatively associated with NZP adoption and this finding is significant in fixed effects specifications for all three school levels. A 10-percentage point increase in a district’s low-income student population is correlated with a 0.93 percentage point decrease in the likelihood that a district adopts an NZP. District location and size is also significantly associated only with high school NZP adoption. Holding size and student characteristics equal, rural, suburban, and town districts were all significantly more likely than city districts to adopt high school NZPs, with suburban districts being the most likely – roughly twenty percentage points more likely to adopt an NZP than city districts of a similar size and with similar student populations.⁶³ We also see that larger districts are more likely to have NZPs – a 1,000-student increase in district size is associated with a 1.01 percentage point increased chance of having a

⁶¹ OLS estimates are qualitatively similar to the marginal effects in models that do not include fixed effects. These are available on request.

⁶² An F-test of the joint significance of the district and student body characteristics is marginally significant in some specifications and not others. In no model does the F-statistic exceed seven.

⁶³ In specifications that include an indicator for NZPs in neighboring districts, there is a significant negative correlation between district policy adoption and an adjacent district’s policy adoption. However, this doesn’t account for districts in geographical proximity that do not share borders. And in specifications with *only* district urbanicity, we cannot detect a significant correlation between urbanicity and NZP adoption.

high school NZP. However, this effect is not significant in the fixed-effect model, suggesting that districts are not more likely to adopt high school NZPs as they get larger.

We also observe correlations between the proportion of English Language Learner (ELL) students and elementary NZP adoption. An increase in ELL students is associated with a slight decrease in elementary NZP adoption in the less restrictive model, but the direction of this relationship reverses in the fixed-effect model, suggesting that – while students in districts with larger proportions of ELL students are less likely to experience NZPs – districts *do* become slightly more likely to adopt elementary NZPs as their share of ELL students goes up – a 10-percentage point increase in a district’s ELL population increases its likelihood of adopting an elementary NZP by 0.20 percentage points.

There is also a positive relationship between the share of underrepresented minority students and the adoption of an NZP, a correlation that is largely driven by elementary school NZPs (see columns 1 and 3). These relationships, however, are not significant when district fixed effects are included in the models; the differences across specifications may be related to unobserved district-level differences, or it may be that there is too little variation in student characteristics over time to identify the effects in the fixed effects specification.

In the elementary school level fixed effect model (column 4), we also see significant negative correlations between elementary level NZP adoption and a district’s proportion of students receiving special education services, as well as the share of students scoring proficient and above on state tests. As the share of these student groups rises in districts, the likelihood of a district adopting an NZP decreases.

5. Discussion and Conclusions

This paper contributes to the literature on grading by directly analyzing policy documents in Washington State and linking them to data on the characteristics of schools and districts that

adopt them. We focus on no-zero policies, which are a loose collection of policies that prohibit teachers from assigning zero grades in a variety of contexts. These policies have become a flashpoint in debates over how schools should assess student achievement. Grounding this analysis in actual policy documents is essential because without them, debates about NZPs risk becoming untethered from the specific district- and school-level decisions that govern how teachers may use grades to assess and motivate students.⁶⁴

We find a significant rise in NZPs in Washington State at all grade levels over the past decade. The number of school districts with any type of NZP rose by over 110%, from just over 30 districts in 2015–16 to nearly 70 in 2024–25. The growth in NZP adoption is greatest in high schools, where we see a 350% increase in NZP policies since 2015. This is important as the high school level is arguably where grades have the greatest immediate impact on students’ post-secondary opportunities.⁶⁵ And because grades for high school coursework are linked to graduation and post-secondary opportunities, they can be highly motivating for students. They also provide important signals to admissions officers and employers as to students’ readiness for college and the workforce.

Interestingly, although NZPs are sometimes anecdotally associated with COVID-19, our findings suggest these policies gained traction well before the pandemic.⁶⁶ We also find that NZP reforms are not as straightforward or uniform as debates about them sometimes suggest.

⁶⁴ While policy documents are important, as we discuss later, we also recognize that they are only one of the many factors that influence teacher grading practices. Policy documents do not capture the amount or quality of training on grading that teachers receive, nor do they capture ways that teachers may try to mediate district or school grading policies within their own classrooms.

⁶⁵ This is not to say that grades in earlier years of schooling are unimportant. Indeed, grades at all levels of schooling signal to parents when students need additional intervention (B-Flation: How “Good” Grades Can Sideline Parents, 2023; Rury & Kalil, 2026). And the interventions provided by parents may be more important for students’ trajectories in the early years of schooling.

⁶⁶ Notably we do not see districts abandoning NZPs; rather it seems that this represents a steady adoption across the decade covered by this study.

Instead, we find that they vary along multiple dimensions, including where they can be found, their application (e.g., final course grades, assignments, or both), scale, and treatment of inadequate evidence of student work.

One of the more striking findings is that grading policies, including NZPs, are often difficult to locate, complex, and sometimes contradictory. School systems post grading policies across a range of locations: school board policies, district documents, and school-level student handbooks,⁶⁷ and individual policies range from a single line to multiple pages. As a result, families and perhaps even teachers may not always know where to find descriptions of how grades work in a given school.

Inconsistency compounds the problem. For instance, in some districts, policies appear in some school handbooks but not others. In one district, a high school NZP appeared in only a quarter of high school handbooks, leaving a question as to how students and parents at the remaining high schools learn about the policy. Additionally, policies sometimes conflict across sources: one district maintained an NZP in its board policy while high school handbooks explicitly permitted zeros. And the language of policies themselves can generate confusion: some eliminate zeros only for certain assessment types (e.g., formative but not summative), while others require students to redo work flagged for academic dishonesty without specifying what happens if they refuse.

It is important to recognize that issues with inconsistent documentation, education-specific language, and policies buried in board documents are not equity-neutral. Parents with more education, more time, and more familiarity with school systems are likely better positioned to find and interpret grading policies, to advocate for their children when policies are unevenly

⁶⁷ This variation in sourcing may also suggest that NZPs in some districts are not being adopted as top-down mandates from school board or district leadership — a claim that merits further qualitative investigation.

applied, and to seek outside information when grading policies are ambiguous. The ways in which families understand NZPs and how they interpret and use grades as grading policies change is an important area for future research.

How teachers use and interpret NZPs is a similarly important topic for future investigation. Proponents of grading reform argue that it isn't enough to simply adopt policies, there must also be a concerted effort to educate and support teachers in rethinking their grading practices (Feldman, 2018, 2023; Guskey et al., 2024; Hess, 2024). There is limited evidence of the extent to which NZPs are accompanied by such training and support. Nor is there much research on how teachers interpret ambiguous NZPs such as the district policy that states students must make up work that violates an academic integrity policy but does not explicitly state that zeros may not be given. Given that NZPs are meant to standardize grading practices, making them more consistent across classrooms, it is possible that teachers' (or schools') different interpretations of ambiguous policy undercut this aim. Further research into these questions of teachers' interpretation and implementation of policy is an additional area for future research.

When we link data on policies and schools, we find NZPs are more common in elementary schools than in secondary schools and tend to operate district-wide rather than varying across schools within a district. Perhaps most surprising, given that NZPs are often framed as equity-promoting reforms, we find that students in districts with higher proportions of low-income students are *less* likely to experience an NZP when accounting for year and district fixed effects.⁶⁸ Another way to put this is that it is possible that students with fewer academic

⁶⁸ We do see some evidence that some traditional markers of disadvantage – the percentage of English Language Learner (ELL) and underrepresented minority (URM) students – are associated with higher levels of elementary NZP adoption, but these findings are not significant for middle and high school NZPs, where grades are arguably more consequential. Additionally, these effects are not consistent across specifications.

challenges and more family financial support are experiencing grading policies that may inflate their grades.⁶⁹ Whether this pattern reflects differences in institutional capacity to adopt new policies, resistance from communities with different educational values, or something else is an important question for future research.

Other questions also remain unanswered. For example, do NZPs make recovery from failure easier and support students toward graduation? And if graduation rates do go up, are students who experience NZPs equally prepared for post-secondary opportunities as students who did not?⁷⁰ Are NZPs equally effective for all students, or do they benefit some student groups more than others? Do NZPs help students avoid losing ground and stay motivated (e.g., Feldman, 2018, 2023; Reeves, 2004), or do they reduce effort incentives in ways that lead to disengagement? And is that effect different for assignment NZPs versus course NZPs?⁷¹ Are grades more closely aligned to test scores when districts implement NZPs (The Equitable Grading Project, 2024), or do NZPs cause grade inflation?⁷² Harkening back to our discussion of

⁶⁹ That is, if NZPs raise GPAs by removing zeros at the course or assignment level, and NZPs are more prevalent in wealthier districts, then students in wealthier districts may be receiving a systematic GPA advantage in college admissions relative to equally performing students in lower-income districts. This compounds existing inequalities in college access rather than reducing them.

⁷⁰ A recent report from UC San Diego's workgroup on admissions found that the high school GPAs of admitted students required to take remedial math were substantively similar to the GPAs of students who did not (Senate-Administration Workgroup on Admissions, 2025). Furthermore, the workgroup found that 42% of students who failed to demonstrate middle school math skills had completed calculus or precalculus (presumably receiving passing grades in their previous math courses to get there). While this study does not look specifically at NZPs, and while it is a correlational study of a single university, it vividly illustrates a possible outcome if grades are artificially inflated by NZPs.

⁷¹ We note that assignment NZPs may change the stakes of *individual* decisions – whether to do tonight's homework, whether to turn in a project, whether to engage with a particular assignment. By raising the floor on any single assignment or removing the possibility for a zero penalty, these policies reduce the short-term cost of non-completion. The student who skips an assignment loses less ground in their overall grade than they would under a traditional zero. This may have positive effects: students who fall behind are not mathematically eliminated from passing, which may sustain engagement and motivation. However, it may also have negative consequences. The reduced penalty for non-completion may weaken the incentive to do the work at all: removing the possibility of a zero for a missing homework assignment or failed test reduces the incentive for a student to complete their work or study for an exam.

⁷² There is considerable evidence of grade inflation following close on the heels of the COVID-19 pandemic from studies that connect grades to student test performance, finding a weakening of the relationship between grades and test scores (Donahue, 2023; False Signals, 2023; Goldhaber & Goodman Young, 2024; Swiderski & Fuller, 2023);

the theoretical underpinnings of grading reform, the intertwining of student motivation and grades as signals makes it impossible to say a priori whether grading reforms are good or bad. Rather, we note that these policies are being adopted rapidly and we know little about how these policies affect the value of grades as a signal of student learning or alter the incentives grades create for students. As future researchers pursue these and other questions, they should bear in mind – and elaborate upon – the policy design nuances highlighted in this paper.

This study has several limitations that are also worth bearing in mind. Most significantly, because we rely on published, espoused *policy*, we likely understate the prevalence of alternative grading *practices*, which may vary considerably from course to course, teacher to teacher, or school to school. Moreover, we cannot capture how policies translate into actual classroom practice. While districts may adopt grading software that mechanically prevents zeros, in some cases implementation may depend on school- or teacher-level decisions such as determinations about what counts as a formative or summative assessment. Second, the data come from a single state that may differ systematically from others in ways that facilitate NZP adoption. For example, Washington implemented a statewide mandatory NZP at the end of 2019–20 in response to COVID-19, potentially influencing subsequent district-level policymaking. Our focus on Washington may also have led us to overlook policy dimensions that are relevant to other states or grading contexts. Finally, our results are descriptive, and we cannot draw causal conclusions about these relationships.

These limitations notwithstanding, this paper contributes groundwork for future studies by documenting the increasing prevalence of NZPs, their considerable variety, and the somewhat surprising association between NZP adoption and districts serving fewer low-income students.

there is also evidence of grade inflation that predates the 2020 pandemic (Gershenson, 2018; Sanchez, 2023; Sanchez & Moore, 2022).

The stakes of continued research on these policies are far from trivial. If future research suggests that NZPs keep struggling students engaged and on track, then their uneven adoption becomes a problem worth addressing. But if future research instead finds they erode the signaling and motivational functions of grades without improving outcomes, then their spread raises concern. Sorting out these competing possibilities should be a priority for researchers and policymakers alike.

References

- Aggarwal, C. C., & Zhai, C. (2012). A Survey of Text Classification Algorithms. In C. C. Aggarwal & C. Zhai (Eds.), *Mining Text Data* (pp. 163–222). Springer US. https://doi.org/10.1007/978-1-4614-3223-4_6
- Alex, P. (2022, July 5). Time to Pull the Plug on Traditional Grading? *Education Next*. <https://www.educationnext.org/time-to-pull-plug-on-traditional-grading-supporters-say-mastery-based-grading-could-promote-equity/>
- Babcock, P. (2010). Real costs of nominal grade inflation? New evidence from student course evaluations. *Economic Inquiry*, 48(4), 983–996. <https://doi.org/10.1111/j.1465-7295.2009.00245.x>
- Barshay, J. (2026, February 9). Easy A's, lower pay: Grade inflation's hidden damage. *The Hechinger Report*. <https://hechingerreport.org/proof-points-grade-inflation-lower-pay/>
- Bazzaz, D., & Long, K. (2020, April 20). Seattle will give high-school students A's or incompletes; colleges urged to adopt generous grading. *The Seattle Times*.
- Betts, J. R., & Grogger, J. (2003). The impact of grading standards on student achievement, educational attainment, and entry-level earnings. *Economics of Education Review*, 22(4), 343–352. [https://doi.org/10.1016/S0272-7757\(02\)00059-6](https://doi.org/10.1016/S0272-7757(02)00059-6)
- B-flation: How “good” grades can sideline parents*. (2023). Gallup.
- Black, P., & Wiliam, D. (1998). Assessment and Classroom Learning. *Assessment in Education: Principles, Policy & Practice*, 5(1), 7–74. <https://doi.org/10.1080/0969595980050102>
- Bowden, B. A., Rodriguez, V., & Weingarten, Z. (2023). *The unintended consequences of academic leniency* (EdWorkingPaper Nos. 23–836). edworkingpapers.com. <https://edworkingpapers.com/ai23-836>
- Bowers, A. J. (2011). What's in a grade? The multidimensional nature of what teacher-assigned grades assess in high school. *Educational Research and Evaluation*, 17(3), 141–159. <https://doi.org/10.1080/13803611.2011.597112>
- Brookhart, S. M. (1993). Teachers' grading practices: Meaning and values. *Journal of Educational Measurement*, 30(2), 123–142.
- Brookhart, S. M., & Guskey, T. R. (2020). Eight essential principles for improving grading. *Educational Leadership*, 78(1). <https://www.ascd.org/el/articles/eight-essential-principles-for-improving-grading>
- Brookhart, S. M., Guskey, T. R., Bowers, A. J., McMillan, J. H., Smith, J. K., Smith, L. F., Stevens, M. T., & Welsh, M. E. (2016). A century of grading research: Meaning and value in the most common educational measure. *Review of Educational Research*, 86(4), 803–848.

- Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C., ... Amodei, D. (2020). *Language Models are Few-Shot Learners* (arXiv:2005.14165). arXiv. <https://doi.org/10.48550/arXiv.2005.14165>
- Buck, D. (2022). *A “no zeroes” grading policy is the worst of all worlds*. Thomas B. Fordham Institute. <https://fordhaminstitute.org/national/commentary/no-zeroes-grading-policy-worst-all-worlds>
- Butler, M. (2025). Alternative grading systems and student outcomes: A comparative analysis of motivation, enjoyment, engagement, stress, and perceptions of final grades. *Assessment & Evaluation in Higher Education*, 50(5), 807–817. <https://doi.org/10.1080/02602938.2025.2475068>
- Butler, R., & Nisan, M. (1986). Effects of no feedback, task-related comments, and grades on intrinsic motivation and performance. *Journal of Educational Psychology*, 78(3), 210–216. <https://doi.org/10.1037/0022-0663.78.3.210>
- Calarco, J. M. (2020). Avoiding Us versus Them: How Schools’ Dependence on Privileged “Helicopter” Parents Influences Enforcement of Rules. *American Sociological Review*, 85(2), 223–246. <https://doi.org/10.1177/0003122420905793>
- Canady, R. L., & Hotchkiss, P. R. (1989). It’s a good score! Just a bad grade. *Phi Delta Kappan*, 71(1), 68–71.
- Carey, T., & Carifio, J. (2012). The minimum grading controversy: Results of a quantitative study of seven years of grading data from an urban high school. *Educational Researcher*, 41(6), 201–208. <https://doi.org/10.3102/0013189X12453309>
- Carifio, J., & Carey, T. (2015). Further findings on the positive effects of minimum grading. *Journal of Education and Social Policy*, 2(4), 130–136.
- Chamberlain, G. (1980). Analysis of covariance with qualitative data. *The Review of Economic Studies*, 47(1), 225–238.
- Denning, J. T., Nesbit, R., Pope, N., & Warnick, M. (2025). *Easy A’s, Less Pay: The Long-Term Effects of Grade Inflation* (Working Paper No. 34952). National Bureau of Economic Research. <https://www.nber.org/papers/w34952>
- Donahue, T. (2023, October 23). Opinion | If everyone gets an A, no one gets an A. *The New York Times*. <https://www.nytimes.com/2023/10/23/opinion/grade-inflation-high-school.html>
- Esquivel, P. (2021, November 8). *Faced with soaring Ds and Fs, schools are ditching the old way of grading*. Los Angeles Times. <https://www.latimes.com/california/story/2021-11-08/as-ds-and-fs-soar-schools-ditch-inequitable-grade-systems>

- Fallan, L., & Opstad, L. (2012). Attitudes towards study effort response to higher grading standards: Do gender and personality distinctions matter? *Journal of Education and Learning*, 1(2), p179. <https://doi.org/10.5539/jel.v1n2p179>
- False signals: How pandemic-era grades mislead families and threaten student learning.* (2023). TNTP.
- Feldman, J. (2018). *Grading for equity: What it is, why it matters, and how it can transform schools and classrooms* (1st edition). Corwin.
- Feldman, J. (2023). *Grading for Equity: What It Is, Why It Matters, and How It Can Transform Schools and Classrooms*. Corwin Press.
<http://ebookcentral.proquest.com/lib/washington/detail.action?docID=30755814>
- Figlio, D. N., & Lucas, M. E. (2004). Do high grading standards affect student performance? *Journal of Public Economics*, 88(9–10), 1815–1834. [https://doi.org/10.1016/S0047-2727\(03\)00039-2](https://doi.org/10.1016/S0047-2727(03)00039-2)
- Flannery, M. E. (2026, March 20). Yes, Grade Inflation Is Real. But Is It a Real Problem? | NEA. *NEA Today*. <https://www.nea.org/nea-today/all-news-articles/grade-inflation-real-it-real-problem>
- Galla, B. M., Shulman, E. P., Plummer, B. D., Gardner, M., Hutt, S. J., Goyer, J. P., D’Mello, S. K., Finn, A. S., & Duckworth, A. L. (2019). Why High School Grades Are Better Predictors of On-Time College Graduation Than Are Admissions Test Scores: The Roles of Self-Regulation and Cognitive Ability. *American Educational Research Journal*, 56(6), 2077–2115. <https://doi.org/10.3102/0002831219843292>
- Geiser, S., & Santelices, M. V. (2007). *Validity of high-school grades in predicting student success beyond the freshman year: High-school record vs. Standardized tests as indicators of four-year college outcomes* (Research and Occasional Papers Series). University of California Berkeley.
- Gershenson, S. (2018). *Grade Inflation in High Schools (2005–2016)*. Thomas B. Fordham Institute.
- Gershenson, S., Holt, S. B., & Tyner, A. (2024). Making the grade: The effect of teacher grading standards on student outcomes. *Contemporary Economic Policy*, 42(2), 305–318. <https://doi.org/10.1111/coep.12637>
- Goldhaber, D., & Goodman Young, M. (2024). Course grades as a signal of student achievement: Evidence of grade inflation before and after COVID-19. *Journal of Policy Analysis and Management*, n/a(n/a). <https://doi.org/10.1002/pam.22618>
- Gonzales, N. (2024, February 22). Schools explore replacing Fs to address demographic grading gaps. *Palo Alto Online*. <https://www.paloaltoonline.com/express-newsletter-content/2024/02/21/schools-explore-replacing-fs-to-address-demographic-grading-gaps/>

- Griffith, D., & Tyner, A. (2025). "Equitable" Grading Through the Eyes of Teachers. Thomas B. Fordham Institute. <https://fordhaminstitute.org/national/research/equitable-grading-through-eyes-teachers>.
- Guskey, T. R. (1994). Making the grade: What benefits students? *Educational Leadership*, 52(2), 14–20.
- Guskey, T. R. (2020). Breaking up the Grade. *Educational Leadership*, 78(1), 40–46.
- Guskey, T. R. (2024). Addressing inconsistencies in grading practices. *Phi Delta Kappan*, 105(8), 52–57. <https://doi.org/10.1177/00317217241251883>
- Guskey, T. R., & Brookhart, S. M. (2019). *What We Know about Grading: What Works, What Doesn't, and What's Next*. Association for Supervision & Curriculum Development. <http://ebookcentral.proquest.com/lib/washington/detail.action?docID=5704240>
- Guskey, T. R., Fisher, D., & Frey, N. (2024). The Unwinnable Battle Over Minimum Grades. *Educational Leadership*, 82(2), 68–72.
- Guskey, T. R., & Jung, L. A. (2013). *Answers to Essential Questions About Standards, Assessments, Grading, and Reporting*. Corwin.
- Guskey, T. R., & Link, L. J. (2019). Exploring the factors teachers consider in determining students' grades. *Assessment in Education: Principles, Policy & Practice*, 26(3), 303–320. <https://doi.org/10.1080/0969594X.2018.1555515>
- Hattie, J., & Timperley, H. (2007). The Power of Feedback. *Review of Educational Research*, 77(1), 81–112. <https://doi.org/10.3102/003465430298487>
- Hess, F. M. (2024, February 27). Traditional Grading May Not Be as Straightforward as It Seems. *Medium*. <https://rickhess99.medium.com/traditional-grading-may-not-be-as-straightforward-as-it-seems-4c5f1e6b5863>
- Horowitch, R. (2025, November 19). 'A Recipe for Idiocracy.' *The Atlantic*. <https://www.theatlantic.com/ideas/2025/11/math-decline-ucsd/684973/>
- Huey, M. E., Silvey, P. R., Vaughan, A. G., & Fisher, A. L. (2022). Assessing the impact of standards-based grading policy changes on student performance and practice work completion in secondary mathematics. *Studies in Educational Evaluation*, 75, 101211. <https://doi.org/10.1016/j.stueduc.2022.101211>
- Hutt, E. (2025). Grading Policies. In D. Harris (Ed.), *Live Handbook of Education Policy Research*. Association for Education Finance and Policy. <https://livehandbook.org/k-12-education/standards-and-accountability/grading-policies/>
- Ji, Z., Lee, N., Frieske, R., Yu, T., Su, D., Xu, Y., Ishii, E., Bang, Y. J., Madotto, A., & Fung, P. (2023). Survey of Hallucination in Natural Language Generation. *ACM Computing Surveys*, 55(12), 248:1–248:38. <https://doi.org/10.1145/3571730>

- Johnson, D. R. (2025). Measuring Grading Standards at High Schools: A Methodological Contribution, an Example, and Some Policy Implications. *Educational Evaluation and Policy Analysis*, 01623737241311537. <https://doi.org/10.3102/01623737241311537>
- Jones, C., & Fensterwald, J. (2021, December 6). *Why Some California School Districts Are Changing How Students Earn Grades* | KQED. <https://www.kqed.org/news/11898137/why-some-california-school-districts-are-changing-how-students-earn-grades>
- Kirschenbaum, H., Napier, R., & Simon, S. B. (2021). *Wad-Ja-Get? The Grading Game in American Education* (50th Anniversary). Maize Books. <https://doi.org/10.3998/mpub.11900733>
- Kramer, S. L., Posner, M. A., Browman, A. S., Lawrence, N. R., Roem, J., & Krier, K. (2023). The Impacts of a Standards-Based Grading System Emphasizing Formative Assessment, Feedback, and Re-Assessment: A Mixed Methods, Cluster Randomized Control Trial in Ninth Grade Mathematics Classrooms. *Journal of Research on Educational Effectiveness*, 18(1), 56–87. <https://doi.org/10.1080/19345747.2023.2287594>
- Kunnath, J. (2022, November 18). The Low-hanging Fruit of Grading Reform: Eliminating the Zero. *The Core Collaborative*. <https://thecorecollaborative.com/the-low-hanging-fruit-of-grading-reform-eliminating-the-zero/>
- Lakshminarayanan, B., Pritzel, A., & Blundell, C. (2017). Simple and scalable predictive uncertainty estimation using deep ensembles. *Proceedings of the 31st International Conference on Neural Information Processing Systems, NIPS'17*, 6405–6416. <https://dl.acm.org/doi/10.5555/3295222.3295387>
- Lewis, D. (2022). Impacts of Standards-Based Grading on Students' Mindset and Test Anxiety. *Journal of the Scholarship of Teaching and Learning*, 22(2). <https://doi.org/10.14434/josotl.v22i2.31308>
- Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., Küttler, H., Lewis, M., Yih, W., Rocktäschel, T., Riedel, S., & Kiela, D. (2021). *Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks* (arXiv:2005.11401). arXiv. <https://doi.org/10.48550/arXiv.2005.11401>
- Link, L. J., & Guskey, T. R. (2019). How traditional grading contribute to student inequities and how to fix it. *Curriculum in Context*, 45(1).
- Link, L. J., & Guskey, T. R. (2022). Is standards-based grading effective? *Theory Into Practice*, 61(4), 406–417. <https://doi.org/10.1080/00405841.2022.2107338>
- Manning, C. D., Raghavan, P., & Schütze, H. (2008). *Introduction to Information Retrieval*. Cambridge University Press.

- Morris, S. R., & Johnson, A. H. (2024). Could minimum grading enhance high school graduation rates and cost-effectiveness across Arkansas? *Education Policy Analysis Archives*, 32. <https://doi.org/10.14507/epaa.32.8496>
- Morris, S. R., Maranto, R. A., & McKenzie, S. C. (2023). Examining Teacher Support for Grading Equity. *Journal of Research in Education*, 32(2).
- Morris, S. R., McKenzie, S. C., Wai, J., & Maranto, R. A. (2025). An Investigation of Ninth Grade Students' Perceptions of Equitable Grading Practices. *Journal of Research Initiatives*, 8(5).
- Mozenter, Z. D. (2019). *Essays on the effects of teacher grading standards and other teaching practices* [Ph.D., The University of North Carolina at Chapel Hill]. <https://www.proquest.com/docview/2240074332/abstract/246BCFDF8C5D4CD6PQ/1>
- Nierenberg, A. (2023, December 5). Nearly everyone gets A's at Yale. Does that cheapen the grade? *The New York Times*. <https://www.nytimes.com/2023/12/05/nyregion/yale-grade-inflation.html>
- O'Connor, K., & Wormeli, R. (2011). Reporting Student Learning. *Educational Leadership*, 69(3), 40–44. (66901499).
- OSPI. (2020). *Student Learning & Grading Guidance: Frequently Asked Questions*. Office of Superintendent of Public Instruction.
- Piper, K. (2025, August 19). *When grades stop meaning anything*. <https://www.theargumentmag.com/p/when-grades-stop-meaning-anything>
- Pollio, M., & Hochbein, C. (2015). The Association between Standards-Based Grading and Standardized Test Scores as an Element of a High School Reform Model. *Teachers College Record*, 117(11), 1–28. <https://doi.org/10.1177/016146811511701106>
- Reeves, D. B. (2004). The case against the zero. *Phi Delta Kappan*, 86(4), 324–325. <https://doi.org/10.1177/003172170408600418>
- Reeves, D. B., Jung, L. A., & O'Connor, K. (2017). What's Worth Fighting Against in Grading? *Educational Leadership*, 74(8), 42–45. (122878943).
- Rury, D., & Kalil, A. (2026). *Interpreting Performance: Evidence on Signal Weighting in Human Capital Investment* (SSRN Scholarly Paper No. 6174458). Social Science Research Network. <https://doi.org/10.2139/ssrn.6174458>
- Sanchez, E. I. (2023). *Evidence of grade inflation since 2010 in high school English, mathematics, social studies, and science courses*. ACT.
- Sanchez, E. I., & Moore, R. (2022). *Grade inflation continues to grow in the past decade*. ACT.

- Schneider, J., & Hutt, E. (2014). 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>
- Senate-Administration Workgroup on Admissions. (2025). *Final Report*. University of California San Diego.
- Shute, V. J. (2008). Focus on Formative Feedback. *Review of Educational Research*, 78(1), 153–189. <https://doi.org/10.3102/0034654307313795>
- Swiderski, T., & Fuller, S. C. (2023). Student GPA and test score gaps are growing—And could be slowing pandemic recovery. *Brookings*. <https://www.brookings.edu/articles/student-gpa-and-test-score-gaps-are-growing-and-could-be-slowing-pandemic-recovery/>
- Tan, B. J. (2023). The consequences of letter grades for labor market outcomes and student behavior. *Journal of Labor Economics*, 41(3), 565–588. <https://doi.org/10.1086/719994>
- The Equitable Grading Project. (2024). *Can we trust the transcript? Recognizing student potential through more accurate grading* [White Paper]. The Equitable Grading Project.
- Townsend, M. B. (2022). Why I Stopped Giving Zeros. *Edutopia*. <https://www.edutopia.org/article/why-i-stopped-giving-zeros/>
- Townsley, M., & Varga, M. (2018). Getting High School Students Ready for College: A Quantitative Study of Standards- Based Grading Practices. *Journal of Research in Education*, 28(1), 92–112.
- Tyner, A. B. (2024). Unintended Effects of “Equitable” Grading: Reforms aimed at levelling playing fields may lead instead to lowered expectations for student performance. *School Administrator*, 81(11), 38–41.
- Zhao, P., Zhang, H., Yu, Q., Wang, Z., Geng, Y., Fu, F., Yang, L., Zhang, W., Jiang, J., & Cui, B. (2024). *Retrieval-Augmented Generation for AI-Generated Content: A Survey* (arXiv:2402.19473). arXiv. <https://doi.org/10.48550/arXiv.2402.19473>.

Tables and Figures

Table 1. Descriptive Statistics of Districts With Elementary, Middle, And High School NZPs Weighted by District Student Population

	All Schools		Elementary Schools		Middle Schools		High Schools	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	NZP	No NZP	NZP	No NZP	NZP	No NZP	NZP	No NZP
District Characteristics								
% City	39.4	41.3	36.0***	42.8	30.8***	42.1	28.7***	42.3
% Rural	4.0***	12.4	3.1***	12.1	6.2**	9.7	1.8***	9.7
% Town	5.2***	17.6	4.7***	16.9	3.1***	14.5	0.2***	14.6
% Suburb	51.4***	28.7	56.2***	28.2	60.0***	33.7	69.3***	33.4
# Students (thousand)	20.0***	12.6	19.1***	13.6	17.3***	15.0	26.6***	14.1
	(10.9)	(12.5)	(8.0)	(13.7)	(8.3)	(12.9)	(14.4)	(11.5)
Student Characteristics								
% Proficient on State Test	52.0***	46.1	52.5***	46.4	45.9***	48.7	49.4	48.2
	(15.3)	(12.8)	(15.7)	(12.8)	(14.2)	(14.0)	(15.0)	(14.0)
% Special education	14.6***	15.1	14.3***	15.2	14.8***	14.9	14.8	14.9
	(2.7)	(2.3)	(2.7)	(2.3)	(2.1)	(2.5)	(2.3)	(2.5)
% Low income	42.1***	51.0	41.8***	50.5	48.9	47.7	45.0***	48.2
	(20.7)	(17.8)	(21.3)	(17.8)	(18.8)	(19.4)	(19.3)	(19.2)
% English Language Learner	12.1	12.7	12.0*	12.7	15.1***	12.2	14.9***	12.3
	(8.2)	(10.7)	(8.2)	(10.5)	(9.7)	(9.9)	(9.0)	(9.9)
% Underrepresented minority	51.1***	46.3	51.6***	46.4	56.6***	46.9	57.1***	47.3
	(17.8)	(20.6)	(17.9)	(20.3)	(18.8)	(19.6)	(15.6)	(19.8)
N Districts	68	227	53	242	30	253	18	234

Note: 1) All statistics are weighted by district student population 2) Stars indicate significant differences between the means in Columns 1 and 2, the means in Columns 3 and 4, the means in Columns 5 and 6, and the means in Column 7 and 8: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. 3) N Districts represents the number of schools with NZPs and without NZPs in the final year of our sample. Because districts adopt NZPs in different years, the number of districts with and without NZPs varies across our panel.

Table 2. Descriptive Statistics of Schools With and Without NZPs in Districts With Cross-School NZP Variation, Weighted by School Size

	All Schools		Elementary Schools		Middle Schools		High School	
	(1) NZP	(2) No NZP	(3) NZP	(4) No NZP	(5) NZP	(6) No NZP	(7) NZP	(8) No NZP
District Characteristics								
# Students	743.7 (503.9)	741.4 (432.6)	459.1 (118.5)	475.8 (88.9)	655.2** (149.0)	743.3 (117.6)	1873.5* (56.8)	1604.7 (249.4)
Student Characteristics								
% Proficient on State Test	49.5 (17.2)	47.9 (17.5)	53.4 (18.2)	50.5 (17.9)	37.7 (11.0)	39.1 (11.8)	64.9 (1.5)	53.3 (19.3)
% Special education	15.3 (2.7)	15.2 (4.5)	16.2 (2.3)	16.4 (4.9)	15.7 (2.4)	14.9 (2.9)	11.7 (0.7)	11.6 (2.4)
% Low income	45.4 (23.4)	46.6 (21.6)	44.5 (25.0)	45.2 (20.1)	56.4 (18.5)	62.0 (19.5)	22.4 (1.3)	27.2 (8.4)
% English Language Learner	9.0 (9.0)	10.5 (9.0)	7.5 (7.6)	9.7 (7.7)	13.7 (10.3)	17.5 (9.3)	2.7 (0.4)	2.3 (1.5)
% Underrepresented minority	42.3 (19.1)	45.0 (20.8)	33.4 (13.9)	39.4 (15.6)	57.0 (20.4)	64.6 (23.8)	36.1 (2.5)	33.1 (4.1)
N Schools	11	23	6	18	4	4	1	1

Note: 1) All statistics are weighted by school student population 2) Stars indicate significant differences between the means in Columns 1 and 2, the means in Columns 3 and 4, the means in Columns 5 and 6, and the means in Column 7 and 8: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ 3) N Schools represents the number of schools with NZPs and without NZPs in the final year of our sample. Because schools adopt NZPs in different years, the number of schools with and without NZPs varies across our panel.

Table 3. OLS Coefficients of District Covariates on the Likelihood of a District Adopting a No Zeros Policy, Weighted by District Size

	All NZPs		Elementary NZPs		Middle School NZPs		High School NZPs	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(# Students (1000s))	0.0084 (0.0070)	-0.0837 (0.0545)	0.0037 (0.0087)	0.0155 (0.0168)	0.0008 (0.0040)	0.0010 (0.0194)	0.0101*** (0.0028)	-0.0676 (0.0580)
Rural (ref. City)	0.1025 (0.1324)		0.0205 (0.1491)		0.0491 (0.0907)		0.1642*** (0.0542)	
Suburb (ref. City)	0.1727 (0.1194)		0.1916 (0.1233)		0.1267 (0.1162)		0.2030** (0.0957)	
Town (ref. City)	0.0036 (0.1305)		-0.0270 (0.1372)		-0.0370 (0.0900)		0.1064** (0.0533)	
% SpEd	0.0105 (0.0148)	-0.0113 (0.0084)	-0.0073 (0.0171)	-0.0095** (0.0043)	0.0029 (0.0124)	-0.0205 (0.0168)	0.0049 (0.0127)	-0.0312 (0.0190)
% Low income	-0.0052 (0.0041)	-0.0093*** (0.0035)	-0.0029 (0.0044)	-0.0059* (0.0035)	-0.0022 (0.0026)	-0.0052* (0.0027)	-0.0008 (0.0022)	-0.0065* (0.0035)
% ELL	-0.0105 (0.0081)	0.0151 (0.0129)	-0.0137* (0.0080)	0.0202* (0.0105)	-0.0049 (0.0062)	-0.0050 (0.0080)	-0.0020 (0.0057)	-0.0034 (0.0084)
% URM	0.0090* (0.0050)	-0.0095 (0.0058)	0.0096* (0.0052)	-0.0046 (0.0031)	0.0049 (0.0032)	0.0100 (0.0079)	0.0017 (0.0018)	0.0013 (0.0100)
% Proficient on State Test	-0.0004 (0.0033)	0.0036 (0.0056)	0.0004 (0.0030)	-0.0043** (0.0018)	-0.0044 (0.0027)	0.0004 (0.0027)	-0.0030 (0.0019)	0.0080 (0.0063)
Fixed Effects		Dist & Yr		Dist & Yr		Dist & Yr		Dist & Yr
Constant	-0.1577 (0.3865)	2.3588*** (0.8983)	0.1332 (0.4255)	0.6822* (0.3952)	0.1707 (0.2585)	0.2328 (0.4374)	-0.0988 (0.1785)	1.5110 (0.9949)
N Districts	295	295	295	295	283	283	252	252
R-squared	0.1838	0.8531	0.1673	0.9138	0.0797	0.7641	0.1806	0.6920

Note: 1) Standard errors in parentheses are clustered by district. 2) p-values from two-sided t-tests: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. 3) SpEd = Special Education, ELL = English Language Learner, URM = Underrepresented Minority. 4) Specifications with all NZPs also include dummy variables indicating whether a district has a high school and whether a district has a middle school. These are not significant in any specification.

Figure 1. Increase in NZPs in Washington State Over Time

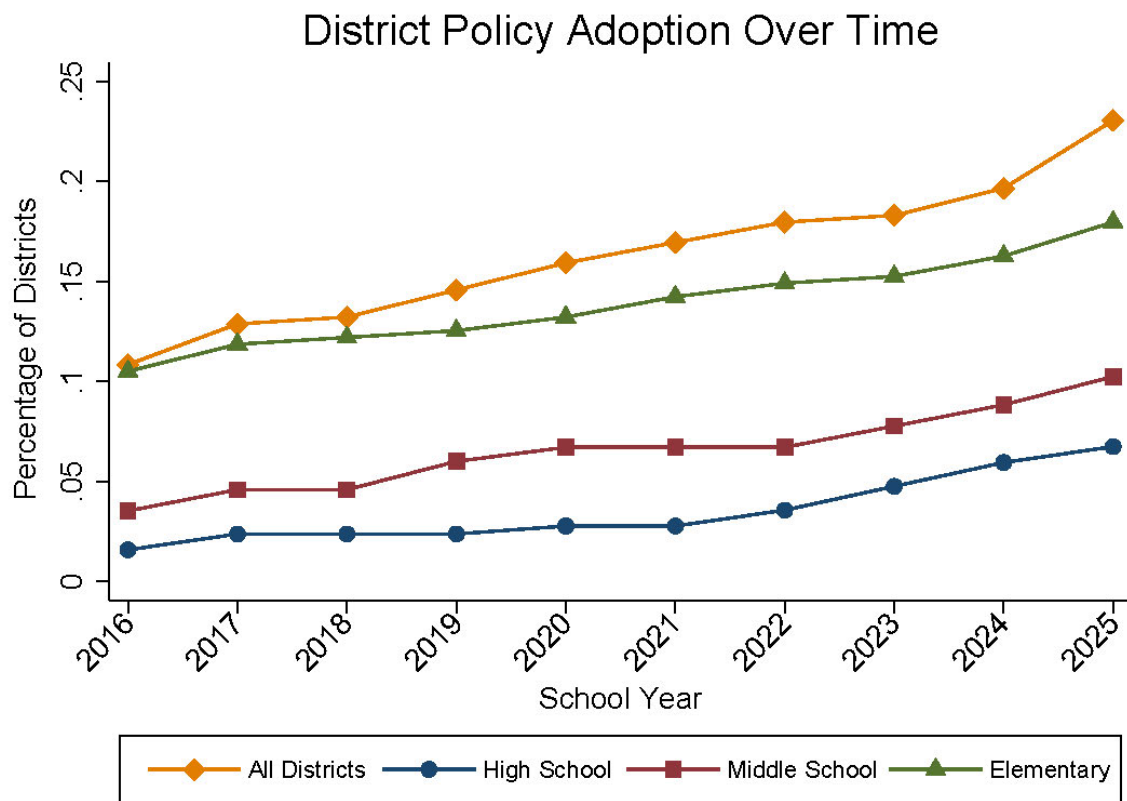
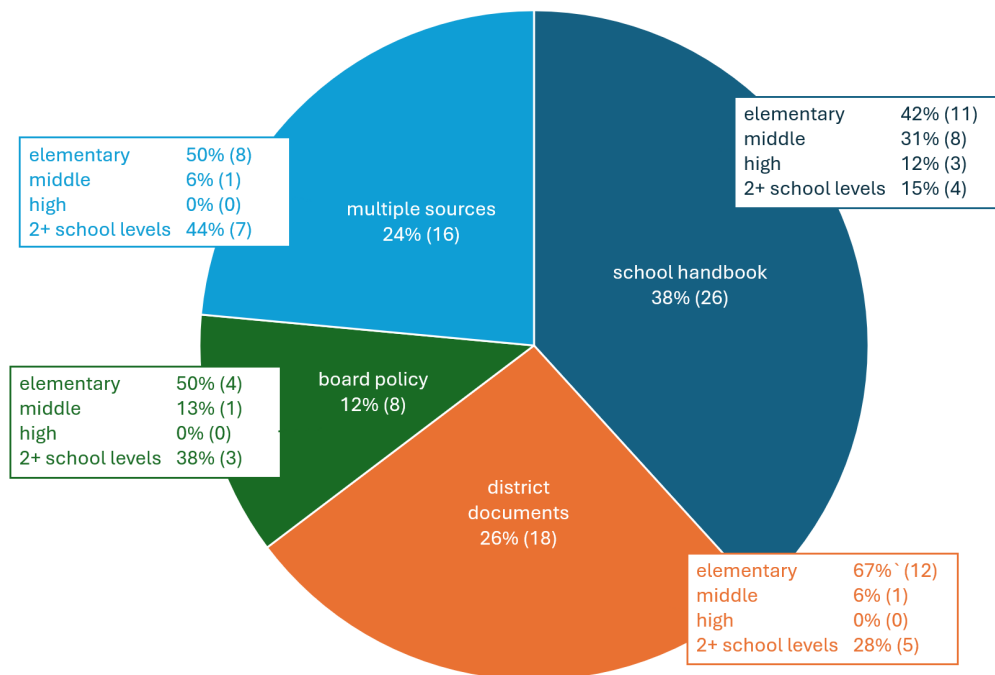
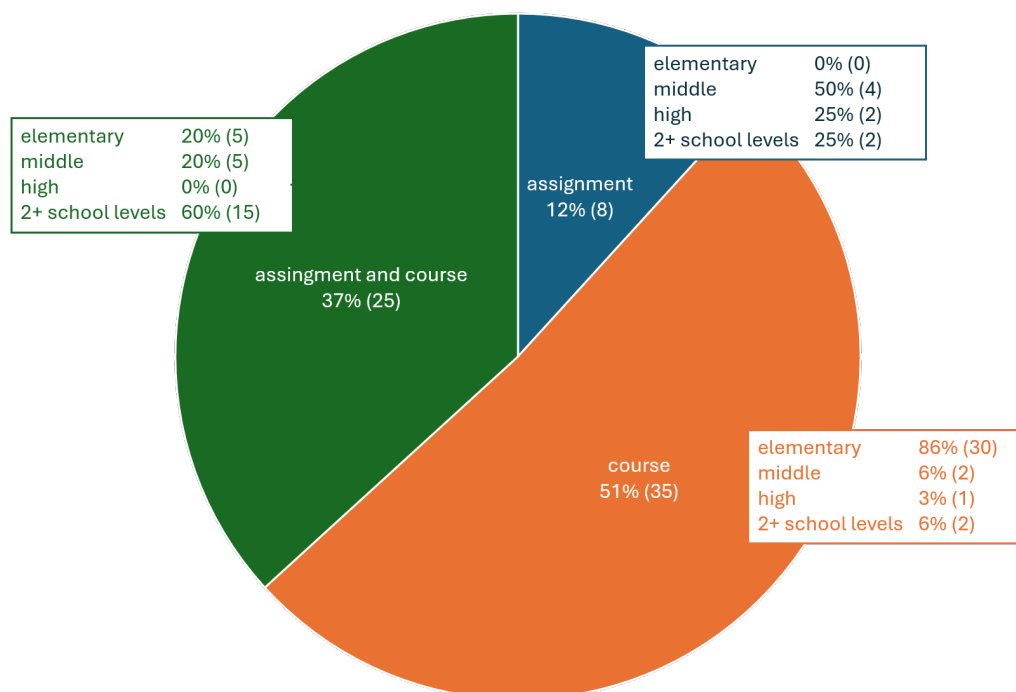


Figure 2. Distribution of NZP typology across grade levels

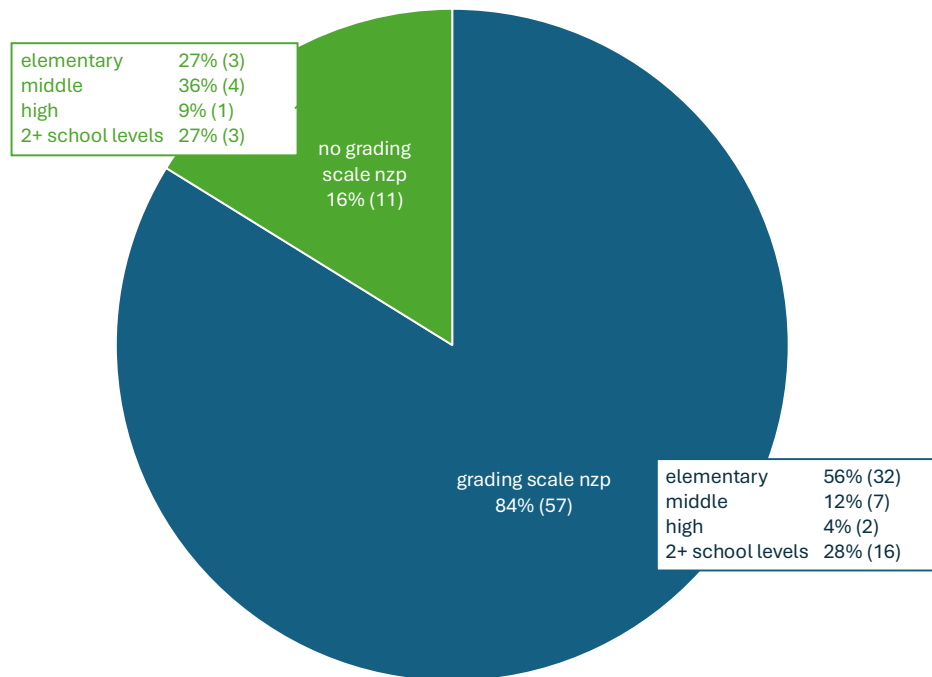
Panel A. NZP Origin



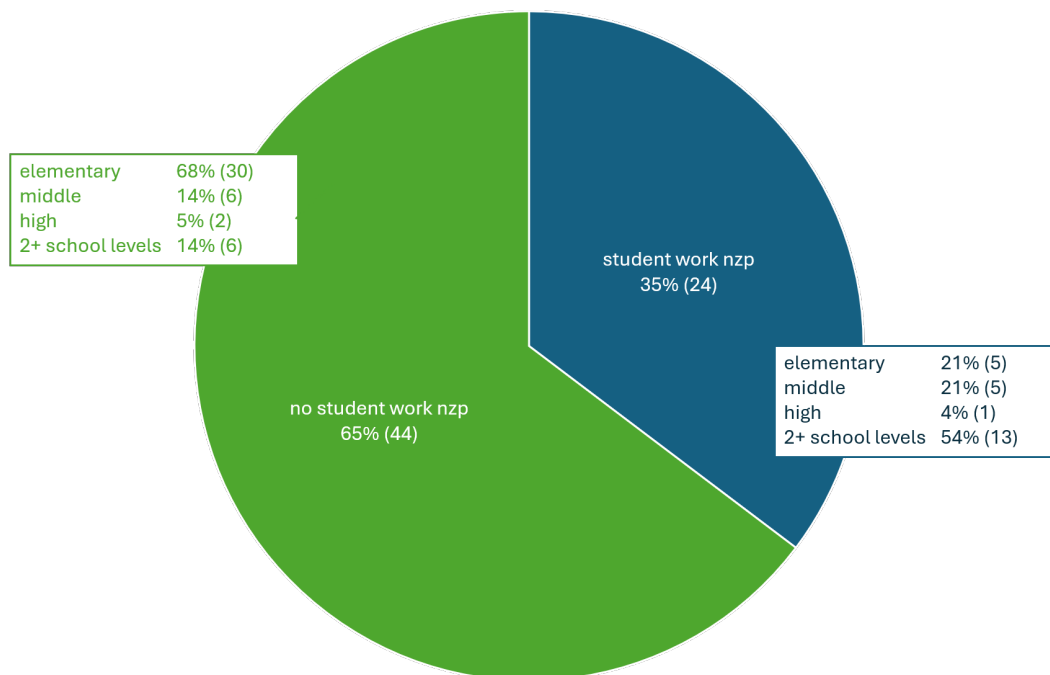
Panel B. NZP Application



Panel C. Grading Scale NZPs



Panel D. Student Work NZPs



APPENDIX A. Handbook Data Collection

The first pass of handbook searches entailed searching district-level websites for school-level handbooks. When we found that that did not give us the coverage we needed, we conducted a second search of school-level handbooks. We did this at the beginning of the 2025-26 school year and found that some school websites had a placeholder that the school handbook was “coming soon.” Thus, we conducted a third search in the middle of the 2025-26 school year to identify any handbooks that we had missed. On this third pass, we also queried a search engine with the school name, district name, and “handbook” to find any files that might be buried in district websites.

Because there is no universal standard for handbooks across districts, and because we encounter districts with one or two schools as well as districts with over twenty schools, we developed the following logic in determining if a policy is at the school or district level

1. If district documents (board policies and district handbooks) include an NZP, the district has a **district-level NZP**.
2. If district documents do not include an NZP, but school handbooks reference a district-level NZP (e.g., “district leadership requires...” or “as per district rules...”), the district has a **district-level NZP**.
3. If district documents do not include an NZP, but *all* school handbooks reference some level of NZP, the district has a **district-level NZP**.
4. If district documents do not include an NZP, but *some* school handbooks reference an NZP and the remaining handbooks make no reference to an NZP (e.g., the handbooks do not reference grading at all), we make an arbitrary distinction that the district has a **district-level NZP** if more than half of the

schools at a given level mention the same NZP and the district has **school-level NZPs** if half or fewer of the schools at that level mention an NZP.

5. If district documents do not include an NZP, but some school handbooks reference an NZP and any the remaining handbooks have a contradictory policy (e.g., one school has a non-zero grading floor policy, and another school explicitly uses zeros) the district has **school-level NZPs**.

APPENDIX B. Technical Appendix and AI Queries

Our document collection efforts result in a corpus of 7,401 files in varying file formats. To analyze this volume of documents, we use large language models (LLMs) to query each document for relevant grading information. We then develop a systematic document processing pipeline using open-source extraction software and a retrieval-augmented generation system (RAG) to allow us to query document collections. In this appendix we discuss the rationale for using LLMs and RAG and detail the pipeline for querying documents.

We elect to use LLMs for this purpose rather than other forms of machine learning and document classification because traditional document classification methods – such as keyword matching, term frequency-inverse document frequency (TF-IDF), and bag-of-words – tend to rely on techniques like word counts and matches to keywords to generate classifications. These techniques typically require extensive example datasets to train on that have already been manually labeled and classified (Aggarwal & Zhai, 2012; Manning et al., 2008). Since we did not have a preexisting labeled dataset to produce classifiers reliably and because language varies significantly between documents and document types, making keyword searches unreliable, we turned to performing semantic searches with LLMs to extract answers from our corpus.

To understand the power of semantic searches using LLMs, it's helpful to compare it to how humans understand language in social contexts. A phrase like “I’m freezing” could be interpreted by a reader in a variety of ways:

- “I’m cold” (a typical complaint)
- “I am unable to move” (physical or metaphorical paralysis)
- “I am below 32 degrees Fahrenheit” (a literal physical state)

Traditional keyword-based classification methods treat "freezing" as a token to be matched, recognizing its presence in a document, but not its meaning within context. LLMs, by contrast, interpret the phrase contextually, processing the surrounding sentence or paragraph to determine which meaning applies. This mirrors how humans naturally resolve ambiguity through understanding conceptual relationships rather than surface-level keyword matching. In this way, LLMs improve upon traditional methods by capturing meaning above and beyond syntax, enabling more accurate and contextually grounded document classification and retrieval (Zhao et al., 2024).

Since LLMs are trained to predict patterns derived from a massive variety of datasets, they are also capable of "zero-shot" classification, removing the need to have thousands of labeled examples to train classification. This ability, however, raises concerns about the risk of hallucination – a phenomenon in which LLMs generate confident but factually incorrect outputs (Ji et al., 2023). A common condition contributing to hallucination is when an LLM is treated more as a database than a pattern predictor. If one searches for information in a database, it is either there or it returns a null result. Since LLMs are pattern predictors, they do not know what they don't know and can produce confident answers even when the relevant information is absent from their training data.

To mitigate the risk of hallucination, we implemented a RAG system. Rather than relying solely on the LLM's training, RAG first retrieves relevant document fragments from a curated database of district policies and handbooks in response to each query and provides them as context within the query, grounding the model's responses in verifiable source material (P. Lewis et al., 2021). Rather than relying only on the training an LLM has received, RAG allows an LLM to use its reasoning to analyze the submitted text.

An example of misusing an LLM as a knowledge base would be to ask an LLM directly: does this district have an NZP? In the case of NZP, the LLM likely has no training data specific to the district in question and may hallucinate a plausible but fabricated answer. With RAG, a typical prompt takes more of the form: "Analyze the following texts and determine, based solely on this material, whether this district meets NZP criteria: {texts}." Grounded in the supplied source material, the LLM can then render a judgment that is traceable to specific documents rather than generated from training alone.

The RAG approach offers several key advantages: it grounds the model's outputs in verifiable source material, reducing hallucinations significantly; it enables the system to access updated and domain-specific information without requiring costly model retraining when documents are added or updated; and it provides auditability by allowing users to trace answers back to their source documents.

Our RAG pipeline broadly consists of the following: a webapp to organize documents by district, school, date, and document tag, processes to extract various filetypes to a standardized format, dissection of these extractions into meaningful fragments (of a paragraph or less), conversion of each fragment into embeddings using the VoyageAI "voyage-law-2" encoding model, and storage of these results in a vector store (ChromaDB). For document retrieval, we encode a query with the VoyageAI encoding model, select the most similar text fragments from the relevant files in ChromaDB (selecting the top 10 results), rerank these results with the VoyageAI rerank-lite model (keeping the 5 most relevant paragraphs), and then send these results as context supporting the query to Anthropic's Claude Sonnet 4.6 and Claude Opus 4.7 for judgments.

While LLMs are capable of zero-shot classification within this RAG architecture, we observed improved response quality when queries were coupled with detailed system prompts incorporating several general labeled examples – an approach known as few-shot prompting (Brown et al., 2020). These system prompts provide additional context to guide the model's decision making for each query type. We also instructed the LLM to return confidence ratings and explanatory notes for each judgment to support downstream validation.

As we describe in the body of the paper, validation of the results was performed at multiple levels to ensure accuracy. Verdicts with low confidence ratings that indicated the presence of an NZP were flagged for human review. Since we ran the same queries through multiple models, when the models disagreed, we also flagged these responses for human review. This multi-model approach functions as a form of ensemble validation, where disagreement between models serves as a signal of ambiguity warranting closer human scrutiny (Lakshminarayanan et al., 2017). Results were further validated by carefully reading the returned notes and comments for all queries that produce a positive indication for an NZP and by similarly examining a random selection of responses for queries that yield an uncertain or negative response for an NZP.

APPENDIX C. Additional Tables and Figures

Table C1. Descriptive Statistics of Districts with Elementary, Middle, and High School NZPs

	All Schools		Elementary Schools		Middle Schools		High Schools	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	NZP	No NZP	NZP	No NZP	NZP	No NZP	NZP	No NZP
District Characteristics								
% City	18.7***	8.9	19.3***	9.0	16.2**	10.6	13.4	12.3
% Rural	42.0***	59.9	43.7***	59.1	43.1***	56.7	34.1***	52.7
% Town	10.0***	17.5	8.4***	17.6	7.2***	16.9	1.2***	18.5
% Suburb	29.3***	13.7	28.6***	14.2	33.5***	15.8	51.2***	16.6
# Students (thousand)	8.3***	2.8	8.4***	3.0	7.6***	3.6	12.0***	4.1
	(9.9)	(5.3)	(9.5)	(5.6)	(8.7)	(6.4)	(13.5)	(6.4)
Student Characteristics								
% Proficient on State Test	46.3***	43.3	46.3***	43.4	43.7	43.6	42.8	43.4
	(14.7)	(13.1)	(15.0)	(13.1)	(13.1)	(13.6)	(13.9)	(13.6)
% Special education	14.9**	15.4	14.8***	15.4	15.0	15.5	15.1	15.6
	(3.1)	(4.5)	(3.2)	(4.4)	(2.3)	(4.2)	(2.3)	(3.6)
% Low income	51.4	52.9	51.8	52.8	55.8	53.2	54.4	53.9
	(21.3)	(21.9)	(21.4)	(21.9)	(19.3)	(21.1)	(19.0)	(19.6)
% English Language Learner	11.0***	8.8	11.0***	8.9	13.0***	8.9	9.5	9.6
	(11.9)	(12.5)	(12.0)	(12.5)	(12.6)	(11.9)	(10.0)	(11.8)
% Underrepresented minority	43.6***	35.5	43.8***	35.7	47.3**	36.2	42.2	37.9
	(23.5)	(24.2)	(23.3)	(24.2)	(23.6)	(23.8)	(20.8)	(23.4)
N Districts	68	227	53	242	29	254	17	235

Note: 1) Stars indicate significant differences between the means in Columns 1 and 2, the means in Columns 3 and 4, the means in Columns 5 and 6, and the means in Column 7 and 8: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. 2) N Districts represents the number of schools with NZPs and without NZPs in the final year of our sample. Because districts adopt NZPs in different years, the number of districts with and without NZPs varies across our panel.

Table C2. Descriptive Statistics of Schools With and Without NZPs in Districts with Cross-School NZP Variation

	All Schools		Elementary Schools		Middle Schools		High School	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	NZP	No NZP	NZP	No NZP	NZP	No NZP	NZP	No NZP
District Characteristics								
# Students	554.2 (326.5)	585.4 (302.8)	432.0 (109.5)	456.9 (93.1)	607.5** (174.2)	722.1 (124.8)	1872.3* (56.3)	1568.9 (245.3)
Student Characteristics								
% Proficient on State Test	48.6 (17.5)	47.1 (17.4)	53.1 (18.2)	48.8 (17.9)	37.9 (10.9)	39.6 (12.6)	64.9 (1.5)	52.7 (18.6)
% Special education	15.9 (2.5)	16.2 (4.9)	16.2 (2.4)	16.9 (5.2)	15.9 (2.4)	14.9 (2.9)	11.7 (0.7)	12.0 (2.4)
% Low income	47.0 (23.8)	48.9 (21.2)	44.5 (25.4)	47.2 (20.4)	55.1 (18.8)	60.9 (20.0)	22.4 (1.3)	28.4 (8.3)
% English Language Learner	9.1 (9.0)	10.7 (8.6)	7.2* (7.8)	9.5 (7.6)	13.7 (10.0)	17.1 (9.5)	2.7 (0.4)	2.5 (1.5)
% Underrepresented minority	40.2 (19.6)	43.8 (20.0)	32.7** (14.5)	39.1 (15.8)	55.1 (21.3)	62.9 (23.7)	36.1 (2.5)	33.5 (4.3)
N Schools	12	48	6	38	5	9	1	1

Note: 1) Stars indicate significant differences between the means in Columns 1 and 2, the means in Columns 3 and 4, the means in Columns 5 and 6, and the means in Column 7 and 8: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ 2) N Schools represents the number of schools with NZPs and without NZPs in the final year of our sample. Because schools adopt NZPs in different years, the number of schools with and without NZPs varies across our panel.

Table C3. OLS Coefficients of School Covariates on the Likelihood of a School Adopting a No Zeros Policy, Weighted by School Size

	All NZPs		Elementary NZPs		Middle School NZPs	
	(1)	(2)	(3)	(4)	(5)	(6)
(# Students (1000s))	-0.0639 (0.1734)	0.3360 (0.5846)	-0.3308 (1.1532)	-0.1914 (0.3026)	-1.0956** (0.3560)	-0.1331 (0.6152)
% SpEd	0.0007 (0.0128)	-0.0000 (0.0091)	-0.0041 (0.0152)	0.0056 (0.0037)	0.0096 (0.0275)	-0.0002 (0.0250)
% Low income	0.0014 (0.0049)	0.0043 (0.0053)	0.0025 (0.0057)	0.0009 (0.0032)	-0.0002 (0.0089)	0.0057 (0.0095)
% ELL	-0.0042 (0.0114)	0.0048 (0.0047)	-0.0000 (0.0090)	-0.0031 (0.0031)	-0.0180 (0.0244)	0.0265 (0.0164)
% URM	-0.0015 (0.0044)	-0.0229** (0.0096)	-0.0035 (0.0048)	-0.0060 (0.0039)	0.0039 (0.0099)	-0.0515* (0.0249)
% Proficient on State Test	0.0016 (0.0030)	0.0012 (0.0024)	0.0025 (0.0062)	-0.0008 (0.0012)	-0.0078 (0.0078)	-0.0011 (0.0166)
Fixed Effects		Dist & Yr		Dist & Yr		Dist & Yr
Constant	0.0841 (0.3971)	0.6776 (0.7262)	0.3100 (0.7777)	0.4453 (0.3433)	1.2796* (0.5776)	2.8543 (1.5249)
N Districts	10	10	5	5	4	4
N Schools	34	34	24	24	8	8
F-statistic	0.3495	2.9176	0.4552	0.4257	6.4893	1.9670
Prob > F	0.9245	0.0215	0.8339	0.8542	0.0133	0.1984
R-squared	0.024	0.780	0.033	0.948	0.154	0.769

Note: 1) Standard errors in parentheses are clustered by school. 2) p-values from two-sided t-tests: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. 3) SpEd = Special Education, ELL = English Language Learner, URM = Underrepresented Minority. 4) Specifications with all NZPs also include dummy variables indicating the school level. These are not significant in any specification.