

Grading Policy Evaluation & Review

By Russell Carlock, Ed.D.

June 11, 2026

TABLE OF CONTENTS

ACKNOWLEDGEMENTS.....	2
EXECUTIVE SUMMARY.....	3
I. INTRODUCTION.....	6
II. METHODS.....	10
III. COMPARATIVE POLICY ANALYSIS.....	10
IV. REGRESSION ANALYSIS.....	18
V. TEACHER SURVEY.....	33
VI. STUDENT FOCUS GROUPS.....	43
VII. PARENT FOCUS GROUPS.....	47
VIII. CONCLUSIONS.....	54
IX. RECOMMENDATIONS.....	55
X. REFERENCES.....	59
XI. APPENDIX A: Full Results from Teacher Survey.....	61

ACKNOWLEDGEMENTS

Thank you to the teachers, students, and parents who contributed perspectives and feedback to produce this report. The study informing this report was produced with the support of a Research Practice Partnership between Albemarle County Public Schools and the University of Virginia. The feedback and recommendations on methods and analysis by university researchers Asha Muralidharan and Nina Schoonover were invaluable. Much gratitude as well goes to Chris Gilman and Sara Dusenberry for their astute work in survey creation and data collection. Finally, thanks to the many ACPS staff who reviewed and provided commentary through multiple drafts to make the report as clear and useful as possible to inform everyday practice and improve student learning.

EXECUTIVE SUMMARY

In 2020, the Albemarle County Public Schools' School Board passed its first comprehensive grading policy that attempted to institute principles of Standards Based Grading (SBG). This policy came under review in 2026, and the School Board called for a comprehensive study to review the policy and its associated guidelines. This report includes results and recommendations from a comparative policy analysis, quantitative analysis of grades' relationship to standardized assessment scores, a survey to all ACPS teachers, and six focus groups with students and parents.

Findings

Comparative policy analysis: This analysis revealed that many Virginia divisions aligned with ACPS in its adoption of SBG practices. **However, the procedures around implementing SBG grading practices do not currently align with peer divisions.** Specifically, ACPS has no specific standards regarding:

- implementing reassessments
- evaluating late work
- differentiating between unattempted student work versus student work that is submitted but is far below standards
- communicating grades to parents
- ensuring minimum standards in digital gradebook maintenance such as a minimum number of assignments for a marking period

In addition, ACPS is one of only two divisions that use a three point scale for elementary school, and the only division among peers that does not differentiate between honors and college level courses when awarding bonus quality points for GPA calculation.

Quantitative analysis: Across all student groups, the alignment between grades and test achievement in secondary reading has held relatively static, but students of different achievement levels differ in their alignment. In secondary schools, the alignment between grades and achievement has improved for lower-achieving students, but grades have become less valid indicators of achievement for higher achieving students. For elementary schools, grades have become less valid for all students. **This provides evidence that the 2020 grading changes have had the intended impact for secondary students who historically have struggled academically, but has had the opposite effect for more advantaged students.** Since 2019, both test achievement and grades in elementary reading and secondary math have dropped for nearly all groups.

Teacher survey: There was general support among ACPS teachers for the core beliefs of SBG. However there are deep concerns about how SBG has been implemented in the 2020 grading policy and guidelines. **Teachers believe that the grade floor of 50, a lack of parameters governing retakes, and the exclusion of grades on work habits have**

contributed to reduced motivation among students and a heavier workload on staff. Teacher-proposed solutions included:

- assigning grades lower than 50 for work that is not attempted.
- limiting retakes to students who need them most, capping the grade for which a retake counts, limiting students to one retake per assessment, and requiring remediation or study prior to a retake .
- re-incorporating grades on work habits with uniform and reasonable penalties and deadlines for late work and providing separate grades specifically on work-readiness skills that do not count towards GPA.

Student focus groups: In focus groups with students (n=47), **students noted inconsistency in application of the grading guidelines across classrooms and expressed a desire for greater consistency.** Students agreed that:

- Retakes can be beneficial or detrimental to learning depending on whether or not feedback and remediation are provided. Retakes can improve learning but students supported some limitations.
- Formative assessments should be required prior to summative assessments with regular teacher feedback between assessments as well as prior to retakes.
- Grades should differentiate between attempted assignments and those not attempted at all.
- Grace periods after deadlines to turn in work should exist, with a limit on that grace period as well as reasonable penalties for late work.

Students were divided on the grade floor of 50 percent, citing potential for both positive and negative impacts on motivation.

Parent/Guardian focus groups: Parents and guardians (n=28) **expressed support for the core tenets of SBG:** aligning grades to standards and providing spiraling curricula that allow students to revisit standards using different modalities and demonstrate their understanding of standards throughout the year; however **parents/guardians also noted many problems with implementation of the current grading procedures**, and generally desired:

- More understanding of students' work habits, either incorporated into an overall grade or reported separately.
- Limits on retakes, with individual feedback and re-assessment using different modalities prior to retakes.
- Greater consistency in policies around late work.
- A 4 point scale in elementary school grading to which all reported assessment data could be aligned.
- A minimum number of assignments per marking period to adequately describe student progress.

Parents were also divided on a grading floor of 50 for attempted work, but found **consensus in giving lower grades for assignments not attempted.**

Conclusions

ACPS's 2020 grading policy changes have succeeded in moving the division towards the principals of SBG, however are not meeting the broader goals of consistency, validity, and improved student learning due to issues with implementation. Below are conclusions from each phase of the analysis:

- While grade alignment for lower-performing students has remarkably improved, the opposite is true for

higher performing students. The evidence of grade inflation amongst higher-performing groups has diminished the overall district-wide gains in grading validity.

- Students, parents, and teachers embrace many principles of Standards Based Grading. Participants shared that elements like properly implemented retakes and a grade floor of 50 can improve learning and support students in certain contexts when important limitations are in place. However, inconsistent implementation with few parameters decreases student motivation and compromises grading validity.
- Comparative policy analysis revealed ACPS as an outlier among peer divisions in its lack of specificity on retakes, late work, and grading scales, contributing to implementation problems.

Recommendations

To address implementation inconsistencies and improve grading validity, **ACPS should make grading guidelines into formal, enforceable School Board procedures.** The call for procedures should address the most salient issues raised by the community: limitations on retakes, differentiating between attempted and unattempted assignments, and clear guidelines for late work. The School Board should also adopt language around minimum expectations for parent communication directly in its policy.

Recommendations for Updates to Procedural Standards for 2026-2027

- **Implement Reassessment Guardrails:** Limit students to one retake per summative assessment, only for students scoring below a letter grade of B, and capped at a maximum grade of the minimum threshold for a letter grade of B. Require feedback before retakes.
- **Establish a Differentiated Grade Floor:** Adopt a grading scale of 40-100 with a floor of 40 for "not attempted" work and 50 for "completed" work. This creates a clear six-category grading scale that addresses community concerns while attempting to preserve the validity improvements for lower-performing students.
- **Increase Grading Transparency:** Officially publish the secondary grading scale within the procedures.
- **Pacing and Systems:** Mandate a minimum number of graded assignments per quarter and describe uniform, rolling-semester gradebook setups across the division.
- **Late Work:** Define clear, division-wide parameters and consistent penalties for late submissions.
- **Parent Engagement:** Require proactive communication regarding the grading system at the start of the year, and mandate direct teacher notification to parents when a student is at risk of failure.

Recommendations for Updates to Procedural Standards for 2027-2028

- Adopt a 4-point grading scale for elementary schools and align all assessment data to this scale.
- Create a set of elementary grading procedures specific to the needs of elementary students and staff.
- Adopt procedures for reporting on student work habits separate from academic achievement and mandate that grades for work habits be reported separately from academic mastery.

I. INTRODUCTION

Albemarle County Public Schools Grading Policy

The Albemarle County School Board adopted a new grading policy and guidelines in 2020 that aligned with a national movement towards the principles of Standards Based Grading (SBG). Proponents of SBG tout its improved accuracy, enhanced ability to communicate progress to students and families; and reduction of bias, variation, and subjectivity in grading relative to the traditional point based system. Standards Based Grading is also called *proficiency based grading*, *competency based grading*, and *equitable grading*, and shares foundational concepts with Mastery Learning. Though there are nuances in these terms, they share a focus on three core principles:

- grades should be a reflection of a student’s learning of specific standards;
- grades on academic standards should be separated from grades on process (or non-cognitive) standards such as effort, timeliness, and work skills;
- and grades should be reported on a limited number, usually four to seven, categories to improve accuracy.

Albemarle County Public Schools adopted its new policy with the goals of better communicating student learning and providing more precise feedback on progress towards standards to students and parents.

The 2020 policy is now up for review, and this report will provide data and analysis to inform the School Board’s decision making.

Research on Grading Practices

In the early 20th century, most American high schools used the 100 point percentage grading scale, with no letter grades, to evaluate students; while elementary schools used narrative reports to communicate student performance to parents. The earliest research on grading focused on the reliability of grades on the 100 point scale as they related to assessments of content knowledge, and this research found substantial variation and little reliability in teachers’ grades. It was found that schools using categorical grading systems were more effective at measuring student achievement than those using the percentage scale (F. J. Kelly, 1914; Rugg, 1918). For this reason, secondary schools began to adopt the categorical grading system of A-E or F, with some schools opting for the full categorical system, but most translating percentage scores into letter grades, known as “percentage equivalent categorical grading.” Research from over 100 years has consistently found the traditional 100 point scale to be less reliable in reporting on student achievement compared to a categorical system (Brookhart et al, 2016).

In the latter half of the 20th century and early parts of the 21st, research focused on the sources of variation in grading, and found that teacher-assigned grades were often amalgams of cognitive and non-cognitive factors, such that teachers assigned a grade partly on a student’s mastery of the curriculum, but also on work habits, engagement, and process. Teachers used grades on non-cognitive factors to influence student motivation and completion of work, and this had

moderate but significant positive impacts on achievement (Willingham et al, 2002; Kelly, 2008). An attenuating drawback of this amalgam of cognitive and non-cognitive factors, however, was continued variability across teachers, as some teachers valued certain non-cognitive factors more than others. It was also difficult to communicate cognitive and non-cognitive factors separately in the traditional system (Brookhart, 1991).

In further exploring grading on cognitive and non-cognitive factors, researchers reported it was possible to separate these domains to preserve the independent value of each. For example, teachers could assign grades on academic achievement to provide a more accurate measure of mastery of the content, while also providing one or more grades on non-cognitive factors such as engagement and work habits. This system is used in all schools in Sweden and shows promise in providing a more accurate way to report on content mastery while preserving the tool of non-cognitive grades to increase student motivation (Thorsen & Cliffordson, 2012).

Contemporary research has focused on the adoption of Standards Based Grading - a system that combines a categorical grading system with a focus on reporting on mastery of specific standards at a specific time and separating grades on cognitive and non-cognitive factors. Although rigorous research on the effectiveness of SBG is sparse, some studies show that when teachers report grades on individual standards and support students to track their growth towards mastery, learning improves (Haystead & Marzano, 2009). Researchers have also found that parents generally prefer this system of reporting (Swan et al., 2014). A study of teachers using SBG reported that teachers generally agreed with the principles of SBG and believed it led to improved communication with parents; however they also shared that it was difficult to implement, time consuming, and initially had a negative impact on student motivation (Knight & Cooper, 2019).

The recent impetus for the national movement towards Standards Based Grading has two primary drivers. First, the inaccuracy and inconsistency of traditional grading practices spurred a call for a new system by parents, students, and educators (Guskey, 2024). Second, the adoption of standards based education, specifically the *No Child Left Behind Act*, the *Every Student Succeeds Act*, and the adoption of *Common Core Standards* across most states, made Standards Based Grading more appealing as both movements shared similar goals - to reduce variation across schools in quality and content and ensure a minimum level of cognitive rigor for student learning (Zhang, 2024).

Challenges in Implementation

While the problems SBG purported to solve were clear, implementing the new system has proven challenging for school divisions over the last decade as they wrestled with how to communicate changes in grading to parents, prepare teachers to use a new way of grading, and consider the details and logistics of implementing the principles. What followed has been varied and inconsistent adoption of SBG across divisions and a strong pushback against some aspects of implementation by both teachers and parents, leading some divisions to amend or abandon reform efforts.

Three of the most controversial aspects of divisions' attempts to adopt principles of SBG have been the elimination of grades for practice work, requiring retakes for students, and the adoption of a minimum grade, usually 50, to truncate the point based grading system for greater accuracy. Opponents argue that not grading practice like homework and allowing retakes lowers motivation in students to complete homework and study for assessments, which in turn lowers their overall performance. Similarly, opponents of a minimum grade argue that it does not teach accountability to students to turn work in on time and makes it too easy to pass a class (Guskey et al., 2020).

The above-mentioned research on the role of non-cognitive grades in student motivation provides support to these claims, as does more contemporary research on attempts to implement SBG without continuing to report on non-cognitive skills separately. A study on secondary mathematics found that when grades for work skills were eliminated, student motivation to complete practice work declined. The same study noted that when retakes were allowed for all students, it reduced students' motivation to study for assessments (Huey et al., 2022). Teachers also point out that allowing students to retake assessments adds substantial time to their workload, even if they agree it may provide more accurate information on student learning.

Researchers on grading point out that none of these measures - prohibiting grading practice work, allowing unlimited retakes, or setting a minimum grade of 50 - are inherent to or essential for Standards Based Grading. Rather, they represent attempts to forcibly fit the principles of SBG onto the points based system, often leading to confusion (Guskey et al., 2024).

A true adoption of SBG would continue to report grades for work skills such as timeliness and completion of practice while separating these from grades on purely academic standards. Similarly, SBG does not require retakes. Instead, it advocates that a student's grade be updated continually to represent the most up-to-date and accurate measure of the learning of a standard at that moment in time. Thus, rather than retakes, students could wait until the next assessment, study hard, do well, and this would take the place of a previous poor grade to the extent that it represented the same standards.

Retakes can also be limited to a certain time-frame and for certain levels of performance, noting the emphasis on a grade representing a student's learning of a standard at the reported time period, and expediting remediation for students who need it the most. Similarly, SBG does not require a minimum grade of 50, but rather requires the adoption of four to seven categories to replace the 100 point system. Rubrics created by teachers can define the levels of performance for each standard or domain. Thus, much of the pushback against these policies has been about implementation guidelines that are not required by the core principles of standards based reform.

Researcher Thomas Guskey (2024) provided the grading statement from the *American School in Paris* as a strong model for describing a true standards-based grading system:

The primary purpose of grading is to effectively communicate student achievement toward specific standards, at this point in time. A grade should reflect what a student knows and is able to do. Students will receive separate feedback and evaluation on their learning habits, which will not be included in the academic achievement grades.

The *American School in Paris* uses a seven point grading scale that is aligned with the *International Baccalaureate* (IB) grading system. The IB system grades on seven categories ranging from “very poor” to “excellent.”

Standards Based Grading in Virginia

Albemarle County Public Schools finds itself in the middle of these policy debates in Virginia. Virginia Beach and Fairfax are instructive as two of the first Virginia school divisions to adopt the principles of standards based grading. [A Virginia Department of Education Superintendent’s Bulletin from 2021](#) encouraged all Virginia School divisions to consider reforms related to SBG, and cited Virginia Beach and Fairfax as examples (Lane, 2021). Both divisions adopted guidelines that included a minimum grade of 50 for secondary schools, separating non-cognitive grades from academic grades, and requiring retakes for assessments. Neither fully adopted SBG, but instead used these guidelines to create an amalgam of the standards based system with the points system. Both divisions faced pushback from the community and teachers focused mostly on minimum grade and retake policies. Both divisions have since revised their policies to address these concerns.

Evaluating ACPS’s Grading Policy

The ACPS School Board has an opportunity to learn from both its constituents and the successes and challenges of other divisions navigating these complex questions. This report seeks to present the voices of teachers, students, and families, while also examining quantitative metrics to understand if there is evidence that the policy is having its intended impact.

The following questions guided the analysis in this report:

1. How does ACPS’s grading policy and guidelines compare to those of peer divisions in Virginia?
2. How, if at all, has the relationship between students’ grades and achievement changed since adoption of the policy in 2020?
3. How do teachers rate the effectiveness of the policy and the current grading handbook on achieving the goals of accuracy, improving student learning, and consistency?
4. How do teachers, students, and families describe the impact of the policy and current grading handbook on their work and experience in ACPS’s education system?

II. METHODS

Analysis proceeded under a mixed-methods approach, beginning with a comparative policy analysis of ACPS with peer districts in Virginia. This analysis used deductive coding with themes identified from the literature that allowed for comparison across district policies with the goal of identifying areas of alignment and discrepancy in ACPS's policies relative to its peers. Fourteen peer districts' policies were searched using the terms, "Grading," "Grades," "Student Evaluation," and "Student Progress." Only two divisions, Louisa and Fluvanna, did not yield clear policies on grading using these terms. The remaining 12 districts included Arlington, Fairfax, Stafford, Harrisonburg, Charlottesville, Roanoke, Virginia Beach, Loudoun, York, Henrico, Augusta, and Chesterfield.

Next, a regression analysis of student grades and achievement on standardized assessments across multiple years was performed to test the hypothesis that student grades become more valid predictors of achievement after adoption of the grading reforms, and to note any differences in this relationship across student groups. This analysis included over 7,800 unique student records of SOL testing data for reading in grades 6, 8, and 11 in the school years 2018-19, 2021-22, and 2024-25 and for mathematics (n=5,243) in grades 8 and 9 for the same years. To analyze elementary school grading, the study used an aggregate score of reading standards on the annual report card and reading SOL test scores in grade 5, also for the school years above, and included more than 6,000 unique student records.

Qualitative analysis also involved survey data collected from all teachers in the school division (n = 391). Deductive coding of the qualitative data from the survey took place during two rounds. Deductive codes were generated through the comparative policy analysis and focused on teacher feedback on grading policy and guidelines. These same codes guided analysis of data generated through focus groups of students, families, and teachers. Transcripts of focus groups were generated with the software [Otter.ai](#).

III. COMPARATIVE POLICY ANALYSIS

Based on a review of the literature, as well as an analysis of existing policies, regulations, and handbooks among peer districts, the following codes were developed: policy/regulations, philosophy, elementary scale, secondary scale, gradebook setup, grade weighting, minimum number of assignments, reassessment, minimum grade, late work, parent communication.

Policy/Regulation

The first distinction among school division policies on grading is the extent to which grading practices are dictated in policy versus regulations or guidelines published separately. Albemarle was among the majority of eight divisions that published a general purpose statement on grading that defined grades and their role in education without further detail. Seven other divisions provided additional policy on grades, with Fairfax and Augusta providing the most detail on grading practices in both elementary and secondary schools. Of the seven divisions that provided additional

information, all published grading scales for elementary, secondary, or both and included expectations around communication of grades to parents.

Most divisions provide a separate procedural document for implementation of the policy. These are most often called “regulations” or “policy implementation procedures.” Albemarle and Charlottesville call their documents “guidelines,” while Virginia Beach calls it a “handbook” for elementary schools.

Albemarle provides less detail in its policy than its peers, and this is most apparent in its lack of specificity around its grading scale and expectations for communicating grades to parents. It is also outside of the norm in providing “guidelines,” rather than regulations or policy implementation procedures. The School Board may wish to consider the level of expectation communicated with each of these terms for consistency in implementation.

When aggregating all regulations/guidelines and policies across school divisions to account for a total difference in clarity and detail, Albemarle stands as an outlier in the lack of clearly defined procedures provided to schools and teachers on how to implement its policy. While it has a research-based policy informed by SBG, its guidelines are often vague and lack metrics for accountability and standardization, leading to the potential for confusion and inconsistency for teachers and students in implementation. For each thematic category below, commentary is made on where ACPS lacks the clarity and detail of many peer divisions.

Philosophy

Divisions differed in whether, and to what extent, they adopted the principles of Standards Based Grading (SBG). Albemarle was among a small majority (7 of 13) of divisions that adopted clear language of Standards Based Grading in their policy. Virginia Beach was an anomaly in that it used SBG language for elementary schools, but not for secondary schools, due to its school-based process at the secondary level. Loudoun adopted language that was explicitly counter to SBG, and the remaining divisions had neutral language that did not refute or endorse principles of SBG.

A notable missing element of Albemarle’s language among its peers was that it did not require the reporting of grades on life skills and work habits. Four of the seven schools using SBG used language that explicitly required reporting on these, in addition to grades on academic achievement.

Overall, Albemarle’s language around the philosophy in SBG principles was a strength in that it was clear and succinct relative to peer divisions.

Elementary Scale

School divisions differed across two domains in elementary grading scales: the number of categories to rate student work, and the grade levels for which a numeric standards based system versus a letter grade content based system were used. Eight of 11 divisions that published elementary grading scales used a four point scale for standards based grading.

Only Albemarle, Charlottesville, and York used a three point scale. Below are two examples of four point scales from Arlington and Stafford (see Figures 1 and 2)

Arlington Public Schools Standards-based Reporting Performance Levels	
(4) Extending the Standard	In addition to meeting the standard, the student shows in-depth understanding and application of the standards taught this quarter.
(3) Meeting the Standard	The student shows full and consistent understanding of the standards taught this quarter.
(2) Approaching the Standard	The student shows partial and/or inconsistent understanding of the standards taught this quarter.
(1) Developing the Standard	With support, the student shows initial understanding of the standards taught this quarter.

Figure 1. Arlington Public Schools Standards-based Reporting Performance Levels

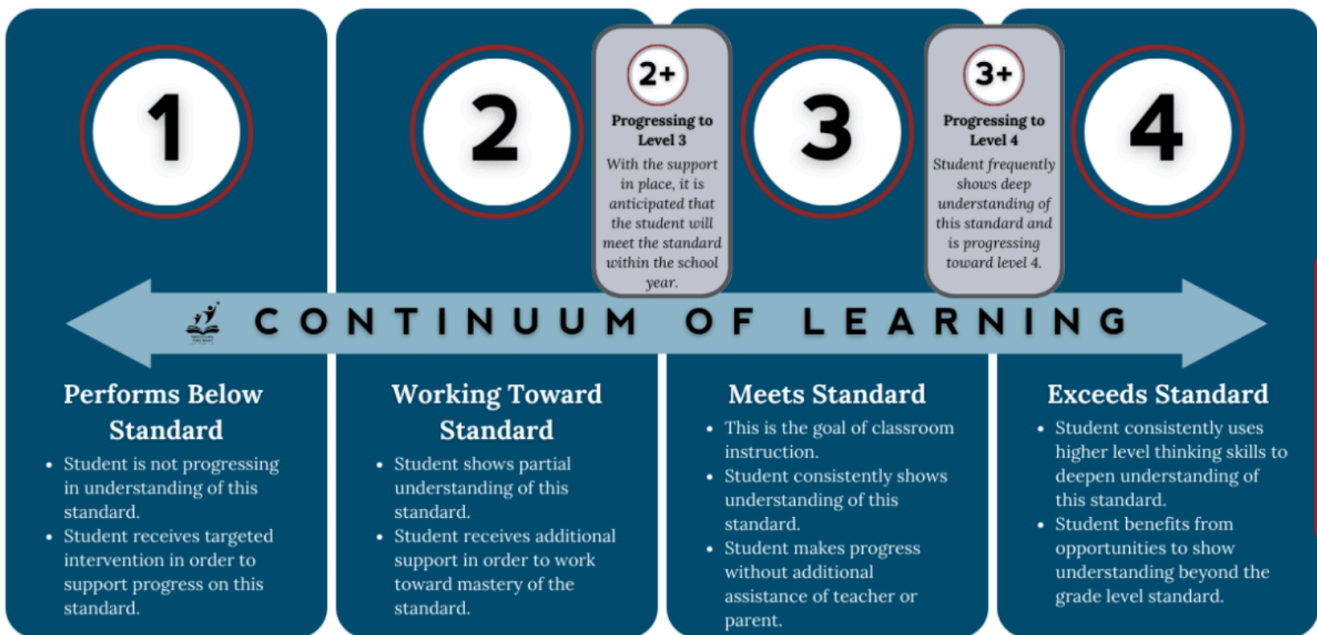


Figure 2. Stafford County Public Schools Four Point Scale

There was variation in when divisions switched from standards-based three or four point categorical grading systems to letter grading systems more similar to what is used in secondary schools. A plurality (five of 11) of divisions, including Albemarle, used SBG for all elementary students in grades K-5. Three divisions used SBG for K-1, and letter grades for all other grades. Two divisions used SBG for K-2, then moved to letter grades. Only Charlottesville provided SBG for K-4, then switched to letter grades in fifth grade.

Given its outlier status in using a three point scale for elementary SBG, research supporting SBG systems that use four to seven categories, and the alignment of a four point scale to Virginia’s categories of achievement on its standards,

Albemarle County may wish to consider moving to a four point grading scale for elementary schools for better accuracy and feedback on learning.

Secondary Scale

Areas of contrast in secondary grading scales include the grade range and the inclusion of +’s and -’s.

A plurality of peer divisions (5/11) continue to use a 0-100 point grading scale. Three divisions (Fairfax, Albemarle, and Charlottesville) use a 50-100 point scale. Augusta uses the 50 point scale for middle school, but continues to use a 0-100 point scale for high school courses. Arlington records final grades using the same quality points as the GPA system, creating an eight point categorical scale.

Most (eight of 13) divisions include secondary grades with +’s and -’s and five include only the letter grade. Most divisions set the failing grade at below 60; however three set this threshold at 65.

Nearly all divisions that continue to use a 0-100 point scale provide some provision for students to raise their grade to the lowest passing grade on the scale through remedial measures and if the student shows “reasonable effort” through attempts at completing work. This distinguishes the 0 as a grade reserved for when there is no evidence of students completing assignments.

Gradebook Setup & Weighting

Areas of contrast in gradebook setup include the standardization of gradebook types and the degree of division-wide mandate regarding category weighting.

A subset of peer divisions has moved toward higher levels of procedural standardization; for example, Fairfax and Arlington mandate a “yearlong rolling gradebook” to facilitate a cumulative view of student mastery. Albemarle and the majority of other peer divisions (9 of 13) do not specify a mandatory gradebook type in policy or guidelines, deferring these setup decisions to individual schools or teachers, or technology departments responsible for setting up division wide gradebooks.

The most significant divergence occurs in the methodology for category weighting. Four of the 13 divisions (Fairfax, Arlington, Augusta, and Chesterfield) mandate specific, division-wide weights for assessment categories to ensure consistency across buildings. For example, Fairfax requires a 70% summative and 30% formative split, while Augusta provides a four-part mandate (40% Major Assessments, 30% Minor Assessments, 20% Participation, and 10% Practice). Conversely, Albemarle and the remaining peer divisions provide no division-wide guidance on weighting, except for limiting the weight of homework and practice to between 10% and 20% of the final grade.

Nearly all divisions that lack a division-wide weighting mandate (such as Virginia Beach) still require a formal "bottom-up" consensus process where grade-level or department PLCs must agree on a consistent setup that is then approved by the principal. This distinguishes Albemarle as a qualitative outlier, as its handbook lacks the procedural requirement for documented consensus, leading to potential variance in how the same course is mathematically calculated across different classrooms or schools.

Minimum Number of Assignments

Another area of contrast in secondary grading is the standardization of the minimum number of graded assignments required to validate a marking period grade.

A subset of peer divisions (5 of 13) establishes a specific, division-wide numerical floor for assignments to ensure a consistent body of evidence. For instance, Fairfax and Arlington mandate a minimum of seven assignments per quarter for all classes. Charlottesville provides a range of 4-9. This procedural requirement is designed to ensure that a student's final grade is a reliable reflection of multiple performance indicators rather than a result of a few high-stakes data points.

Other divisions, such as Virginia Beach, utilize a framework that defers the specific count to school-based leadership or Professional Learning Communities (PLCs), requiring them to establish and document a consistent threshold. This approach provides content-specific flexibility while ensuring that students and families within a specific school or course have a predictable expectation for the volume of work required.

Albemarle is among the divisions that provide no specific division-wide mandated minimum threshold for assignments. In the absence of a mandated minimum, these divisions lack a centralized safeguard against "thin" gradebooks—where a student's entire quarter grade could theoretically rest on a very limited number of assessments.

Reassessment Policies, Regulations, and Guidelines

Areas of contrast in secondary reassessment include student eligibility thresholds, maximum score caps, and the frequency of retake opportunities.

A plurality of peer divisions (6 of 15) provides highly structured, division-wide guardrails to prevent grade inflation while still allowing for mastery. For example, Fairfax, Arlington, and Henrico implement a grade cap on reassessments (90%, 80%, and 70% respectively), and restrict eligibility to students who earned below these caps on the initial attempt. In these divisions, the policy is designed to support students who have not yet reached higher levels of proficiency.

Other divisions, such as Virginia Beach, delegate the specific parameters of reassessment to the school or Professional Learning Community (PLC) level. However, these divisions still require that those parameters be documented and applied consistently across a grade level or department.

Albemarle and Charlottesville stand out due to their high degree of permissiveness and lack of procedural caps. Both divisions provide no division-wide limit on the number of retakes, no maximum score cap (students can earn up to 100%), and no eligibility restrictions based on the initial grade. This distinguishes Albemarle and Charlottesville as qualitative outliers; while the quantity of philosophical guidance on mastery is high, the quantity of procedural limitations is nearly zero.

Minimum Grade

Areas of contrast in secondary minimum grade policies and regulations include the point at which a grade floor is applied (assignment level vs. marking period level) and the requirement of a "reasonable attempt" to trigger the minimum grade.

A plurality of peer divisions (5 of 13) utilizes a conditional 50% floor to prevent the "mathematical cliff" of a zero while still requiring evidence of effort. Fairfax, Charlottesville, and Loudoun all specify that a minimum grade of 50 is only assigned if a student makes a "reasonable attempt" or submits the work. In Fairfax and Charlottesville, a zero is explicitly permitted for work that is not attempted or submitted.

Other divisions utilize the marking period as the primary level of intervention. Henrico is unique in offering a "remediation plan" that allows students who are failing to raise their marking period grade to a 65% by completing additional work, though this is limited to one marking period per year for high school credit-bearing courses. Similarly, Loudoun and Charlottesville apply its 50% floor only to quarter and final grades—not individual assignments—to ensure that a single poor quarter does not irrecoverably damage a student's ability to pass the course.

Albemarle stands out as a qualitative outlier. While peer divisions like Fairfax and Charlottesville use "reasonable attempt" clauses to distinguish between struggling learners and non-submitting students, Albemarle's guidelines set a minimum grade of 50 on all assignments and marking periods. This lack of a submission requirement obscures the difference between a student who attempted the material and one who did not.

Arlington provides a system most aligned with SBG in that all quarter and final grades are converted to quality points. This means that while a student may receive 0s on individual assignments, resulting in an F for a quarter grade, the quality points will remain on an 8 point scale, leading to a system that allows distinction between students who turn in work versus those who do not; while also preserving the mathematical accuracy of SBG.

Late Work

Areas of contrast in secondary late work practice include the application of penalties and the establishment of hard deadlines for submission.

A majority of peer divisions (7 of 13), including Fairfax, Arlington, York, and Chesterfield, allow or mandate point-based penalties for late work to incentivize time management. Arlington specifically defines a penalty cap of 10% to ensure that the final grade still reflects academic mastery while maintaining a procedural "cost" for lateness. This approach creates a distinction between students who demonstrate mastery within the instructional window and those who do so after the fact.

Five divisions, including Arlington, Augusta, Loudoun, York, and Chesterfield establish hard, division-wide "sunset clauses"—specifying, for example, that work will not be accepted more than two weeks after the due date or past the end of a specific unit. This provides teachers with a centralized mandate to refuse months-old work, preventing the back-grading surges that can occur at the end of a marking period.

Albemarle and Charlottesville stand out as qualitative outliers in their degree of permissiveness regarding both penalties and deadlines. Unlike Chesterfield or Virginia Beach, which allow for school-based late work structures, Albemarle and Charlottesville's guidelines discourage penalties and lack a division-wide expiration date for assignments.

Parent Communication of Grades

Areas of contrast in parent communication include the specific triggers for intervention (such as grade drops or failing marks) and the mandatory speed of entry for grades in digital gradebooks.

A majority of peer divisions (9 of 13) establish clear, pre-emptive triggers for communication that go beyond the quarterly report card. Chesterfield and Arlington are notable for their proactive requirements: Chesterfield requires teachers to attempt contact if a grade drops by two letter grades by the midpoint of a quarter, while Arlington mandates an "advisory report" be sent three weeks prior to the end of a marking period for any student performing at an unsatisfactory level. Similarly, Fairfax, Virginia Beach, and Henrico all require parental contact if a student is at risk of failing, prior to issuing a report card.

A second area of contrast is the procedural mandate for digital gradebook maintenance. Divisions like Loudoun, Augusta, York, and Arlington provide highly specific timelines for grade entry to ensure parents have real-time data. Loudoun and Arlington require grades to be entered within 10 days of the due date, Augusta requires within three days, and York allows between three and five days depending on the type of assignment. Charlottesville sets a minimum update frequency of every two weeks and requires "meaningful feedback in a timely manner."

Albemarle is an outlier in this category, as it does not specify communication triggers or grade entry timelines for teachers. While other divisions like Chesterfield use the communication policy to reinforce the separation of academic and non-academic factors, requiring separate feedback to parents on "academic effort" and "21st-century skills," Albemarle's policy lacks clear procedural requirements for teacher communication with parents. This distinguishes Albemarle as a qualitative outlier; without a division-wide mandate on how often grades must be updated or what specifically triggers a phone call home, the burden of determining sufficient communication falls entirely on individual teachers, potentially leading to significant variance in the family experience across the division. By not specifying a grade-entry window, Albemarle lacks a standard to ensure that the "Accuracy" and "Communication" goals stated in its grading policy are being met in practice.

Grade Point Average (GPA) Calculations

Areas of contrast in secondary GPA calculations include the tiered weighting of honors versus college level coursework and specific accountability requirements for awarding bonus points in college level classes.

A plurality of peer divisions (6 of 13) provide a standardized, tiered weighting system. Divisions such as Fairfax, Chesterfield, Roanoke City, Henrico, Loudoun, and Harrisonburg all utilize a consistent formula: a 0.5-point boost for honors courses and a 1.0-point boost for college-level courses (including Advanced Placement, International Baccalaureate, or Dual Enrollment). This tiered approach is designed to create a distinction that acknowledges the varying levels of rigor between local advanced curricula and external, college-level standards. Two divisions provide bonus points only for college level courses.

Specific procedural requirements for these points also vary across divisions. Arlington is an example of accountability weighting, as it mandates that students must actually sit for the AP exam to receive the 1.0 point boost.

Albemarle stands out among its peers in that it awards one additional quality point for honors *and* college-level courses.

Overall Findings

The comparative policy analysis reveals that ACPS's grading policy and handbook, while aligned with most of its peers in adoption of principles of SBG, is a qualitative outlier in many aspects of the procedural implementation of these principles as well as in grading guidelines outside of the scope of SBG. Specifically, ACPS has few or no specific metrics for implementation of reassessments, evaluating late work, or differentiating between student work that is not attempted at all versus student work that is submitted but is far below standards. Additionally, ACPS has no specific standards in its policy or guidelines for communication of grades to parents or protocol for ensuring minimum standards in digital gradebook maintenance. Finally, ACPS is one of only two divisions that use a three point scale for elementary school, and the only division among peers that does not differentiate between honors and college level courses when awarding bonus quality points for GPA calculation.

IV. REGRESSION ANALYSIS

One measurable goal for SBG-inspired grading systems is whether or not they lead to better alignment between teacher-assigned grades for academics and external measures of achievement such as standardized tests. To understand whether or not the adoption of ACPS's grading policy and guidelines in 2020 have led to a stronger positive relationship between grades and achievement, or not, regression analyses were conducted on datasets in elementary and secondary reading, as well as in secondary mathematics, on datasets from 2018-2020, 2021-2022, and 2024-2025. Datasets included student demographic variables, teacher assigned final grades for a course in numeric and letter category form, school, and the first-attempt SOL assessment score for the course.

To analyze potential differential impacts of the grading policy and guidelines on students based on their prior achievement, SOL test scores for the same content area for the previous year those scores were available were also included. To ensure the relationship was not artificially strengthened due to the truncation of the grading scale from 0-100 to 50-100 after 2018-2019, analyses were conducted both on the actual scores as well as on datasets that converted all scores below 50 in 2018-2019 to 50s. Similarly, prior to the grading changes, extra credit led to some students receiving percent grades as high as 115, and these were converted to 100. Independent analyses were conducted by data scientist Russell Carlock with Albemarle County Public Schools and on deidentified data by University of Virginia Center for Community Partnerships data scientist, Asha Muralidharan, to ensure alignment of findings.

Secondary Reading Grades and Achievement

The secondary reading dataset consisted of 7,877 individual student records from one school year prior to the grading policy changes (2018-2019), and two school years after its adoption (2021-2022, 2024-2025). Figure 3 summarizes the data collected, which includes data from students in grades 6, 8, and 11 in English Language Arts. All levels except for college level courses in 11th grade were included, as the curricula for Advanced Placement and Dual Enrollment English are not aligned directly to the Virginia standards of learning assessed.

Secondary Reading Dataset

Characteristic	2018-2019, N = 2,793 ¹	2021-2022, N = 2,659 ¹	2024-2025, N = 2,547 ¹
gender			
F	1,358 (49%)	1,288 (48%)	1,172 (46%)
M	1,435 (51%)	1,371 (52%)	1,375 (54%)
grade			
6	1,039 (37%)	891 (34%)	939 (37%)
8	979 (35%)	992 (37%)	948 (37%)
11	775 (28%)	776 (29%)	660 (26%)
econ_dis	729 (26%)	695 (26%)	764 (30%)
race_ethnicity			
AS	149 (5.3%)	128 (4.8%)	158 (6.2%)
BL	331 (12%)	339 (13%)	322 (13%)
HL	302 (11%)	372 (14%)	476 (19%)
MR	155 (5.5%)	171 (6.4%)	203 (8.0%)
WH	1,856 (66%)	1,649 (62%)	1,388 (54%)
swd			
N	2,452 (88%)	2,347 (88%)	2,184 (86%)
Y	341 (12%)	312 (12%)	363 (14%)
el			
N	2,556 (92%)	2,420 (91%)	2,302 (90%)
Y	237 (8.5%)	239 (9.0%)	245 (9.6%)
grade_alpha			
A	1,157 (41%)	1,208 (45%)	1,021 (40%)
B	908 (33%)	752 (28%)	795 (31%)
C	477 (17%)	425 (16%)	440 (17%)
D	202 (7.2%)	204 (7.7%)	220 (8.6%)
F	49 (1.8%)	70 (2.6%)	71 (2.8%)
grade_num	84.7 (11.1)	85.3 (11.3)	84.3 (11.5)
score	443.6 (61.8)	448.8 (66.6)	438.1 (69.2)

¹ n (%); Mean (SD)

Figure 3. Secondary Reading Dataset

Descriptive statistics in Figure 3 and density plots in Figure 4 do not show substantial differences in grade distribution, numeric grades, or average SOL test reading scores from before and after the implementation of the grading policy and guidelines; however this high level analysis can mask substantial differences in the relationship between grading and achievement for different groups of students.

Grade Validity in Secondary Reading

Distribution of actual SOL scores for students by grade earned in class

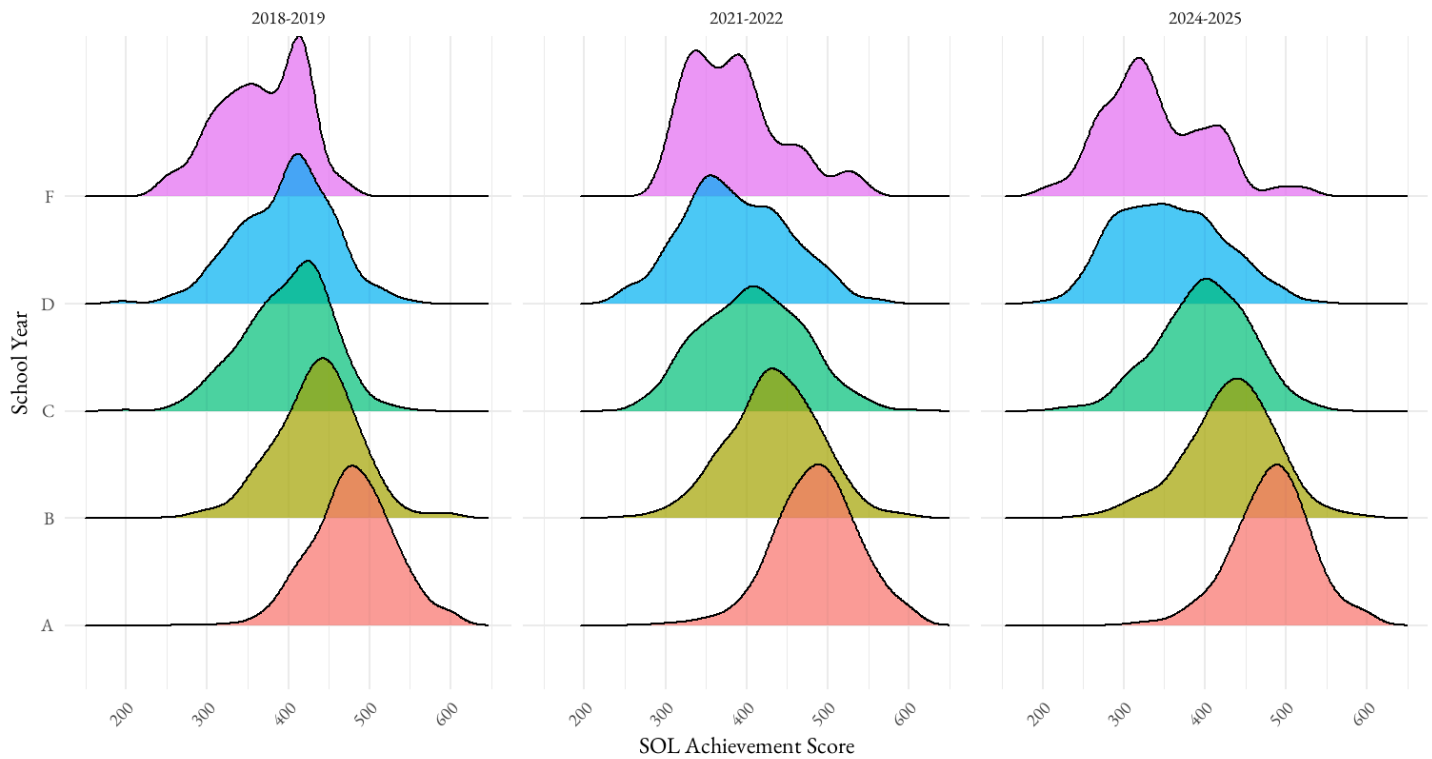


Figure 4. Grade Validity in Secondary Reading

It is possible, for example, that changes in the grading policy impacted different groups of students differently, resulting in little change in overall descriptive statistics, but important differences among students. It is also possible to overlook changes in the relationship of grades to achievement by looking only at averages and grade distribution. The regression analyses examined the relationship between these variables using each point in the dataset, and explored differences among student groups.

An Ordinary Least Squares (OLS) regression model was used to measure the shift in predictive validity of classroom grades in English Language Arts classes on standardized reading scores:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 (x_1 \times x_2) + \sum \gamma_i c_i + \epsilon$$

All continuous variables were transformed into Z-scores to standardize variables. The reference year was set at 2018-2019 to establish a pre-policy baseline, allowing the model to treat subsequent years as treatments to measure for any shift in the relationship between grades and standardized test scores. The key part of the analysis was the interaction term `std_grade:year` - a positive result provided evidence for grades becoming more tightly coupled with achievement. A negative result would indicate decoupling between grades and achievement.

Stratified modeling across demographic membership groups allowed for the testing of heterogeneous treatment effects, or whether there was evidence of the grading policy and guidelines affecting different groups of students differently. Adjusted r-squared values allowed for the measurement of how much variance in achievement was explained by grades in each of the years. Results are in Figure 5.

Secondary Reading: Relationship Between Grades and Achievement

<i>Predictors</i>	Overall	Econ Disadv	Non-Econ Disadv	ELs	Non-ELs	SWD	Non-SWD	BL	HL	WH
	<i>Estimates</i>	<i>Estimates</i>	<i>Estimates</i>	<i>Estimates</i>	<i>Estimates</i>	<i>Estimates</i>	<i>Estimates</i>	<i>Estimates</i>	<i>Estimates</i>	<i>Estimates</i>
Baseline Intercept	0.34 ***	-0.46 ***	0.11 ***	-0.56 ***	0.04 **	-0.80 ***	0.09 ***	-0.49 ***	-0.39 ***	0.11 ***
Grade (Baseline 2018-19)	0.39 ***	0.31 ***	0.54 ***	0.40 ***	0.53 ***	0.36 ***	0.50 ***	0.33 ***	0.43 ***	0.51 ***
year2021-2022	0.05 **	0.07	0.03	-0.16	0.06 *	0.18 *	0.02	0.02	0.11	0.05 *
year2024-2025	-0.04	0.03	-0.06 *	-0.52 ***	-0.04	0.28 ***	-0.10 ***	-0.01	-0.02	0.00
genderM	0.02									
econ_disYes	-0.27 ***									
race_ethnicityBL	-0.53 ***									
race_ethnicityHL	-0.22 ***									
race_ethnicityMR	-0.14 **									
race_ethnicityWH	-0.08 *									
eIY	-0.68 ***									
swdY	-0.64 ***									
std_grade:year2021-2022	0.04 *	0.05	0.04	0.02	0.03	0.08	0.06 *	-0.00	-0.01	0.07 *
std_grade:year2024-2025	0.07 ***	0.20 ***	0.07 *	-0.01	0.07 **	0.21 **	0.12 ***	0.23 ***	0.17 **	0.00
Observations	7819	2105	5714	681	7138	964	6855	962	1108	4806
R ² / R ² adjusted	0.465 / 0.464	0.199 / 0.197	0.293 / 0.292	0.269 / 0.263	0.321 / 0.321	0.224 / 0.220	0.318 / 0.317	0.223 / 0.219	0.277 / 0.274	0.276 / 0.275

* $p < 0.05$ ** $p < 0.01$ *** $p < 0.001$

Figure 5. Secondary Reading Results

Results indicate that grades have overall become more aligned to achievement since the adoption and implementation of the grading policy; however this improvement in alignment has been different for different groups of students. While economically disadvantaged students, students with disabilities, Hispanic/Latino, and Black students have significantly more accurate grades than previously; the relationship between grades for other groups, such as White students and ELs, has not changed significantly. The scatterplot of grades and SOL test scores across the three years (Figure 6), faceted by race and ethnicity, demonstrate these results. For all groups except for White students, the relationship between grades and achievement has become stronger, indicated by the increasing positive slope of each year's line.

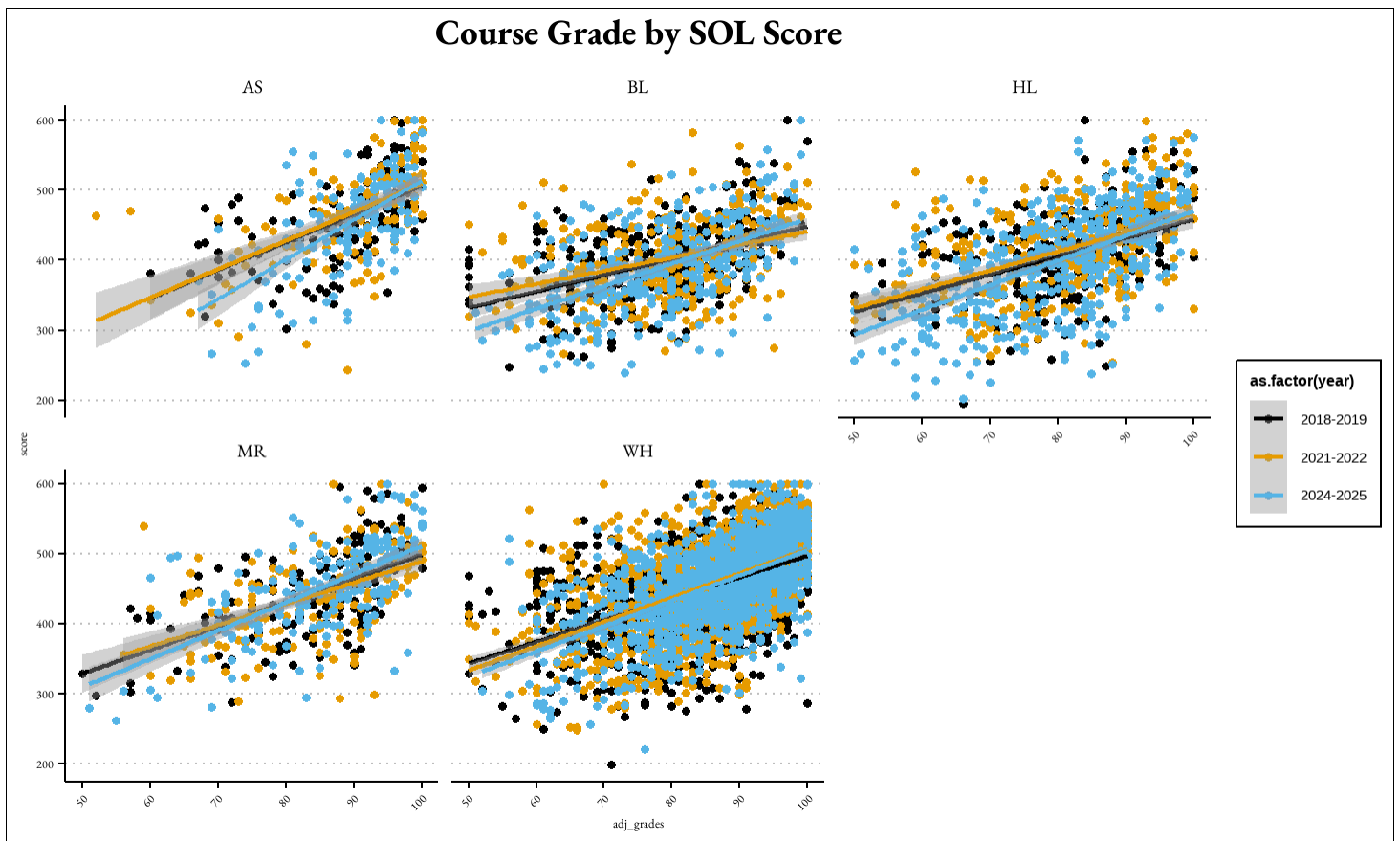


Figure 6. Course Grade by SOL Score

While this analysis allows us to conclude that overall grades have become more valid measures of student achievement in secondary reading, it leaves out an important method for grouping students to understand further how the grading policy may have different impacts on different students: by prior achievement. It stands to reason that students at different parts of the achievement spectrum may benefit from changes in the grading policy differently. For example, students who struggle academically may benefit more from re-teaching and re-assessment than students who are already doing well. To test for evidence of this, previous years' SOL test scores were used to divide groups into quartiles, and then regression analysis was used to measure if there were differences in the changes in the relationship between grades and achievement for these groups.

Constructing a dataset that used the previous test score in reading limited this analysis to students who had a previous score, excluding many transfer students. As transfer students on average tend to face more challenges in learning, this dataset represents a higher performing group on average than the previous dataset. That fact notwithstanding, it still provides valuable evidence on how the grading policy may be operating differently for different groups of students based on their prior achievement.

Figure 7 provides the results from the regression analysis on students grouped by previous achievement in reading. This analysis gives further evidence of the grading policy's differential impact on students and finds that all of the

improvement in grade reliability has occurred among the bottom 50th percentile of students. Students at the lower levels of achievement have experienced a statistically significant improvement in the accuracy of their grades since the adoption of the grading policy and guidelines while students at the higher ends of achievement have experienced the opposite: their grades are now less accurate predictors of achievement than before.

Grade Accuracy by Prior Achievement Quartiles (Tiers)

<i>Predictors</i>	Overall <i>Estimates</i>	Tier 1 (Lowest) <i>Estimates</i>	Tier 2 <i>Estimates</i>	Tier 3 <i>Estimates</i>	Tier 4 (Highest) <i>Estimates</i>
Baseline Intercept	0.21 ***	-0.72 ***	-0.15 ***	0.23 ***	0.62 ***
Grade (Baseline 2018-19)	0.52 ***	0.26 ***	0.26 ***	0.29 ***	0.38 ***
year2021-2022	-0.03	0.19 **	0.21 ***	0.12 *	0.22 **
year2024-2025	-0.05 *	0.12	0.07	0.19 ***	0.21 ***
Econ Disadv	-0.33 ***	-0.09 *	-0.16 ***	-0.12 *	-0.06
eLY	-0.50 ***	-0.28 ***	-0.08	-0.20	-0.01
swdY	-0.59 ***	-0.36 ***	-0.21 ***	-0.15	-0.16
std_grade:year2021-2022	0.09 **	0.11	0.05	0.13 *	-0.08
std_grade:year2024-2025	-0.05 *	0.11 *	0.06	-0.11 *	-0.27 ***
Observations	4541	1136	1135	1135	1135
R ² / R ² adjusted	0.497 / 0.496	0.312 / 0.307	0.248 / 0.243	0.185 / 0.180	0.137 / 0.131

* $p < 0.05$ ** $p < 0.01$ *** $p < 0.001$

Figure 7. Grade Accuracy by Prior Achievement Quartiles

A reasonable hypothesis to explain this trend is that the grading policy and guidelines are working as intended for students at the lower ends of the achievement spectrum while at the same time creating opportunities for students at the higher end to inflate their grades relative to their actual learning. One way to test this hypothesis was to analyze grade distributions by these same tiers and note how they have changed from before the policy and guidelines to after.

Figure 8 visualizes this analysis and demonstrates that since 2018-2019, students at the lower end of the achievement spectrum are more likely to earn failing grades while students at the higher end are more likely to earn As, demonstrating a tendency toward grade inflation among high achievers, and grading that better matches students' achievement performance among historically low achievers.

Differential Impact Based on Historical Achievement

Proportion of letter grades awarded, grouped by students' historical achievement tiers

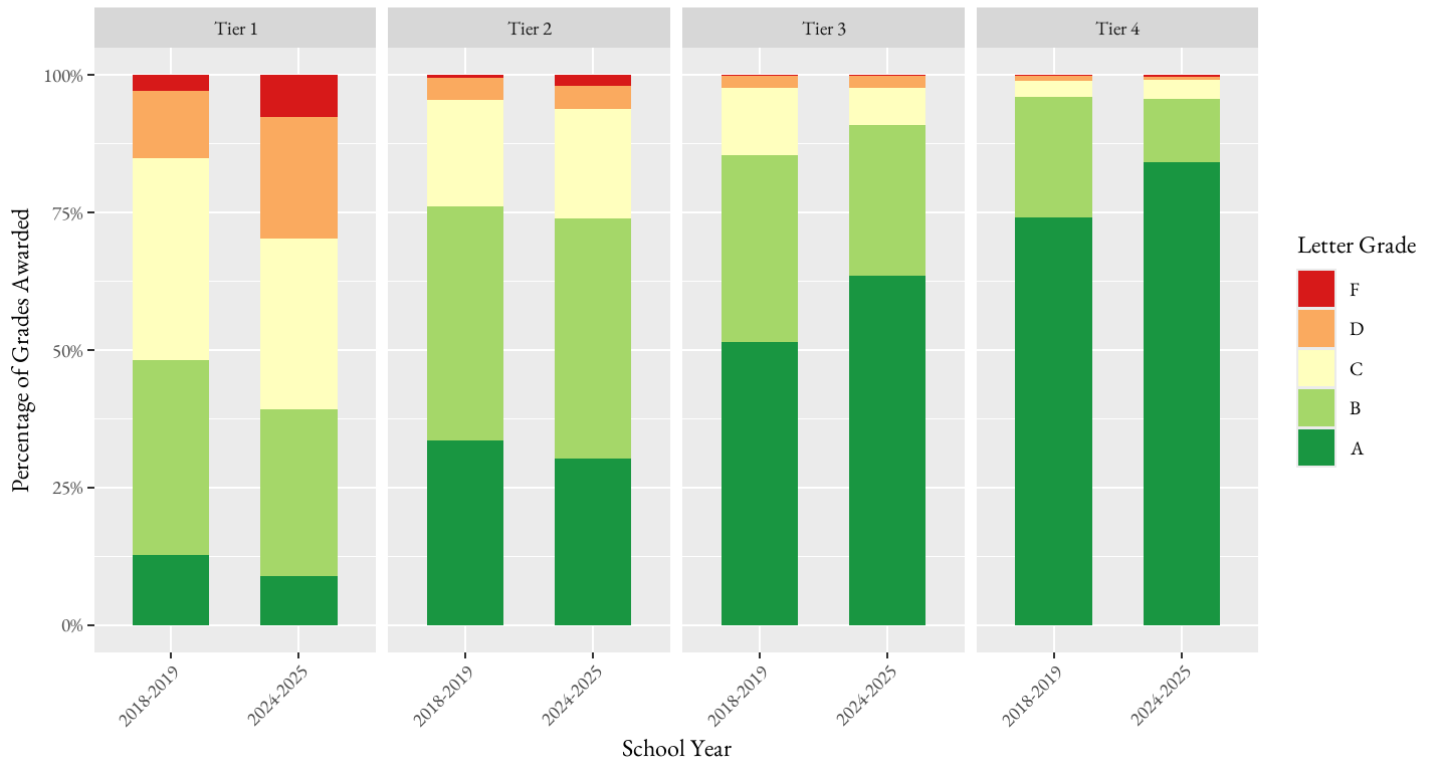


Figure 8. Differential Impact Based on Historical Achievement Reading

An analysis of SOL test scores for students in this group supports this interpretation. While lowest quartile students' SOL test scores have neither risen nor dropped substantially, their grades *have* dropped, indicating that teachers are now more likely to provide low achieving students with grades that match their low achievement. In sum, since the grading policy and guidelines changed in 2020, lower performing students' grades have become more accurate and lower, while higher performing students' grades have become less accurate and higher. This outcome has exacerbated equity gaps in secondary reading grades.

Now, we turn to the mathematics dataset to see if these same trends hold true.

Secondary Mathematics Grades and Achievement

The secondary mathematics dataset consisted of 5,243 individual student records from one school year prior to the grading policy changes (2018-2019), and two school years after its adoption (2021-2022, 2024-2025). Figure 9 summarizes the data collected, which included data from students in grades 8 and 9 in all math courses. These grades were chosen as students above grade nine often do not take a standardized math assessment, and some students in lower grades took SOL tests not aligned with their grade level in some years.

Secondary Mathematics Dataset

Characteristic	2018-2019, N = 1,721 ¹	2021-2022, N = 1,804 ¹	2024-2025, N = 1,717 ¹
gender			
F	848 (49%)	898 (50%)	838 (49%)
M	873 (51%)	906 (50%)	879 (51%)
grade			
8	969 (56%)	1,018 (56%)	961 (56%)
9	752 (44%)	786 (44%)	756 (44%)
econ_dis			
N	1,265 (74%)	1,265 (70%)	1,145 (67%)
Y	456 (26%)	539 (30%)	572 (33%)
race_ethnicity			
AS	80 (4.6%)	78 (4.3%)	86 (5.0%)
BL	207 (12%)	227 (13%)	223 (13%)
HL	214 (12%)	341 (19%)	353 (21%)
MR	104 (6.0%)	134 (7.4%)	119 (6.9%)
WH	1,116 (65%)	1,024 (57%)	936 (55%)
swd			
N	1,494 (87%)	1,578 (87%)	1,463 (85%)
Y	227 (13%)	226 (13%)	254 (15%)
el			
N	1,588 (92%)	1,559 (86%)	1,505 (88%)
Y	133 (7.7%)	245 (14%)	212 (12%)
grade_alpha			
A	656 (38%)	697 (39%)	494 (29%)
B	588 (34%)	530 (29%)	564 (33%)
C	313 (18%)	315 (17%)	346 (20%)
D	112 (6.5%)	189 (10%)	195 (11%)
F	52 (3.0%)	73 (4.0%)	118 (6.9%)
grade_num	83.7 (12.2)	83.2 (12.2)	80.7 (12.0)
score	443.4 (50.8)	420.3 (54.2)	426.4 (51.9)

¹ n (%); Mean (SD)

Figure 9. Secondary Mathematics Dataset

Descriptive statistics in Figure 9 show an increase in failures and a reduction in the number of As for students in 8th and 9th grade mathematics since the adoption of the policy; however student achievement also dropped during this time period, making it difficult to disentangle the impact of the policy from dropping achievement resulting from school closures and online learning. Similar to the reading data, both Figure 9 and density plots in Figure 10 show mild, but not substantial changes in distributions of test scores by grade from before and after the policy changes. For this reason, it is important to investigate evidence for potential differential impacts on groups of students.

Grade Validity in Secondary Mathematics

Distribution of actual SOL scores for students by grade earned in class

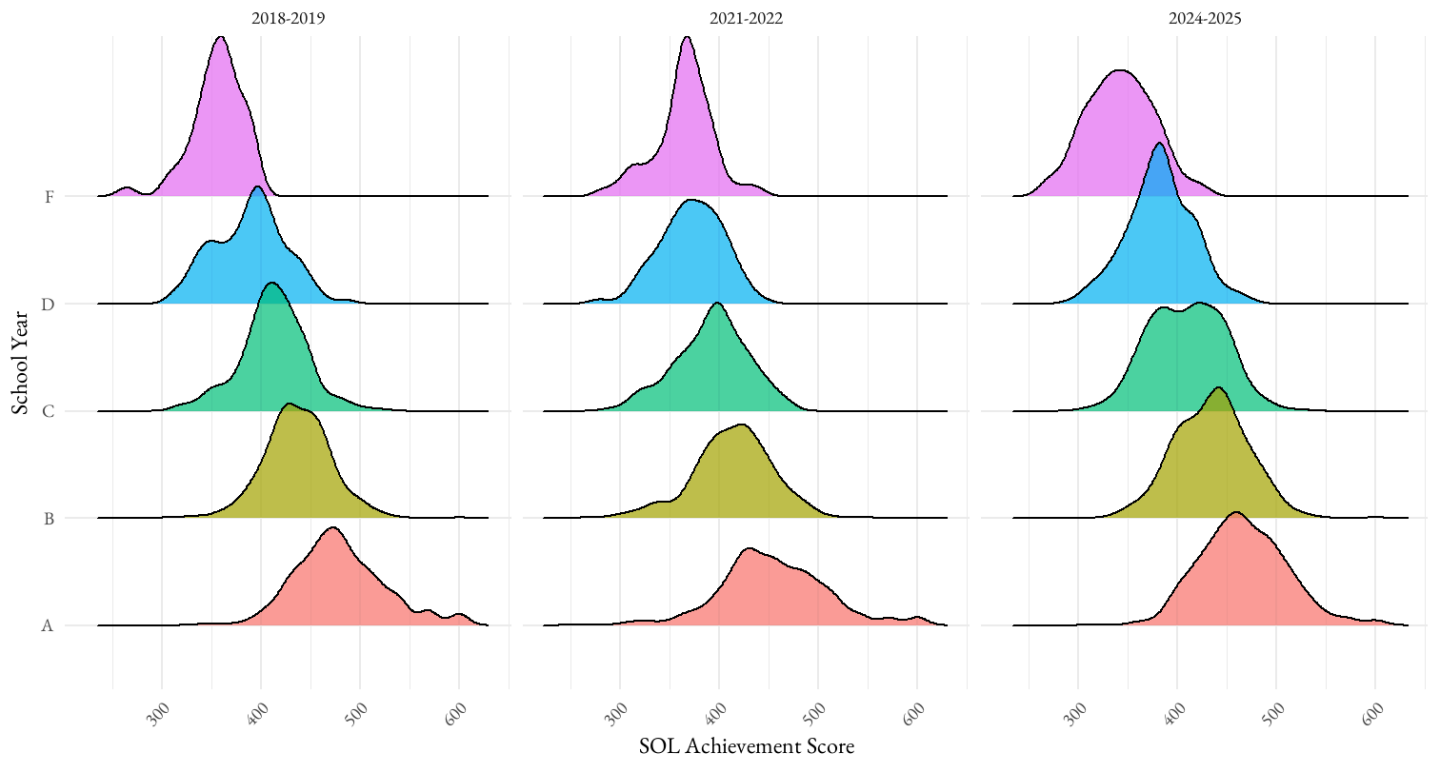


Figure 10. Grade Validity in Secondary Mathematics

The regression analysis examined the relationship between these variables using each point in the dataset to explore differences among student groups. This analysis controlled for the overall drop in achievement, allowing for an understanding of the relationship between grades and achievement independent of pandemic era losses. Figure 11 shows results of the regression model.

Secondary Mathematics: Relationship Between Grades and Achievement

	Overall	Econ Disadv	Non-Econ Disadv	ELs	Non-ELs	SWD	Non-SWD	BL	HL	WH	Male	Female
<i>Predictors</i>	<i>Estimates</i>	<i>Estimates</i>	<i>Estimates</i>	<i>Estimates</i>	<i>Estimates</i>	<i>Estimates</i>	<i>Estimates</i>	<i>Estimates</i>	<i>Estimates</i>	<i>Estimates</i>	<i>Estimates</i>	<i>Estimates</i>
Baseline Intercept	0.52 ***	-0.12 ***	0.22 ***	-0.35 ***	0.22 ***	-0.26 ***	0.23 ***	-0.22 ***	-0.12 *	0.28 ***	0.19 ***	0.13 ***
Grade (Baseline 2018-19)	0.57 ***	0.38 ***	0.78 ***	0.69 ***	0.64 ***	0.46 ***	0.64 ***	0.35 ***	0.63 ***	0.64 ***	0.57 ***	0.76 ***
year2021-2022	-0.38 ***	-0.43 ***	-0.37 ***	-0.36 ***	-0.40 ***	-0.51 ***	-0.41 ***	-0.35 ***	-0.36 ***	-0.41 ***	-0.39 ***	-0.41 ***
year2024-2025	-0.12 ***	-0.06	-0.10 **	-0.02	-0.14 ***	0.08	-0.17 ***	-0.05	-0.08	-0.14 ***	-0.10 *	-0.15 ***
genderM	0.09 ***											
econ_disY	-0.19 ***											
race_ethnicityBL	-0.50 ***											
race_ethnicityHL	-0.32 ***											
race_ethnicityMR	-0.32 ***											
race_ethnicityWH	-0.20 ***											
eLY	-0.44 ***											
swdY	-0.38 ***											
std_grade:year2021-2022	-0.05 *	0.01	-0.13 ***	-0.38 ***	-0.01	-0.11	-0.04	-0.00	-0.20 **	0.01	0.03	-0.15 ***
std_grade:year2024-2025	-0.00	0.23 ***	-0.10 **	-0.09	0.03	0.17 *	0.05	0.22 ***	-0.03	0.01	0.17 ***	-0.10 *
Observations	4976	1449	3527	540	4436	632	4344	613	853	2948	2512	2464
R ² / R ² adjusted	0.488 / 0.487	0.374 / 0.372	0.407 / 0.406	0.312 / 0.305	0.430 / 0.429	0.365 / 0.360	0.415 / 0.415	0.364 / 0.359	0.373 / 0.370	0.405 / 0.404	0.417 / 0.416	0.439 / 0.438

* $p < 0.05$ ** $p < 0.01$ *** $p < 0.001$

Figure 11. Secondary Mathematics Results

One item to note is that even prior to the pandemic and the adoption of the 2020 grading policy, math grades were more valid measures of achievement compared to reading grades. This greater baseline precision means the bar is higher to show statistically significant improvement in grading accuracy among math teachers.

These results show there has been no significant *overall* change in the relationship between grades and achievement in secondary mathematics since adoption of the policy; however, the math data tells a very similar and *even stronger story* as the reading data when separating out differential impacts on groups: While historically lower-performing groups' grades became significantly better predictors of achievement, including economically disadvantaged students, students with disabilities, Black students, and male students; other groups of students' grades became significantly less predictive of achievement, e.g., non-economically disadvantaged students and female students. The numbers of students in the dataset whose grades became less predictive of achievement canceled out any gains in accuracy among other students, leading to a net measure of 0 when evaluating the overall relationship.

Figure 12 demonstrates the improvement in accuracy for grades among economically disadvantaged students and the simultaneous further decoupling of grades from achievement for economically advantaged students. While all of these students' achievement dropped dramatically during the time of the pandemic, non-economically disadvantaged students were able to maintain higher grades relative to their drop in achievement.

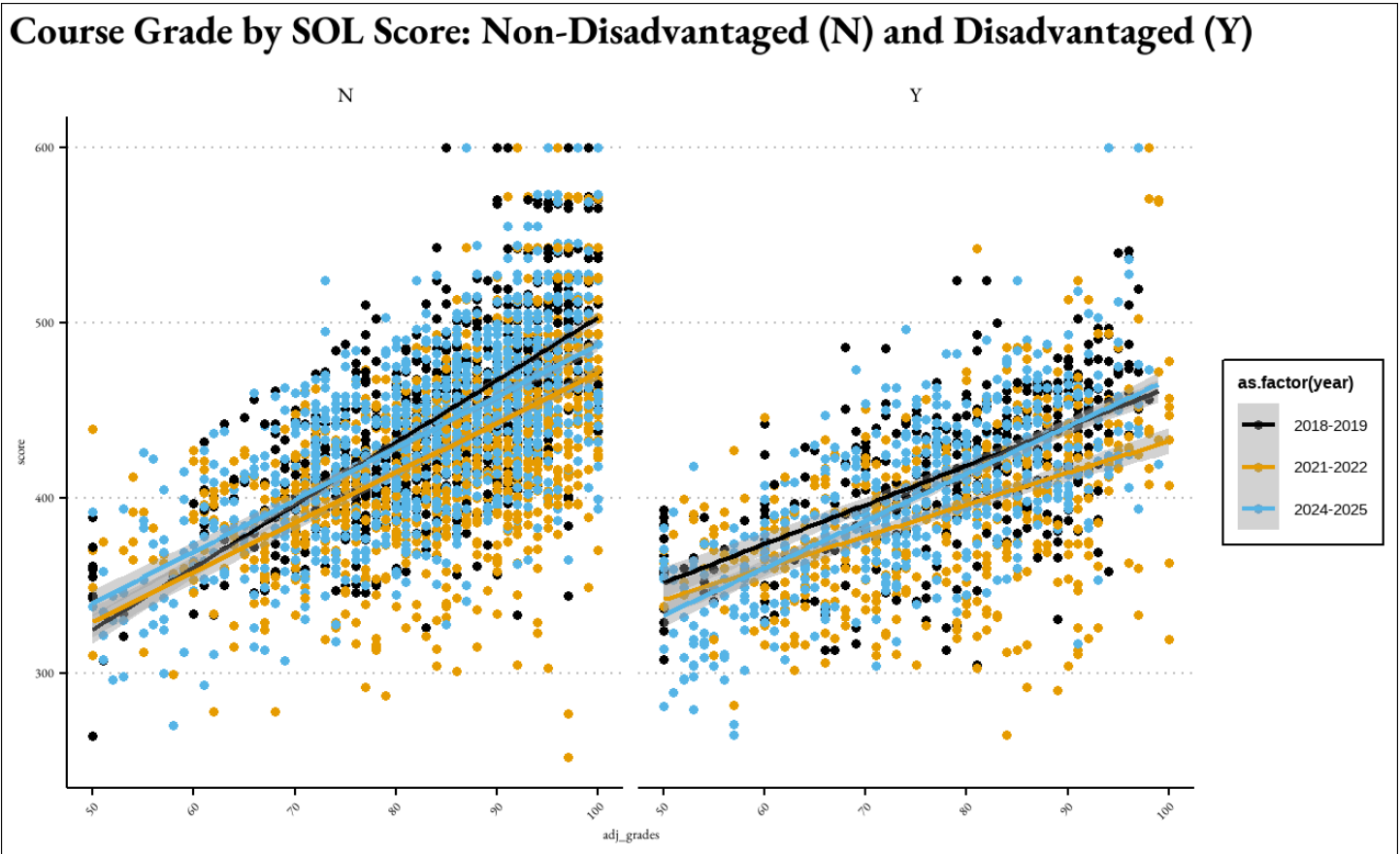


Figure 12. Course Grade by SOL Score

Figure 13 provides the results from the regression analysis on students grouped by previous achievement in mathematics. This analysis gives further evidence for the grading policy's differential impact and shows an even more extreme difference between historically higher achieving students compared to historically lower achieving students in the math dataset. Students at the lower distribution of achievement experienced a statistically significant improvement in the accuracy of their grades, yet students at the higher distribution of achievement experienced twice as much change in a decline of accuracy.

Grade Accuracy by Prior Achievement Quartiles (Tiers)

<i>Predictors</i>	Overall <i>Estimates</i>	Tier 1 (Lowest) <i>Estimates</i>	Tier 2 <i>Estimates</i>	Tier 3 <i>Estimates</i>	Tier 4 (Highest) <i>Estimates</i>
Baseline Intercept	0.24 ***	-0.46 ***	-0.08 *	0.16 ***	0.45 ***
Grade (Baseline 2018-19)	0.60 ***	0.27 ***	0.38 ***	0.41 ***	0.83 ***
year2021-2022	-0.30 ***	-0.25 ***	-0.11 *	-0.16 *	0.30
year2024-2025	-0.12 ***	0.07	-0.06	-0.04	0.10
econ_disY	-0.25 ***	-0.10 *	-0.10 *	-0.08	-0.25 **
eLY	-0.35 ***	-0.26 ***	0.09	-0.02	-0.05
swdY	-0.35 ***	-0.25 ***	-0.17 **	-0.13	0.03
std_grade:year2021-2022	0.06	-0.08	-0.09	0.09	-0.33
std_grade:year2024-2025	-0.03	0.15 ***	-0.01	-0.06	-0.30 ***
Observations	3534	884	884	883	883
R ² / R ² adjusted	0.488 / 0.487	0.330 / 0.324	0.241 / 0.234	0.216 / 0.209	0.274 / 0.268

* $p < 0.05$ ** $p < 0.01$ *** $p < 0.001$

Figure 13. Grade Accuracy by Prior Achievement Quartiles

Figure 14 shows that while grades have dropped for all students in secondary math, they have dropped more precipitously for historically lower achieving students.

Differential Impact Based on Historical Math Achievement

Proportion of letter grades awarded, grouped by students' historical achievement tiers

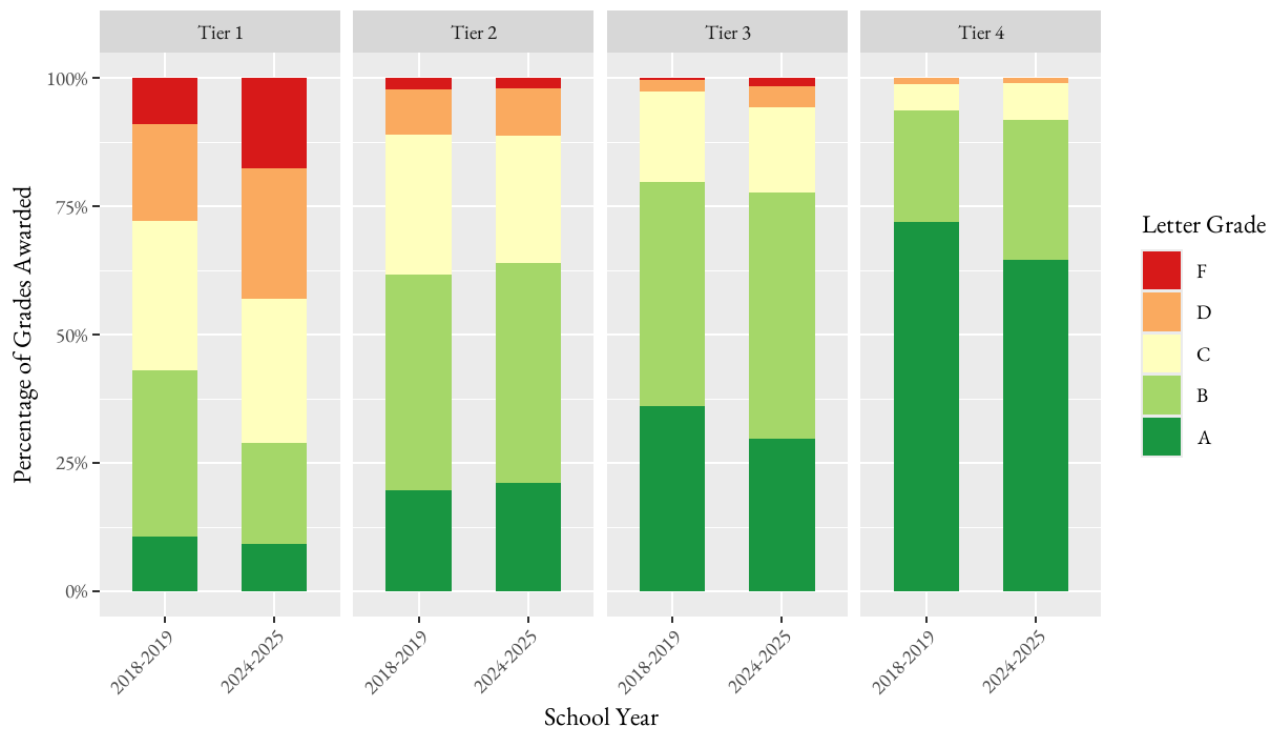


Figure 14. Differential Impact Based on Historical Math Achievement

Summary of Secondary Quantitative Grading Analysis

Both mathematics and reading datasets point to a differential impact of changes in grading policy and practice for students at the higher and lower ends of the achievement distribution. Grades for lower achieving students have become more valid measures of achievement, while grades for higher achieving students have become less valid.

Elementary Reading Grades and Achievement

Elementary report cards already used a standards based system in 2018-2019, recording grades as numbers between one and three for standards assessed in the grade level of the student. To form this dataset, we took all of the grades for reported reading standards for the year and averaged these into a composite reading score of 1-3, then compared this score to the students' achievement on the SOL assessment. The dataset included over 6,000 unique student records.

Descriptive statistics in Figure 15 show that both grades and achievement dropped substantially in elementary reading from 2018-2019 to 2024-2025. The proportion of students earning a “one” on reading grades more than doubled from around 4% in 2019 to 10% in 2025 and the average SOL score dropped from 441 to 429.

Elementary Reading Dataset			
Characteristic	2018-2019, N = 1,928 ^j	2021-2022, N = 1,971 ^j	2024-2025, N = 2,127 ^j
gender			
F	943 (49%)	950 (48%)	1,065 (50%)
M	985 (51%)	1,021 (52%)	1,062 (50%)
grade			
3	946 (49%)	1,015 (51%)	1,088 (51%)
5	982 (51%)	956 (49%)	1,039 (49%)
econ_dis			
N	1,479 (77%)	1,446 (73%)	1,554 (73%)
Y	449 (23%)	525 (27%)	573 (27%)
race_ethnicity			
AS	94 (4.9%)	107 (5.4%)	143 (6.7%)
BL	192 (10.0%)	231 (12%)	242 (11%)
HL	262 (14%)	330 (17%)	390 (18%)
MR	130 (6.7%)	171 (8.7%)	146 (6.9%)
WH	1,250 (65%)	1,132 (57%)	1,206 (57%)
swd			
N	1,711 (89%)	1,742 (88%)	1,839 (86%)
Y	217 (11%)	229 (12%)	288 (14%)
el			
N	1,719 (89%)	1,690 (86%)	1,837 (86%)
Y	209 (11%)	281 (14%)	290 (14%)
grade_alpha			
one	68 (3.5%)	62 (3.1%)	221 (10%)
three	1,157 (60%)	1,229 (62%)	1,213 (57%)
two	703 (36%)	680 (35%)	693 (33%)
grade_num			
	2.8 (0.4)	2.7 (0.4)	2.6 (0.6)
score			
	440.5 (80.2)	425.7 (66.1)	428.8 (73.7)

^j n (%); Mean (SD)

Figure 15. Elementary Reading Dataset

Regression analysis results in Figure 16 show that prior to 2020, with a grade intercept of .66, elementary reading grades were more closely aligned to achievement than secondary reading (.39) and math grades (.57). While secondary reading grades became more valid predictors of achievement after 2020, and secondary math grades maintained their alignment, elementary grades actually dropped significantly in their alignment to achievement. Unlike in the secondary datasets, this decoupling of grades from achievement did not have a differential impact across groups - rather it affected nearly all groups at a similar level. In short, elementary reading grades became less valid for all student groups after 2020. Since there is no indication of outright grade inflation, this decoupling is likely due to achievement dropping faster than grades since the pandemic era.

Elementary Reading: Relationship Between Grades and Achievement

Predictors	Overall	Econ Disadv	Non-Econ Disadv	ELs	Non-ELs	SWD	Non-SWD	BL	HL	WH	Male	Female
	Estimates	Estimates	Estimates	Estimates	Estimates	Estimates	Estimates	Estimates	Estimates	Estimates	Estimates	Estimates
Baseline Intercept	0.44 ***	-0.41 ***	0.12 ***	-0.46 ***	0.06 ***	-0.65 ***	0.05 **	-0.34 ***	-0.33 ***	0.12 ***	-0.04	0.08 **
Grade (Baseline 2018-19)	0.66 ***	0.59 ***	0.86 ***	0.58 ***	0.84 ***	0.41 ***	0.89 ***	0.58 ***	0.69 ***	0.85 ***	0.80 ***	0.88 ***
year2021-2022	-0.14 ***	-0.12 **	-0.18 ***	-0.16 *	-0.15 ***	0.05	-0.17 ***	-0.14 *	-0.21 **	-0.18 ***	-0.16 ***	-0.19 ***
year2024-2025	0.02	0.02	0.03	-0.17 *	0.03	0.38 ***	0.02	0.03	0.10	0.04	0.05	0.01
genderM	-0.08 ***											
econ_disY	-0.34 ***											
race_ethnicityBL	-0.46 ***											
race_ethnicityHL	-0.33 ***											
race_ethnicityMR	-0.25 ***											
race_ethnicityWH	-0.15 ***											
elY	-0.37 ***											
swdY	-0.38 ***											
std_grade:year2021-2022	-0.19 ***	-0.19 ***	-0.17 ***	-0.12	-0.21 ***	-0.04	-0.22 ***	-0.20 ***	-0.21 ***	-0.17 ***	-0.15 ***	-0.23 ***
std_grade:year2024-2025	-0.19 ***	-0.15 ***	-0.19 ***	-0.24 ***	-0.18 ***	0.06	-0.24 ***	-0.08	-0.19 ***	-0.19 ***	-0.17 ***	-0.24 ***
Observations	5903	1501	4402	726	5177	681	5222	650	944	3541	2993	2910
R ² / R ² adjusted	0.518 / 0.517	0.396 / 0.394	0.397 / 0.396	0.376 / 0.371	0.412 / 0.411	0.339 / 0.334	0.428 / 0.427	0.444 / 0.439	0.420 / 0.417	0.348 / 0.348	0.429 / 0.428	0.464 / 0.463

* p<0.05 ** p<0.01 *** p<0.001

Figure 16. Elementary Reading: Relationship Between Grades and Achievement

Figure 17 shows this decoupling of grades from achievement. With each year, the slope measuring the positive relationship between grades and SOL test scores declines.



Figure 17. Average Reading Standard Grade by SOL Score

Summary of Quantitative Analysis of Grades and Achievement

This quantitative analysis found that since 2019 achievement in elementary reading and math has dropped for nearly all groups and grades have also dropped during this time. For secondary reading, achievement and grades have remained relatively static. The relationship between grades and achievement has improved for lower-achieving students in secondary school, but grades have become less valid indicators of achievement for higher achieving students. For elementary school, grades have become less valid for all students. This provides evidence that the 2020 grading changes had the intended impact of improving validity for secondary students who historically have struggled academically, but had the opposite effect for more advantaged students. The changes have had either no impact or a negative impact on validity for elementary school students, as there is no evidence to support a positive impact.

V. TEACHER SURVEY

An anonymous survey on teachers' perspectives on the grading policy and handbook was sent to all teachers in ACPS via the division's monthly newsletter, school newsletters, direct email, and outreach through the local chapter of the Virginia Education Association. Below is a table that shows survey participation rates compared to the population of teachers in ACPS.

Table 1. Teacher Survey Demographics

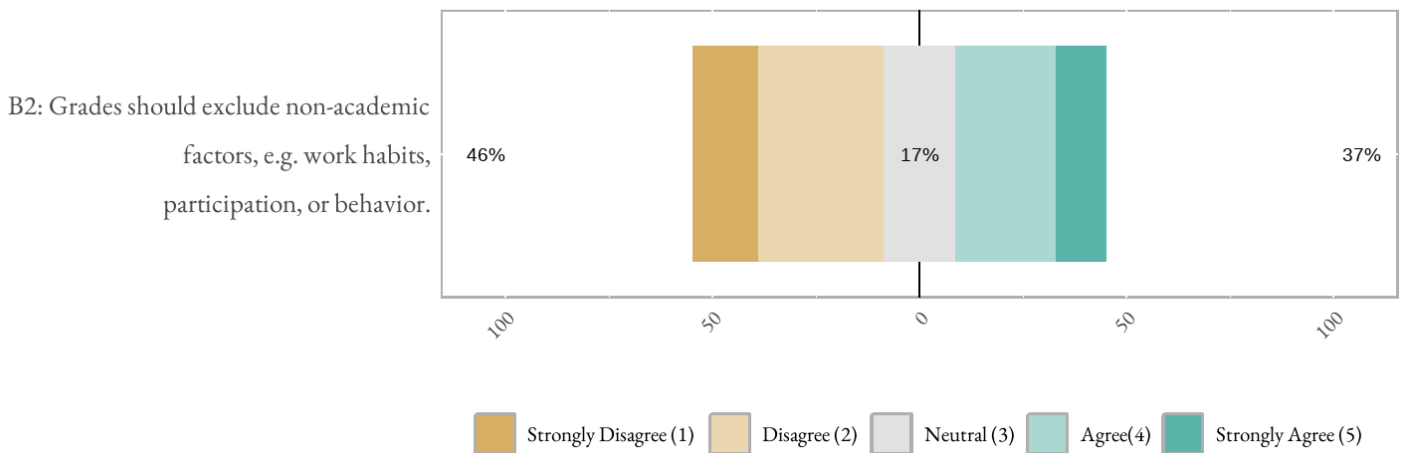
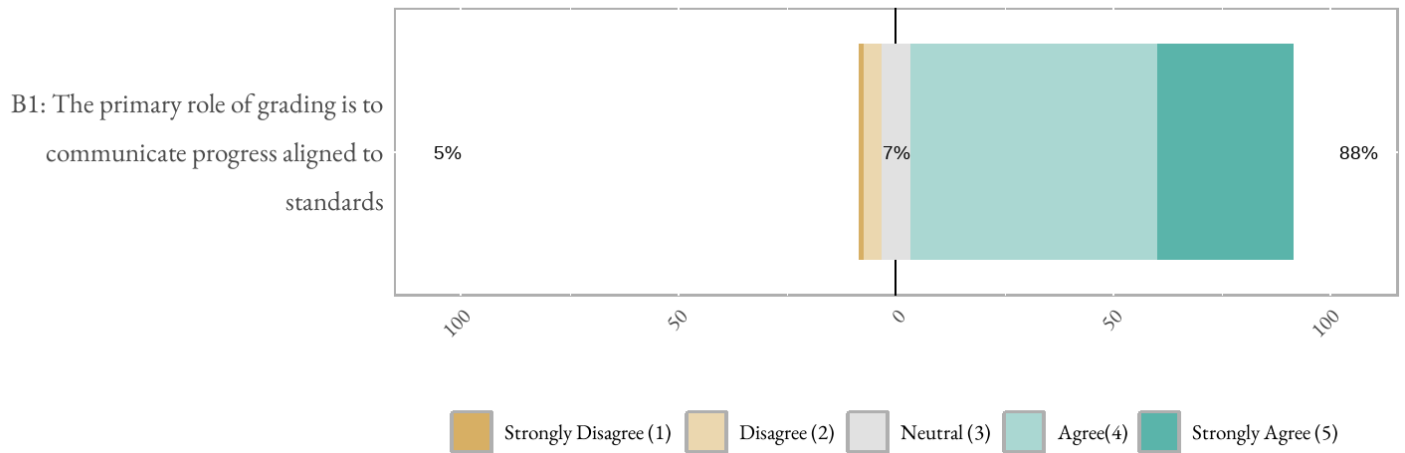
Level	Survey Sample	Population	Sample Percent of Population
Elementary	51	582	9%
Middle	140	255	55%
High	180	337	53%

Years Teaching	Survey Sample	Population	Sample Percent of Population
0-5	52	247	21%
6-10	71	223	32%
11-20	113	357	32%
21-30	112	278	40%
31+	24	69	35%

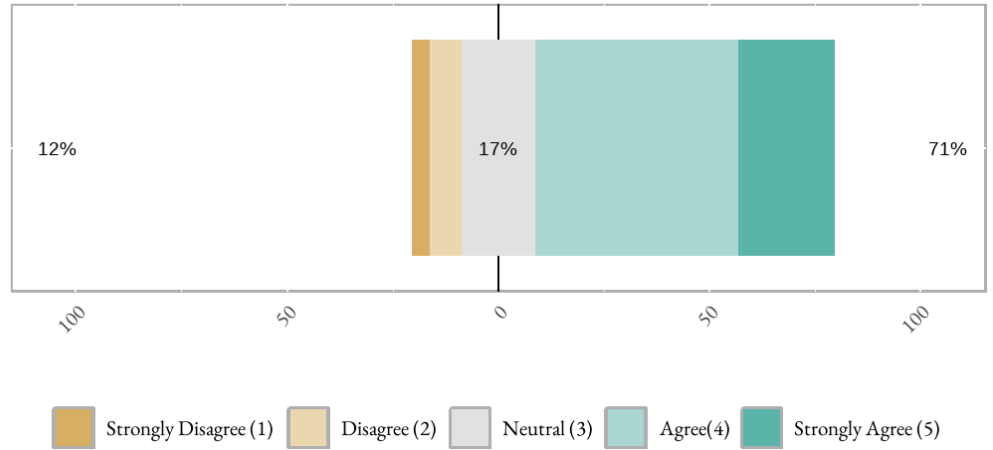
These rates allow us to share general sentiments among many teachers in ACPS; however, they should not be considered representative of all teachers as it is possible that the teachers who chose not to complete the survey have a different perspective than those who did, leading to selection bias. This is especially true for elementary school teachers who had the lowest rates of participation. The sampled population was on average more experienced than the overall population of teachers in ACPS. These limitations notwithstanding, over 50% participation for middle and high school teachers provides moderate confidence that the findings below represent the perspectives of a majority of teachers at the secondary level.

Teacher Beliefs about Grading

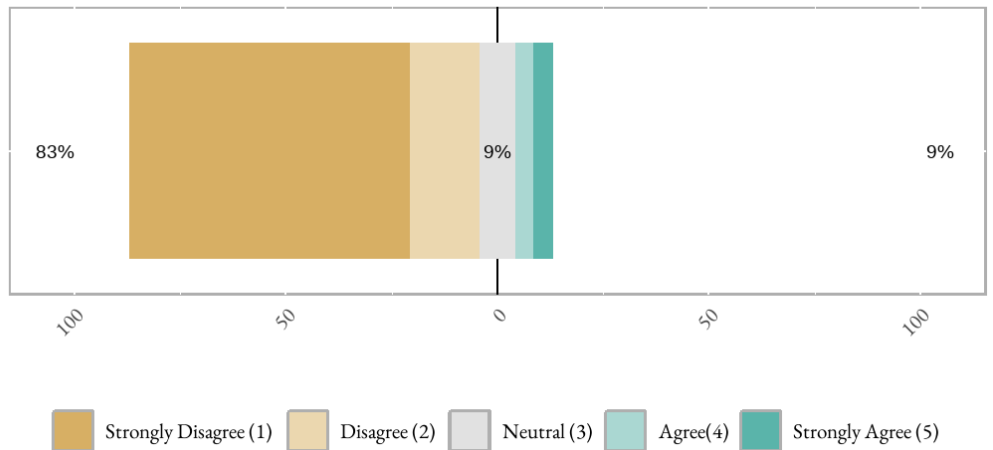
The first set of questions used a Likert scale to measure teachers’ beliefs about grading, focused on aspects of SBG and ways its principles have been implemented in ACPS. This set of questions provides a basis for summarizing results (see Figure 18). The full response data is included in Appendix A.



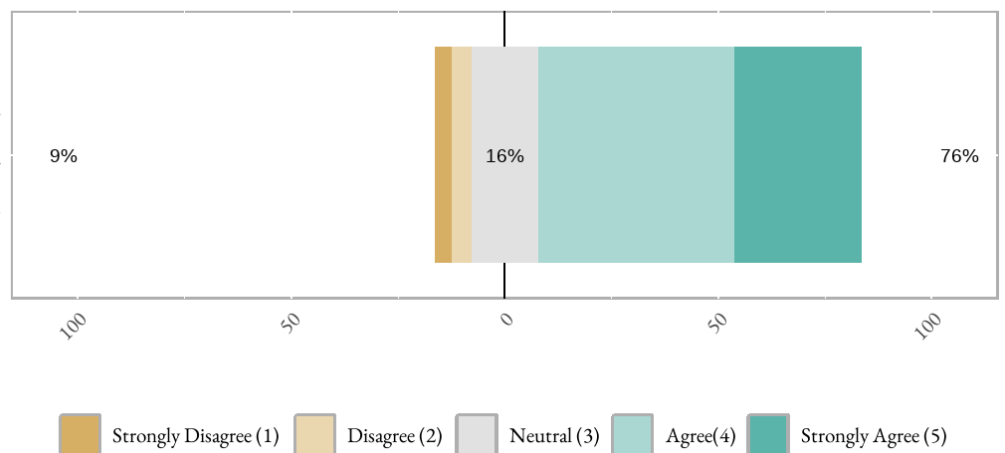
B3: Timely feedback and opportunities to redo work allow students to learn from their mistakes.



B4: Grading should be calculated using a 50% as the lowest grade.



B5: Different grading practices should be used in different developmental periods, e.g. elementary, middle, high



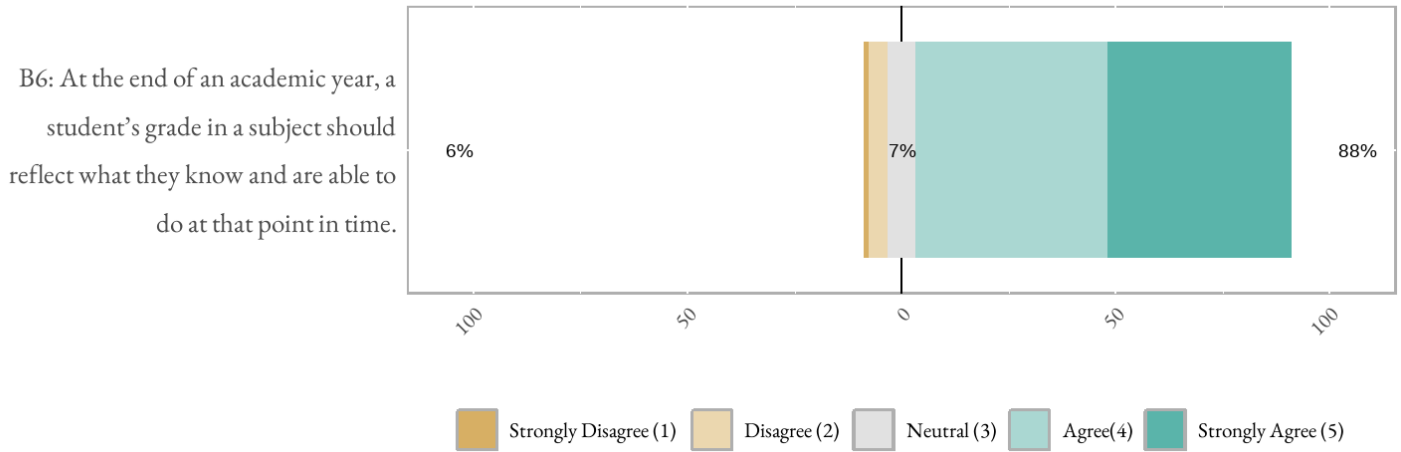


Figure 18. Teacher Survey Results

There is broad support for the principles of SBG among ACPS teachers; however, there is deep concern for some of the ways ACPS is attempting to implement these principles.

Nearly 90% of teachers support two of the three pillars of SBG: ensuring grades are an accurate reflection of what students know and can do in relation to the standards, and ensuring end-of-year grades are a reflection of what students know and can do at that point in time. Teachers' support for Guidelines 1, 2, 6, 7, and 8 highlight this alignment to the core principles of SBG, including greater weight on summative assessments and reduced weight on practice assignments as well as providing opportunities to show learning using multiple measures on standards throughout the school year.

Teachers expressed concern about the way that ACPS has attempted to implement other aspects of SBG, and indicated lack of support for Guidelines 3 and 4. The third pillar of SBG requires a limited set of grading categories, preferably four to seven, to improve reliability. ACPS attempted to move towards this by creating a grade floor of 50 for all assignments; however, this only truncated the existing scale without adopting a categorical system and eliminated the ability to differentiate numerically between work not attempted at all versus attempted work that was inadequate. Teachers strongly disagree with this guideline, with 83% expressing discontent. Teachers also expressed concern about not reporting on process skills such as work habits and effort, with only 37% in agreement, indicating a desire among teachers to provide feedback to students and families on work habit process skills in addition to academic outcomes. Teachers' support for retake guidelines was mixed, with 32% in disagreement and 44% in agreement, indicating a need to review further ways it is working and ways it may be improved.

While teachers support the principles of the policy and most of its guidelines, their concerns with Guidelines 3 and 4 were so strong that it led them to rate the overall policy as poor in its attempts to reach the goals of improving student learning, ensuring consistency in grades, and ensuring validity in grades' alignment to standards.

We now turn to a qualitative analysis of teachers' comments in the survey to gain a deeper understanding of their perspectives on policy and guidelines. We will focus on the guidelines that teachers had the most concern with to better understand how to improve.

A Grade Floor of 50

Among all new developments in grading, the grade floor of 50 was the one to which more teachers were opposed than any other, with 83% of teachers disagreeing with the practice, and a plurality strongly disagreeing. Qualitative data reveals the reasons for teachers feeling this way, the negative impact they believe the practice has on students, and possible improvements.

Differentiation between No Work and Sub-Standard Work

Teachers felt it was unfair to measure the work of students who did not turn in work at all equally with a student who tried an assignment but provided sub-standard work. Teachers wanted a way to clearly communicate through grading to parents and students the difference between unattempted assignments and submitted ones.

“Having a minimum 50 grade FOR WORK THAT HAS BEEN COMPLETED helps those students who really struggle. However, when a student doesn't turn in ANY work for a project or paper - the 50% does not represent a student's understanding. We have no idea what the student knows.”

“Students should not receive a minimum grade of 50% for not submitting anything. I believe that if they legitimately try on every question, then the minimum grade can be a 50%. However, students are refusing to take tests or only answer one question because they know they will get a 50%.”

“The grading policy is inherently inconsistent; a student who does absolutely nothing and isn't even present is given the same amount of credit as a student who is present and able to demonstrate [below standard] understanding of material.”

Impact on Motivation

Teachers cited many examples of ways the grade floor of 50 was detrimental to students' motivation.

“Sometimes they don't turn an assignment in at all, choosing to see how the 50 affects their grade so they can choose whether or not to do the assignment. This creates even larger gaps in their learning (especially when considering reading/writing instruction).”

“While I agree that missing work should not receive a grade of 0% because it pulls a student’s overall average too low; I also believe that the 50% for missing work is too high of a grade. A student can miss a lot of work and know that the 50% rule will allow them to pass that class.”

Mathematical Distortion

Some teachers pointed out that truncating the 100 point scale to a 50 point scale without also changing the translation to letter grade categories that determine a minimum standard for passing effectively created a new categorical system that lowered the passing standard.

“Giving a 50% for missing assignments combined with the barrier for passing being 60% is highly distorting and leads to students passing while completing barely any work...theoretically, a student could earn a 70% or higher in the first semester, then complete zero assignments in the second semester, and still end up at a [passing] 60%.”

“The minimum grade of 50 creates a 6 point system when translated into letter grades. A 7 point system would allow for a distinction between an assignment that is not turned in at all and one that is turned in but is far below the requirements.”

“While a 50% is appropriate for no work turned in, you have to move the passing rate to proficient. To keep the passing rate at 60% does not show or promote learning. You could lower the zero grade to 30% and have passing at 60%, but you can’t have a zero grade at 50% and passing at 60%.”

Possible Solutions

The most common solution offered to these problems was reforming the system to distinguish between work turned in, which could receive a minimum grade of 50%, versus work that is not attempted, which could earn a lower grade. Many teachers wanted to use a 0 for work that was not submitted, while many others suggested a balanced approach by using a lower grade than 50, yet was still above zero, to preserve a system more closely aligned to SBG that did not hamper motivation by assigning too heavy a negative weighting for missing assignments.

“I do not agree with a minimum of 50%, but I do think there should be a minimum amount to not tank the grade such as 40%.”

“Re-evaluate the grading scale. Zero’s aren’t the answer, but 50 is also not communicating progress well either.”

“Reform the 50% floor: to receive a 50%, the student needs to have at least attempted the work (50% can be the lowest a completed assignment can earn.”

“Maybe the solution is assignments that are completely missing get something closer to a 0, or that a certain number of missing summative assignments mean that a student cannot pass the class.”

Retakes

While 71% of teachers agreed that allowing students to redo work after feedback was beneficial to learning, and 88% agreed that end-of-year grades should be reflective of what students know and are able to do at that point in time, they had many concerns around a lack of strong standards in retake procedures to determine when retakes are appropriate. This lack of clarity, in their view, leads to detrimental impacts on student motivation and an undue burden on teacher work time that negatively impacts their ability to provide quality feedback to the students who need it the most.

Support for the Principle

Many teachers support the concepts of mastery learning: allowing students to learn from mistakes and showing their continued learning on standards through revisiting material and demonstrating understanding in new ways.

“I think students should revise work and have opportunities to show mastery.”

“I agree with the redo policy to a certain extent. Encouraging students to revisit material until they master it fits more with brain development than expecting everyone to get it at the same pace.”

“I support allowing students to learn from their mistakes and redo work.”

“Students should absolutely have the opportunity to retake tests or redo high stakes assignments to demonstrate their learning. Their progress should be accurately captured by a grade, and the grade should not just reflect moments in time.”

Concern with Implementation

Teachers expressed concern that having retakes on every assessment without limits led to reduced motivation for students to prepare for the assessment on the first attempt and an untenable workload for staff.

“I have observed that many students use this opportunity strategically, waiting to see their initial grade before deciding to put in effort. Instead of preparing from the beginning, some students focus only on correcting specific mistakes to improve their score, without developing a full understanding of the topic. This limits deep learning and affects academic responsibility.”

“Many students do not prepare for the first assessment and wait to see what material is covered and then use the retest which leads to more stress for students and more work for teachers.”

“The constant redos that are allowed are crushing teachers and building poor habits among students. Kids are learning to just wait for the redo and as a result every summative assignment gets dragged on forever with multiple different timelines as kids redo work and get grades back.”

“Automatic redos are untenable for teachers. Grading is neverending and it is impossible to give full attention to the current unit. Students abuse the system by not preparing for the first assessment or assignment.”

“I can't focus on the kids who truly need 1-1 because there are so many kids trying to get above an 89. Kids are routinely playing catch-up for the retake, and can't focus on the present.”

Possible Solutions

Teachers offered a variety of solutions to improve the retake system while preserving its benefits: limiting retakes to certain students based on academic performance, limiting the number of retakes a student can have on a particular assessment, placing stricter emphasis on the requirement to complete remediation or study prior to attempting a retake, and moving to a true SBG system in which standards spiral throughout the curriculum and teachers can adjust previous grades based on mastery shown on later assessments.

“While I support retakes, there needs to be a time and quantity limit. There also need to be requirements in order to retake - students engaging in remediation or a study guide for example.”

“All work must be submitted on time for an opportunity to retest. [ACPS] should create retests and provide material to complete prior to offering a retest. It would also make sense to incorporate into the policy that a retake (by virtue of the benefit afforded by a second go at the skill) automatically be capped at a 90% . 100% can only be earned if a student is successful on the attempt when assessment is initially administered.”

“I would like to see more parameters on guideline #5 to prevent parents from demanding unlimited test retakes indefinitely. It's not most parents, but there are certainly a few who will bring a highlighted policy with demands.”

“It would be useful to establish clearer criteria for the reassessment (retake) process, including limits on the maximum achievable grade.”

“Teachers are burning out with the number of retakes they are being asked to provide students. Students should be provided only one opportunity to retake after reteaching.”

“I have done retakes and remediation to the best of my ability, and it's my opinion that the time and effort would far better be spent on remediation and spiral back review alone. Students already have opportunity to show improvement through future assessments, and continuously going back in the curriculum while others

move forward only widens the gap and creates stress for students who struggle more. It also de-motivates some students to try their best on assessments, since there will ‘always be a retake,’ and reality hits hard when they go to college or the workplace.”

Including Process Grades

A plurality (47%) of teachers believed that work habits should *not* be excluded from the grading system, compared to 36% who agreed. Teachers who supported grading work habits believe they are important markers of student success that require feedback and communication to students and families.

“Turning in work on time is a life skill students need in order to be successful in college and as a professional in the world of work. We are not properly preparing kids to be successful adults.”

“I think retakes and flexibility have their place, but I also believe that academic behaviors are just as important to evaluate and refine.”

“Work ethic and participation is a key component in the workplace and that has not been graded or demonstrated in this grading policy.”

“Academic content and skills are important, but study habits are important too. The grading policy should be revised to enable teachers to instill strong habits in their students.”

“We have put a policy in place that does not translate past high school. I have heard from my own past students as well as children of colleagues who have gone off to college, and have fallen flat on their face because deadlines are real and there is no second chance on the work they do there.”

Potential Solutions

The most common solution teachers provided was for grades to be assigned for workplace readiness skills such as timeliness and work ethic. Many teachers suggested these could be kept separate from GPA, but could be reported in the gradebook as a clear evaluation for students and parents of work habits important for success beyond K-12 education. Many teachers also suggested assigning reasonable penalties to late work and clear deadlines after which late work would no longer be accepted.

“There should be a work ethic grade for each class. It would not count towards the GPA.”

“Allowing some behavior-based information into the gradebook, even if it is not counted in the final grade, so parents can easily see it on PowerSchool. Student behavior -- being late, disruptive, not participating -- does affect their learning as well as the learning of other students and is absolutely relevant.”

“We need to start directly teaching the [workplace readiness] objectives and giving a separate grade on behavior, work ethic and/or executive functioning skills so they can start to learn them before they get to the workplace and need them but don’t have them.”

“I think it would help student engagement if credit were able to be deducted from late work. I accept late assignments, however, I feel it does not support student learning when they abuse the policy and turn it all in at the last minute just to get credit and it is not helping them reinforce concepts learned during class.”

“[ACPS] should have specific timelines for late work completion to be counted.”

Elementary Grading Scale

Participation rates among elementary teachers were low, however, every comment received from elementary teachers about grading requested improvement to the current grading scale of 1-3. Most of these teachers advocated for a four point scale instead.

“I do not feel that the current 1, 2, 3 grading scale in elementary school accurately communicates student knowledge and skills. It often does not provide enough specificity to reflect where a student truly is in their learning. I would prefer a more detailed scale, such as a 1 to 4 system or an A through F scale, because those options provide clearer distinctions in student performance and better communicate mastery of standards to students and families.”

“I think having a greater scale (4 instead of 3) would be able to more accurately communicate a student’s academic level- a 1 is below grade level expectations, a 2 is approaching, a 3 is meeting, and a 4 is exceeding.”

“The standards that are used in our curriculum are different than what we see in the gradebook in ClassLink. If the standards can be changed so that they are accurate and reflect what is actually being taught in class that would make our lives easier when grading and inputting data.”

Summary of Findings from Teacher Survey

There is broad support among ACPS teachers for the core beliefs of SBG; however there are deep concerns about how these have been implemented in the 2020 grading policy and guidelines. Teachers believe that the grade floor of 50, a lack of parameters governing retakes, and the exclusion of grades on work habits have reduced motivation among students and created a heavier workload on staff.

Teachers offered a variety of solutions to the problems identified. A solution proposed for Guideline 3 is assigning grades lower than 50 for work that is not attempted, but keeping the 50 floor for submitted work. Solutions for

Guideline 5 include limiting retakes to students who need them the most, capping the grade for which a retake counts, limiting students to one retake per assessment, and requiring remediation or study prior to a retake. Solutions for re-incorporating grades on work habits include providing separate grades specifically on work-readiness skills that do not count towards the GPA, and allowing for uniform and reasonable penalties and deadlines for accepting late work.

VI. STUDENT FOCUS GROUPS

Three focus groups were conducted with high school students in ACPS which included the perspectives of 47 students. Students were recruited from 11th and 12th graders in the ACPS student senate, 9th graders in the Starr Hill Pathways program that supports predominantly prospective first-generation college students, and 10th graders in Monticello's Early College Scholars Program. Group conversations were recorded and transcripts were used for thematic coding analysis. Themes that arose among students were similar to those that arose among teachers: Retakes, the grade floor of 50, and late work; however, students also emphasized the importance of receiving regular feedback from all teachers.

Retakes

Overall, students were supportive of the retake policy; however they also noted that it could lead to a lack of motivation in some circumstances and suggested adjustments to counter this. Students spoke particularly about the impact on mental health of knowing that they could revisit high-stakes assessments if they performed particularly poorly, for example, but also noted that if retakes were too freely available, it lessened their motivation to study on the first attempt.

Support for Retakes

Students noted that often they do not have time to study adequately for all of the tests they may have, and the opportunity to retake gives them greater flexibility in managing their time. This flexibility reduces stress and testing anxiety.

“I definitely feel more relaxed knowing that I can do retakes because sometimes I don't have time to study, or, like, I don't know, I just don't get around to it, and then I get a bad grade on a test, and I'm really bummed out. And just knowing that I can redo it and get a better grade, it's really comforting. You feel more relaxed about it, because if I get that grade, I can retake it and maybe get a higher one or try better.”

“Sometimes teachers don't notice that we have other classes that can be harder. We have to study a little bit more for another class and sometimes we can't really find the time to study for this specific test that's coming up. Having retakes relaxes us a little bit more knowing that, okay, I'll try my best on this first test, but if I don't pass, I have the second opportunity of taking the retake, and I could pass, because I have more time to study now that I have my other classes done with.”

Negative Impact on Motivation

While students appreciated the flexibility and stress reduction provided by retakes, they also noted that having access to retakes without limits negatively affected their motivation to prepare for assessments.

“I seen people who don't study for the test, and then they just fail, but they know there's gonna be a retake, so instead of studying for the first test, they just study for the second.”

“The guideline is not as positive as it could be, because there's gonna be a habit that you're not gonna study because you just don't want to do it. And you know you're gonna keep taking retakes every single time because you don't want to study the first time.”

“I think part of the thing with the retake is, if there's students who are not studying for the test, and then they get a bad grade, but they're okay with that, because they know they can retake it, I think that's not necessarily helping their learning.”

“In the majority of classes that I see kids will definitely just be like, they're not showing up for the first version, or just, like, not try and then take the retake. Or sometimes, if it's a two part test, kids may spend the whole class just in the first part, and then when they retake, they'll just retake the second so they get double the time.”

The Importance of Remediation and Feedback

Students noted that remediation and feedback prior to attempting a retake was vital for their learning. They noted inconsistency in retake practices across teachers, and held up teachers who provided detailed feedback prior to a retake as a gold standard for supporting learning. Some students suggested that participating in remediation be required prior to any student attempting a retake, as otherwise retakes seemed counterproductive to learning. Students also pointed out that it was unfair to have a high stakes summative assessment if they hadn't yet had a formative assessment and received feedback on it.

“I'm not studying [for assessments] because I don't know how the hell I'm doing in this class. I don't have no quizzes. I get no feedback when we do take quizzes. And so yeah, I'm going to roll the dice. I'll go in and take the test, and if I get a good enough grade, then I'm moving on. But I don't know how I'm doing in here because I haven't had enough assessments that are just formative, that give me feedback on how I'm doing. It's not that the idea of allowing students to revise their work is a bad idea. It's that something else is going on that's really the issue behind the issue.”

“It can be difficult with some teachers. They'll say, you failed, give you the test back, and they won't really explain what you did wrong and how you fix it. They just say you can retake it. But, like my eighth grade teacher, she did retakes really well. She would have like, a couple versions of the test. She would give you, like

your same test, but you redo the questions you got wrong, she helps you fix it, and then you get another test, like similar questions, but different.”

“I think for people who are actually going to try the retake, you can maybe think more about, okay, what did I not feel like I understood? And then you can go focus on that, rather than studying a broad category of things for a quiz or test.”

“There’s multiple ways of assessing the student. Teachers have to work with individual kids a little bit. Now, I know that’s challenging if you have a lot of students, but if you do the practices up front, like formative assessments, things like that, then it really cuts into the students saying, why you need to study. If I don’t know how I’m doing, I don’t know what to study, so I’ll just take the test and see what happens, and then I’ll take it afterwards. In other words, I’m making my own formative assessment out of the test.”

“I think students should be required to do more of the work upfront to be eligible for the retake.”

Grade Floor of 50

Students were divided on the issue of the grade floor of 50. Some students expressed support for the practice as a 0 seemed a disproportionately heavy penalty for not turning in an assignment. There was also support for the 50 leading to continued engagement in school after a poor grade. Other students felt it was unfair for students who tried and were unsuccessful to get the same grade as a student who did not try at all, and they also noted it could have a negative influence on motivation since students did not need to try on as many assignments.

Support for the Grade Floor

Students who supported the 50% floor cited its potential to keep a student engaged even after getting a bad grade on an assignment or test, and its disproportionate impact on final grades.

“I’d say the 50% is good. If you get a zero on something, or like a 12 or 16 on a test, even if you were to improve and get A’s the rest of the year, that could potentially not allow you to get your grade back up to where you need to. It’s very helpful for students who have a lot going on, and sometimes don’t have as much time as they need.”

“Zeros don’t motivate students because when a student just sees they got a zero, they’re like, ‘oh, there’s no way I’m going to come back from that. So why bother trying?’ So I do agree with the 50. It’s an opportunity to be like, hey, I can fix this. I can get it back up.”

Negative Impacts on Motivation and Differentiating Attempted Assignments from those not Submitted

Students who opposed the grade floor of 50 noted that it could give them a false sense of security with regard to what they know in the class. They also felt it was unfair to assign the same grades to students who try and do not understand and to students who do not try at all. While they supported a 50 for attempted work, they thought it was appropriate to assign a zero to work that is not attempted.

“If I see a 50, I’m like, ‘Oh, everything will be fine. It’s gonna be easy to get back up. It’s not really gonna matter.’ But if I see a zero, I’m like, ‘Okay, I’m gonna need to fix that.’”

“I feel like if you didn’t turn it in, that’s fine to have a zero because you didn’t try. But if you did turn it in and you tried, but you still don’t get a good grade, I feel like you should be able to stay in that range of 50.”

“If I turn in an assignment and I do really bad, and I get a 50, and then someone else just, literally just doesn’t do it, and then we both get a 50, it’s kind of unfair, because I still did the work, even though it was bad and they didn’t do anything, and we share the same grade. But you would think it would make more sense for them to get a zero”

Late Work

Students generally had agreement on a desire for greater consistency among teachers in late work policies, and shared a belief that students should be able to turn work in after a deadline within reason with a grace period of between three days to two weeks. Most students believed it was fair for teachers to apply a penalty for work turned in late as long as it was not so stark as to make completing the work no longer worthwhile.

“Imagine you have a project that’s due tomorrow, but you just get a new assignment that you also have to work on. You’re gonna want to work on the one that’s more important instead of something less that you can wait a little bit on. I think a two week deadline for accepting late work is good, because some people could be buried in work from just one senior class and have to get that done before you have time for a new assignment.”

“I feel like the late policy shouldn’t be too strict, because some students genuinely don’t have time, especially if they’re a student athlete, because some students are, like, going to practice or games and they don’t get home until 11, so they should at least have a two, or three day grace period for that. It feels like sometimes teachers just completely forgot what it was like to be in high school because you have some teachers that have no grace period. Like, you don’t get it on time, that’s it. But then you have teachers who will let you pick something up from three months before and turn it in late.”

“I feel like having some more time after a deadline is fair, but it should bring the grade down a little bit more if you turn it in late. Because I feel like it just gives me more time to work on it, and it gives me more time to focus on a little bit more important things that are due.”

Summary of Student Perspectives on Grading

Students noted first the great inconsistency in application of the grading guidelines across their classrooms and expressed a desire for greater consistency. They had general agreement that retakes can be either beneficial or detrimental to learning depending on how they are applied. Retakes that are provided as part of an assessment system that includes formative and summative assessments with specific, targeted feedback were seen as beneficial to learning; however, retakes that were provided without any remediation and without limits were seen as detrimental to both learning and motivation. Students valued retakes as a way to improve their learning but also supported some limitations on them. They also wanted a requirement for formative assessments prior to summative assessments with regular teacher feedback between assessments and prior to retakes as well. Students were divided on the grade floor of 50 percent, citing both potential for both positive and negative impacts on motivation. Most students supported some way of differentiating between attempted assignments and those not attempted at all. Finally, students supported a grace period after deadlines to turn in work; however, they also saw fairness in putting a limit on that grace period and applying reasonable penalties for work turned in late.

VII. PARENT FOCUS GROUPS

Three focus groups were conducted with parents of students in ACPS, which included the perspectives of 28 parents with students in elementary, middle, and high schools. Some also had children in college who had graduated from ACPS. Students were recruited among Parent Teacher Organization (PTO) members across schools and from the ACPS Family Council. Conversations were recorded and transcripts were used for thematic coding analysis. Parents held similar perspectives to teachers on the overall principles of SBG, retakes, the grade floor of 50, and turning in late work. Parents had additional concerns about clarity, communication, and consistency in grading practices and shared similar concerns as students around the necessity for high quality feedback prior to summative assessments.

Retakes

While most parents expressed some support for the general concept of SBG that allows for students to learn from their mistakes and for grades to reflect growth over time, they had significant concerns about reassessment practices that reduced motivation and allowed retakes without substantial feedback or ability for parents to help students understand concepts on which they were not demonstrating mastery.

Support for Retakes, but Concerns about Teacher Burden

Parents expressed some support for retake procedures that included a clear process for students to receive targeted feedback on their mistakes and demonstrate learning through revision and remediation; however, they also noted the burden this put on teachers.

“The process my child’s teachers have used is if they want to retake a quiz or test, they have to do a retake ticket first, so they have to go back through the test and correct everything that they missed. And then they have to turn that in to meet with the teacher, and then they can retake the test.”

“Last year, my child was having trouble with fractions, they took a test, they got like a 40 on the test. So if it’s to even the playing field, I feel like maybe there needs to be more time for revising and feedback, like the revision piece should be given weight, or the remediation - the feedback time is important.”

“There’s no time, there’s no space in the schedule to accommodate all these kids for feedback and retakes. I think those teachers are stressed.”

Lack of Consistency and Negative Impact on Motivation

Parents’ concerns focused on a lack of consistency in retake practices, a need for more limitations on retakes, and the detrimental impact of unlimited retakes on student motivation.

“In the case of revisions, the policy is so broad that it’s going to be a classroom by classroom decision, of what that means. Some classes let you take multiple retakes and other don’t, and so it just feels like there needs to be more specificity around how you get a retake. Maybe it doesn’t have to be for full credit for the original assignment.”

“My daughter, who’s an incredibly studious kid, she’ll oftentimes get an A, but she might get a 92 and then she’ll see someone who takes a retake, and they might get 100, so she would be bothered that she wouldn’t be able to aim for 100 also.”

“For [my child] the regrading has created an attitude that he has that, like, the first time doesn’t count, and so it’s doubled the amount of work that he has to do, despite knowing that doing it twice is going to be twice as much work. And it’s doubled the amount of work that his teacher has to do.”

“I like the idea that you can try again, but I think there needs to be more structure around how that try again process works, and maybe you try again for 80% or you know, like, there’s got to be something that happens that means there’s incentive to do it right the first time.”

The Importance of Timely Feedback and Multiple Assessment Modalities

Similar to students, parents also emphasized the importance of feedback in the grading process, and expressed concern when students were provided with retakes without first getting feedback from a teacher. Parents would like to get paper copies of tests back when possible, or clear unit study guides that indicate the standards their children need to understand, so they can see for themselves in which areas their children are struggling, and can work in partnership with teachers to support student learning. Many families appreciated the part of the grading guidelines that call for multiple assessments of the same standard using different modalities and updating student progress on those standards, in place of doing retakes.

“My daughter has complained that oftentimes students get their numeric test score, but they don't get the test back to see what they've missed. This is because some teachers do not return tests until all retakes have been completed. So it might be six weeks before the test papers get handed back. My daughter may score high enough that she doesn't need a retake but she still missed a few questions. She finds it frustrating that she can't see which problems she missed or what mistakes to learn from until long after the test.”

“It's hard for parents to help their students because we don't have the tests, we don't have the quizzes to go through the ones that they got wrong. I think that's a real big problem.”

“In terms of having a long term plan, I like the idea that we're going to keep building towards a place where there are more varied requirements of assessments that take place for each class, and no retakes. So I think the answer is yes, we should do something that actually starts to restrict the retake policy, and we should be doing that with the idea that our long-term goal is to figure out how to make sure that the assessments that go into each course are varied and multiple over the course for each standard.”

Grade Floor of 50

Parents were divided on the grade floor of 50, with some expressing an understanding for it as a mechanism to prevent students falling so far behind that they give up. Many others noted that it creates disincentives for full engagement since students can choose not to do many assignments and still pass a class. There was consensus around a need for providing limits to the 50% floor, with nearly all parents supporting a lower grade for work that is not turned in and distinguishing this from work that is attempted.

“A zero will kill you, and then once you have a student who's gotten two zeros, because they were on the football team and they had two away games, and their life went sideways, and it's September, and then they're like, 'I'm never digging myself out of this hole, so why try?' We can't create a 'why try?' situation for kids, so I do feel like that 50% is the thing that keeps kids from falling into the hole they can't get out of, because a zero, a zero undermines your ability to progress. Maybe there's a massaging in here of ways to provide accountability, but also you can't make it so that they can't help themselves if they have a bad two weeks spell.”

“I’m fine if we’re going to say you get a 50, absolutely if there’s a kid who’s, like, really trying in math, but she’s just not getting it even though she’s trying so hard, and she gets a 50. But someone else who just doesn’t do it should get a lower grade.”

Grading Work Habits

Parents expressed strong desire to get feedback on their child’s work habits, either as part of their overall grade or as separate grades specifically on work skills. As part of this, parents also advocated for tighter regulations and consistency on deadlines for turning in late work - while there was support for the ability to turn work in late, they felt deadlines and late penalties were appropriate ways to reinforce work habits in grading. Finally, grading for homework also arose as a theme related to work habits, with the majority of parents advocating for homework to be graded for accuracy rather than completion and to count for more than the current 10%, at least in all high school courses or in courses that receive a one point boost in GPA quality points, e.g. honors, AP, IB, and DE.

“We have to make sure we’re giving feedback on work habits: you’re not turning in the work, you’re not showing the effort. What would guidelines look like to support that? Should we be evaluating and helping with those behavior type things, too? I think just as you’re looking at your feedback for academic skills, there should be feedback on timeliness.”

“The work habits matter. If you show up at work and you don’t know how to sequence your events throughout the day, you’re not meeting your expectations. It’s a disservice for the older high schoolers to send them out into the world with these expectations that work habit and behavior shouldn’t be a part of how you’re assessed, it’s how the world assesses you.”

“As someone who works in higher education, I think [the lack of deadlines] is an extreme disservice to students, because I see this frequently... you have so many students that burn out their first semester because they failed so many classes, because they take this high school idea of I can submit it whenever I want. I got an email yesterday from a student that said, ‘How can I make up my grades?’ when I had submitted grades two weeks ago. You can’t at that point.”

“Perhaps instead of a zero, they can turn in work late. It just is not for full credit, maybe for every day that it’s late, it’s 10% you know? If you got 100% on the assignment, it’s three days late, the highest you can get is a 70%.”

“I think that if you’re going to turn in work late, there should be some penalty for it. I mean, I think you should give kids the opportunity to do the work, you know, like, but they won’t get full credit for it.”

“There’s got to be something that happens in there that means that you do it right the first time, and so this kind of goes along with you get 100 if you do your homework, but you got to do it right the first time.”

“Something that I learned teachers were doing this year is when I would see 100 on homework and then a D on a test and I’m like ‘what’s happened, they were doing great on the homework?’, and then the teacher’s like, ‘well, that’s just to show that they did it, not that they did it right.’ And so I feel like built into this whole theory of grading was this idea that kids are going to be getting feedback and learning and going back and forth, but that doesn’t appear to be how it’s being implemented.”

“The maximum amount of their grade from homework is 10% and I’ve seen it in my own two middle schoolers, they’re really not that motivated to do homework, because if they don’t do it, it doesn’t really cost them that much. I had my daughter tell me the other day, I have one teacher who just grades on completeness, so even if she were to get 20 questions on homework and she were to get all of them wrong, she still gets 100 on that. It really creates perverse incentives, and demotivates students to put much emphasis on homework...when you’re not giving credit or giving very minimal credit for homework, students are going to game the system inherently, they’re smart kids, they know like it’s not worth my time to do this, it’s not worth my time to do that, and then there’s a lesson in there that they’re missing, about like, for when they go beyond high school, so like maybe it’s helping like even the playing field in high school, but is that really helping them for the next level?”

Minimum Number of Assignments

In every focus group, some parents brought up the issue of instances in which students did not have enough grades in a marking period for adequate assessment. They advocated for a required minimum number of grading assignments in a marking period.

“My daughter has one class, this entire quarter, there are no grades entered, either in Power School or Schoology, so I have no idea how she’s doing in that class.”

“My child had a class this semester and there was just one test for the whole semester. That doesn’t allow for feedback and learning.”

“One of my daughter’s teachers only gives three to four quizzes/tests each quarter, and one of them is heavily weighted so it can devastate a student’s grade if they don’t do well. I think there needs to be more diversified ways for students to show they know/understand the material. A combination of quizzes, tests, projects, assignments, classwork, etc.”

“One solution that may check a few boxes could be to have regular/weekly quizzes to check student’s knowledge before a test. This would change the low number of graded items recorded in each semester and additionally would prepare students for the test. Students and teachers can see what content is being comprehended correctly throughout the semester, not waiting for a periodic test that may need many retakes.”

Elementary Grades

Parents of elementary grade students focused on the difficulty in interpreting the three point grading scale and translating it to the other assessment data they received from schools. They had general consensus around a preference for a four-point categorical system and advocated for translating all assessment data into this scale for clarity. These parents felt the current policy was tailored mostly to secondary students, and wanted more guidelines specific to elementary schools. Finally, elementary parents also requested receiving more graded work to come home on paper, so they can see for themselves how they can support their children’s learning at home.

“I think that there should be a little bit more understanding for parents, especially on the grade chart. So that looks like communication of what those twos versus threes mean in terms of student learning.”

“The whole document really speaks to more of the secondary piece, it doesn’t even have the language of the elementary school and just doesn’t apply. I will tell you, for elementary I know a lot of parents that don’t even open up the report card because they don’t understand it.”

“At the elementary school level I believe the scale is zero to three, but there’s no overall policy guiding how teachers are supposed to use those data points. There is some general, like categories for approaching meet standard, exceed standard, or some words that go along with it, but it would be great to see some more detail around how kids kind of get grouped into those different categories.”

“All the HMH assessments come back zero to 100%. I believe all of the math assessments come back zero to 100% or it’s like a 15 out of 17, or you know, there’s a number scale that goes along with it. It’s not graded on a zero to three scale, and so then there has to be a translation from whatever those numbers are in HMH or math or wherever else they’re doing that type of testing that does have a zero to 100 scale into the zero to three scale, It’s very confusing. For HMH specifically, there’s no translation between any of that stuff, and how it shows up on the report card.”

“I often don’t know how to help my kids, because all I see is maybe the grades on certain assignments. I don’t think my kids can see exactly which questions they got wrong. I remember when I was growing up, I’d be able to see what I got wrong, like, oh, how did I screw that up? There is a big problem with everything being digital and nothing being sent home.”

“I like the writing rubric in elementary school. It’s zero to four. It actually tells you, okay, to get a one, this is what you have to do, to get a two, this is what you have to do, three and four, and four is like you’re exceeding the expectation, right, but it’s, it’s more explicit in what is the student doing, and it’s pretty clear to follow the others.”

Communication & Clarity

In all groups, parents expressed a desire for greater consistency, communication, and clarity of grading practices in classrooms. Often they did not know how their children were being assessed and felt confused about the meaning of grades. They suggested a requirement for greater consistency, at least across classes or schools, if not at the division level, and for individual teachers to communicate their grading system to parents at the beginning of the school year.

“There’s a lot of inconsistencies with the grades, you know, between the teachers - different classes that our students are taking, some of them will count classwork, or, you know, homework with the 10%. Some of them don’t even know that they’re allowed to do that. It’s just, it’s not consistent, and so it’s very confusing.”

“Some grades are weighted, but in Powerschool, you can’t really tell how it’s being weighted. That information needs to be included, so you know exactly how things are being accounted for.”

“I think there’s a lack of communication to parents about how, if individual teachers are going to be implementing it differently, or if an entire school is going to be implementing it differently than an entire other school, that communication seems to maybe have fallen off somewhere along the way.”

“Some kind of communication should go out at the beginning of the year, for sure, about exactly what the expectations are, exactly the kinds of formative and summative assessments, and I think that that could be an important part of a phase in, like, we really want to take curriculum and instruction seriously, and we really want you to understand what we’re doing, and you know, partner with us in this.”

“Some teachers send a grading policy home to parents to read and sign, but that has not been consistent. My son hasn’t had that from every teacher, but it has been great with the ones did.”

Summary of Parent Perspectives on Grading

Parents expressed support for the core tenets of mastery learning: aligning grades to standards and providing spiraling curricula that allow students to revisit standards using different modalities and demonstrate their understanding of standards throughout the year. Parents generally desired more feedback on students’ work habits, either incorporated into an overall grade or reported separately.

Retakes were seen by some as productive opportunities to strengthen learning, but only when accompanied by detailed individual feedback and re-assessment using different modalities; however parents noted the burden this placed on teachers and expressed concern that retakes without any limitations negatively impacted students' motivation and overwhelmed teachers.

Parents were divided on a grading floor of 50, but found consensus in giving lower grades for assignments not attempted. They expressed desire for greater consistency in policies around late work and suggested a reasonable grace period between 3 and 10 school days that could also incorporate late penalties. Elementary parents preferred a 4 point scale to which all reported assessment data could be aligned. Parents also suggested a minimum number of assignments per marking period to adequately describe student progress

VIII. CONCLUSIONS

The results of ACPS's grading policy and guidelines in 2020 have been mixed, with success in some areas for some groups, and substantial need for improvement in many others in which the policy is not meeting its goals of consistency, validity, and improving student learning. There is broad consensus among students, parents, and teachers that the policy lacks specific parameters and the guidelines must be changed to address clear problems in implementation.

While grades for lower-performing students have become better indicators of real achievement, the opposite is true for higher performing students. There is some evidence for grade inflation among already advantaged students even, while failure rates for lower performing students have risen. In secondary reading, grades for lower performing students have nearly doubled in their capacity to explain standardized tests results - a remarkable improvement. However the greater noise introduced into grades for higher performing students nearly erases the overall improvement for reading, and does erase the gain in mathematics. Thus, ACPS needs to review how its guidelines apply differently to higher and lower performing students and adjust them accordingly. Recommendations to do so follow in the next section.

Qualitative data reveal that some elements of the policy have had the intended effects. Many students, parents, and teachers report that retakes, when properly implemented, have improved student learning. Similarly, students, parents, and teachers indicated that the grade floor of 50 has been helpful for students who work hard yet struggle with the content, and support them to remain engaged instead of giving up.

These positives notwithstanding, however, these same groups indicate many problems with implementation that lead to the opposite effect: a decrease of student motivation to study for tests and complete assignments. They also note substantial inconsistency among teachers in how retake policies are implemented, noting that when retakes are provided to all students without feedback and remediation, it decreases both motivation and validity.

Overall, ACPS's community embraces many principles of SBG; however the current policy and guidelines do not support consistent or effective implementation of these principles. Comparative policy analysis revealed many areas in which ACPS could strengthen its policy and guidelines. For example, ACPS is an outlier in its lack of specificity and parameters on retakes, grade floors, communication to parents, late assignments, and grading scales. The following section notes recommendations for changes to policy and guidelines based on the findings in this report.

IX. RECOMMENDATIONS

There are opportunities to improve both the policy and guidelines on grading in ACPS. An initial change is to move the guidelines to implementation procedures and add them to School Board documents for easier access and greater clarity of expectations. Both the policy and procedures should also change many of the "should" statements to "shall" to indicate the expectations for policy implementation and reduce inconsistency. Finally, while the procedures should contain most of the details on implementation, there are several places where the policy should provide more parameters for implementation to make clearer the framework under which grading occurs in ACPS.

Policy

While ACPS's policy provides a sound philosophical grounding for SBG, it lacks key elements present in other school districts that provide a stronger foundation for consistent and effective implementation. Below are suggestions for additions to the policy that can be implemented for the 2026-2027 school year.

Parent Communication

Most peer districts stipulate minimum requirements for parent communication of grades in their policy. ACPS should consider similar language: "School staff shall ensure that explanation of the grading system is communicated to students and parents/legal guardians at the beginning of each school year. Schools shall communicate to parents and students on a regular basis an evaluation of the student's progress toward meeting the academic objectives."

Procedures and Review

To add procedural support currently lacking in the policy, the following language is suggested: "The School Division shall create procedures to support staff in implementing this policy, and shall conduct periodic reviews with teacher and student input to ensure procedures are meeting the goals of grading accuracy, consistency, and support for student learning. These procedures shall include requirements and reasonable limitations for student reassessments, a minimum grade that differentiates between attempted and unsubmitted assignments, regulations on accepting late work from students, requirements for communication of grades and notifications of risk of failure to parents, a minimum number of assignments per marking period, and grading scales."

Procedures

In addition to changes in policy, ACPS should also consider changes to its grading procedures. Some changes can be implemented in the 2026-2027 school year, while others will require extensive training, so are recommended for implementation in the 2027-2028 school year. These changes will need to be implemented with extensive communication for students, especially those in high school, as many of them have only ever known grading under the current guidelines, and the recommended changes below will represent a big shift. In the words of a parent focus group member: “While I do think changes to the current system are warranted and will be better for our kids, please remember that those who are in middle and high school have only ever known these policies, and transitions...will require adjustment and relearning on their part. Please consider how best to support them through any transition.”

Recommended Changes for 2026-2027

Reassessment

Given the importance of this issue to the community, as well as quantitative evidence of grade distortion occurring for historically high achieving students, it is recommended that language on reassessment be added to the procedures to provide structure that addresses this issue: “Reassessment practices must be consistent with division procedures and within each school. Students who score below a letter grade of B, or a rubric descriptor of “good” on the secondary grading scale, shall be eligible for one opportunity to reassess their learning on each summative assessment; however, this shall occur only after the student has received feedback on the assessment and completed any required remediation. The teacher shall determine the type and timeframe of remediation and implementation of the retake or revision. Students shall not earn a grade higher than the minimum threshold for a letter grade of B, or a rubric descriptor of “good” on a reassessment. Teachers may alternatively choose to evaluate students’ learning on specific standards throughout the school year using multiple modalities and update grades accordingly.”

Create a Grade Floor that Differentiates between Completed and Unattempted Work

While there was disagreement among teachers, students, and parents about assigning a 0 as the lowest grade in a 100 point scale, there was near consensus that work that is turned in should receive a grade higher than work that is not attempted. Given the deep concern over this issue, especially among teachers, it is recommended that this is addressed in procedures as in divisions such as Fairfax, Loudoun, and Augusta. It is important to note that students who have historically struggled have had improvement in the validity of their grades since 2020, and it is possible that the grade floor has played a role in this; thus, the Board should be careful not to adjust the existing guidelines too drastically, while still addressing community concerns.

One way to do this is to move the grade floor to 40 to create a new descriptive category of “not attempted” that clearly differentiates between unattempted work and submitted work. This would also address concerns in the current guidelines of having a passing grade that is too close to the grade floor. Separating the floor 20 points from the passing grade doubles that distance. A grade scale of 40-100 creates a six category grading system, within SBG research’s recommendation of 4-7 categories, and brings ACPS closer to the International Baccalaureate system in terms of the

number of categories and the benchmark for passing. Unfortunately, it is not possible to create a 0-60 grading scale using ACPS’s current Student Information System (Power School), as it requires grades to be entered in the form of a percentage.

To effect this, ACPS could adopt a modified version of Fairfax’s language: “To create a six category grading system, the lowest grade that may be assigned to any completed work is fifty (50), provided the student has made a reasonable attempt to complete the assignment. A ‘reasonable attempt’ occurs when a student submits an assignment with some evidence that an attempt was made to demonstrate knowledge. Teachers may assign a forty (40) if the student has not made a reasonable attempt to complete the work. The lowest grade for a reasonable attempt on an assignment will be a ten (50).”

Publish Grading Scales with Descriptors

Most divisions provide grading scales in their policy or procedures as a clear statement of how grades are produced and what they mean. Below is a sample secondary grading scale based on current practice and improvements suggested from the community that could be added to procedures prior to the 2026-2027 school year. The procedure should also describe the category “Limited Understanding” as the current threshold for receiving credit in a high school course, but that mastery is considered a grade of “Satisfactory” or higher. The scale below sets the lowest grade to forty (40) to support clearer communication about when assignments are not attempted and to double the distance between a not-attempted score and a passing score for a course.

Table 2. Secondary Grade Scale

Letter Grade	Number Scale	Description	Quality Points	Quality Points for High School Honors, AP, IB, and DE Courses
A	90-100	Excellent	4	5
B	80-89	Good	3	4
C	70-79	Satisfactory	2	3
D	60-69	Limited Understanding	1	2
F	50-59	Very Poor	0	0
F	40-49	Not Attempted	0	0

Late Work

Seven peer divisions provide guidelines on accepting late work from students, and doing so in ACPS's procedures would address concerns among teachers, parents, and students. ACPS should consider adopting language similar to Arlington: "Late work shall be accepted to document learning/mastery; however late work submitted after the specified due date but prior to the summative assessment for the teaching unit shall be penalized up to 30 points from the value of the assessment. Each school shall set consistent, reasonable guidelines for accepting late work and applying any penalties. If the unit extends across marking periods, the cut off for some assignments may be the end of the marking period. Assignments not submitted prior to the end of a teaching unit shall result in a grade of "not attempted" (40) and summative assessments that are not made up within five school days may receive a grade of "not attempted" (40); except for in circumstances that produce excused absences, in which case the student may receive one additional day to make up the assignment or assessment for each day of excused absence."

Minimum Number of Assignments

Six peer districts provide a minimum number of assignments leading to grades during a specified period of time. Such a measure would address parent and student concerns about not receiving enough graded assignments in a course to have their understanding assessed using multiple modalities and allowing for change over time. The most common requirement is seven graded assignments per quarter, "Teachers will include at least seven assignments per quarter in the gradebook to assess student learning."

Required Parent Notification for Risk of Failure

Six of our peer districts detail circumstances in which teachers must communicate with parents about grades if their child is not meeting proficiency expectations. Parents in ACPS noted the need for improved communication with teachers about grades, especially in situations where failure may occur. ACPS should consider adopting language similar to Virginia Beach, Fairfax, and Chesterfield: "Teachers are required to communicate with parents or caregivers regarding student progress if a student is at risk of failing or has earned a grade of D or F on their report card."

Gradebook Setup

Seven peer districts describe the gradebook setup used by teachers in their grading procedures available to the public. Currently, ACPS's documentation on gradebook setup is internal and distributed through the Technology Department. ACPS may consider adopting language similar to York, Arlington, Fairfax, and others: "Final grades in secondary courses will be calculated with each semester equaling 50% of the final grade. Grades will be rolling through each semester and teachers will be able to update grades in both semesters throughout the year."

Recommended Changes for 2027-2028

The following changes are recommended for Elementary Grading; however, these will require substantial training for staff, so are suggested to be implemented for the 2027-2028 school year.

Elementary Grading Procedures

ACPS should publish procedures on grading specific to elementary schools. Both teachers and parents describe a need for greater clarity in elementary grading practices and requested alignment for all assessments and grade reports sent home to be aligned to one scale for easier understanding. This is also an opportunity for ACPS to advance the principles of SBG further, as elementary schools have a grading system that is closest to an authentic SBG system, yet also have the least amount of guidance for how to implement it. Adding a fourth category to the elementary grading scale presents an opportunity for the division to bring greater clarity and consistency on how to implement SBG.

Table 3. Elementary Grading Scale

Standard Grade	Description
4	Excellent
3	Proficient
2	Below Standard
1	Far Below Standard

Provide Feedback on Work Habits

Adopt procedures for elementary and secondary schools to report on student work habits separate from academic achievement and mandate that grades for work habits be reported separately from academic mastery.

Move to a Categorical Grading System in Secondary Schools

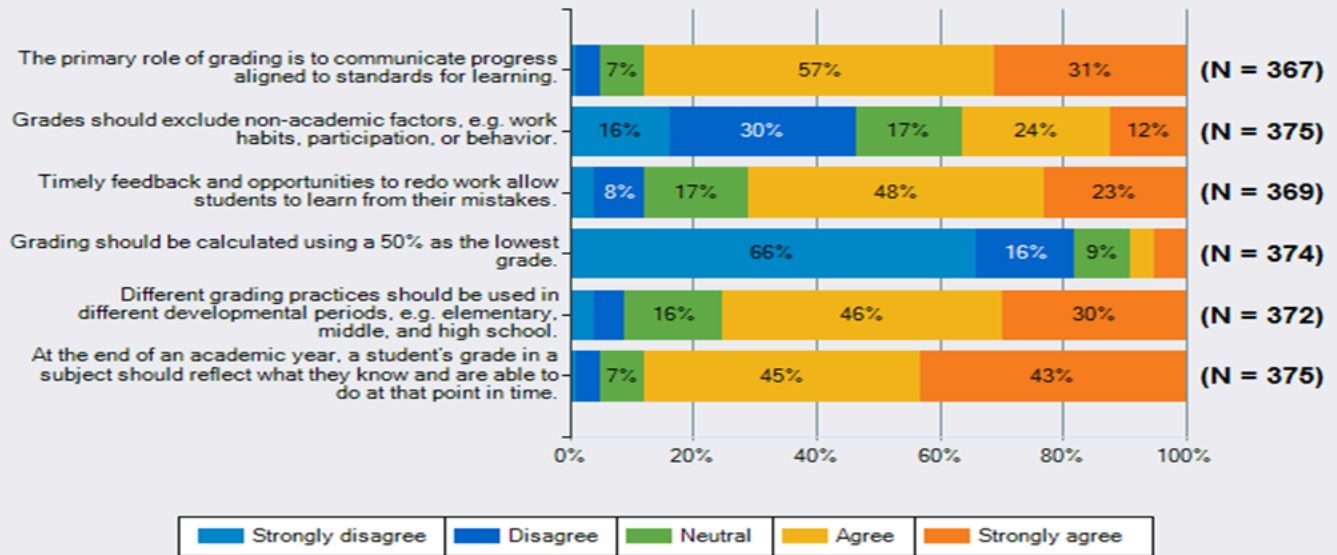
Provide professional learning and training for secondary teachers on implementing six-point categorical grades with rubrics as their base grading system, rather than percent grades. Teachers would score students' work with a letter grade derived from a rubric, which would translate into a numerical grade, rather than entering a numerical grade that translates into a letter grade. Research shows this improves validity and reliability of grading, while also supporting teacher collaboration to build common rubrics. Under this system, students without a minimum number of submitted summative assessments to show understanding would not earn credit for a course. This system would incorporate the improved reliability and consistency of categorical grading systems, and align with post-secondary systems, including Advanced Placement and International Baccalaureate programs.

X. REFERENCES

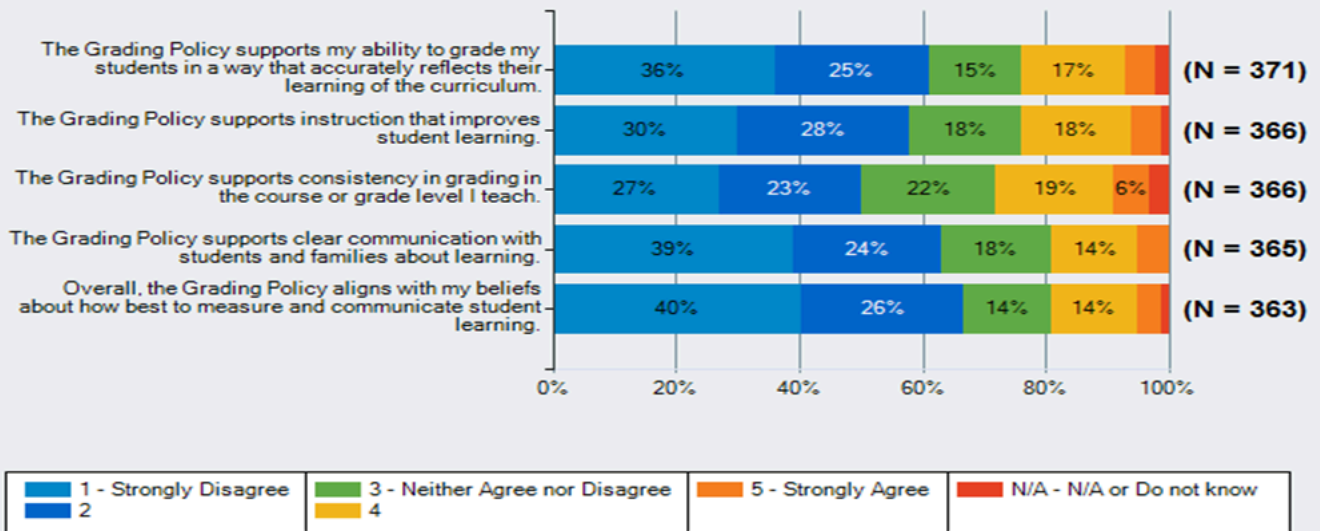
- Brookhart, S. M. (1991). Letter grading practices and validity. *Educational Measurement: Issues and Practice*, 10(1), 35-36.
- Brookhart, S. M., Guskey, T. R., Bowers, A. J., McMillan, J. H., Smith, J. K., Smith, L. F., Stevens, M. T., & Welsh, M. J. (2016). A century of grading research: Meaning and value in the most common educational measure. *Review of Educational Research*, 86(4), 803–848.
- Kelly, F. J. (1914). *Teachers' marks: Their variability and standardization*. Contributions to Education, No. 66. Teachers College Press
- Guskey, T. R., Townsley, M., & Buckmiller, T. M. (2020). The impact of standards-based learning: Tracking high school students' transition to the university. *NASSP Bulletin*, 104(4), 257–269.
- Guskey, T. R. (2024). Addressing inconsistencies in grading practices. *Kappan*, May 2024.
- Guskey, T.R., Fisher, D., & Frey, N. (2024). The unwinnable battle over minimum grades. *Educational Leadership*, October 2024..
- Haystead, M. W., & Marzano, R. J. (2009). Meta-analytic synthesis of studies conducted at Marzano Research Laboratory on instructional strategies. Englewood, CO: *Marzano Research Laboratory*. Retrieved from http://www.marzanoevaluation.com/files/Instructional_Strategies_Report_9_2_09.pdf
- Huey, M.E., Silvey, P., Vaughan, A., & Fisher, A. (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). <https://doi.org/10.1016/j.stueduc.2022.101211>
- Kelly, S. (2008). What types of students' effort are rewarded with high marks? *Sociology of Education*, 81, 32-52.
- Knight, M., & Cooper, R. (2019). Taking on a new grading system: The interconnected effects of standards-based grading on teaching, learning, assessment, and student behavior. *NASSP Bulletin*, 103(1), 65–92.
- Lane, J. (2021, Jan. 15). [Primer on traditional and equitable grading practices](#). Superintendent's Email. *Virginia Department of Education*.
- Rugg, H. O. (1918). Teachers' marks and the reconstruction of the marking system. *Elementary School Journal*, 18, 701-719.
- Swan, G. M., Guskey, T. R., & Jung, L. A. (2014). Parents' and teachers' perceptions of standards-based and traditional report cards. *Educational Assessment, Evaluation and Accountability*, 26, 289-299.
- Thorsen, C., & Cliffordson, C. (2012). Teachers' grade assignment and the predictive validity of criterion-referenced grades. *Educational Research and Evaluation*, 18, 153-172.
- Willingham, W. W., Pollack, J. M., & Lewis, C. (2002). Grades and test scores: Accounting for observed differences. *Journal of Educational Measurement*, 39, 1- 37.
- Zhang, B. (2024). A didactic explanation of standards-based grading. *Educational Research and Development Journal*, 27(2), pp. 42-56.

XI. APPENDIX A: Full Results from Teacher Survey

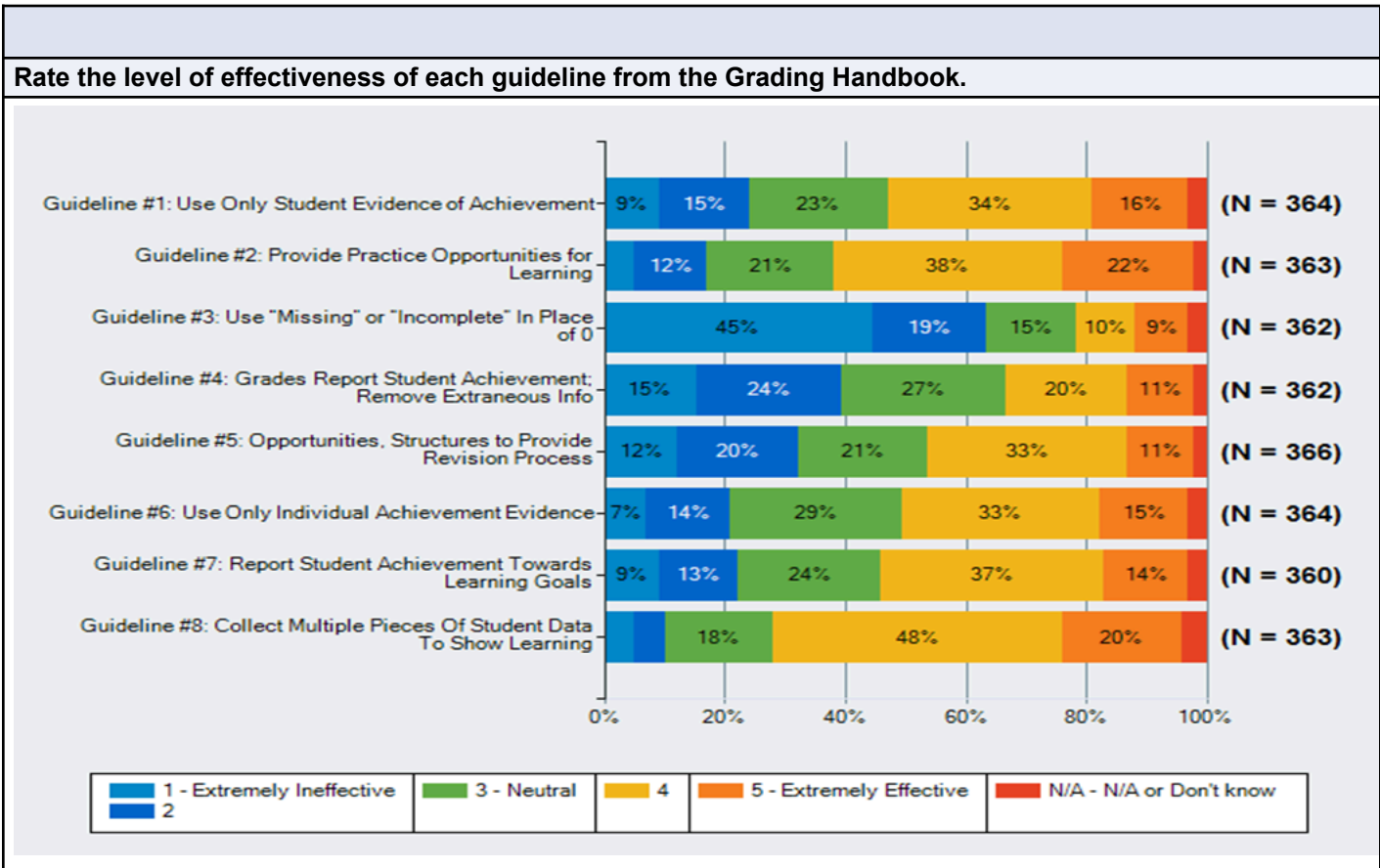
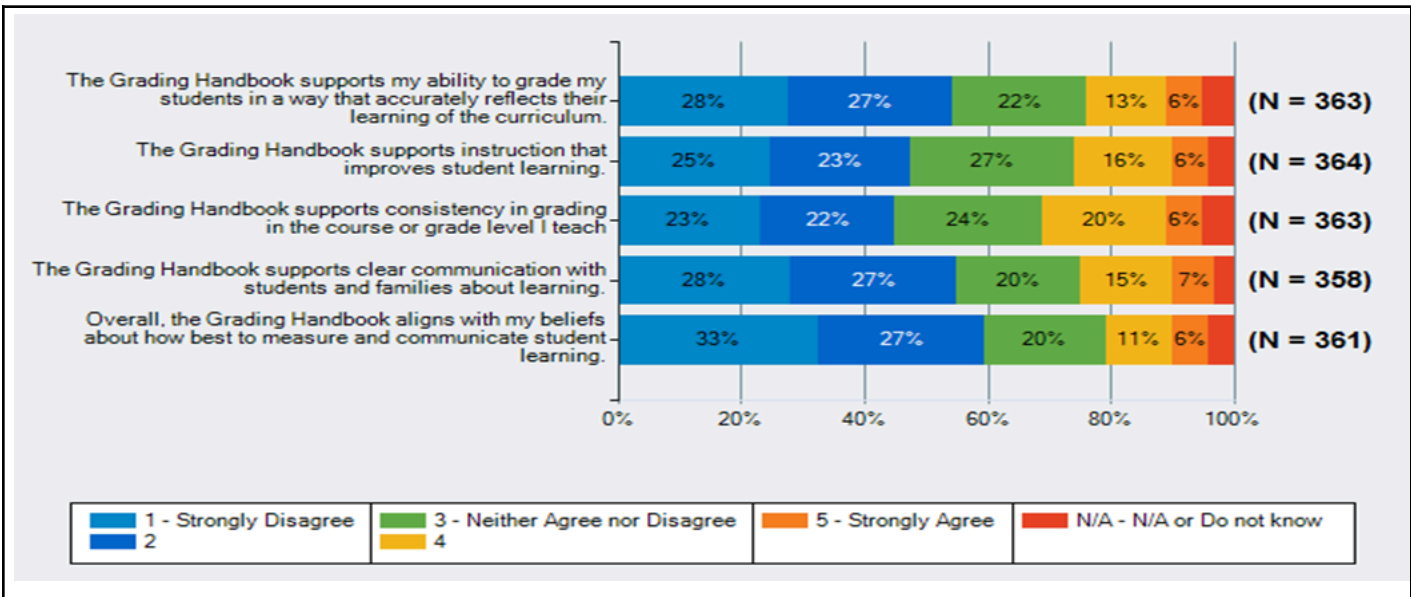
Rate your level of agreement with each statement.



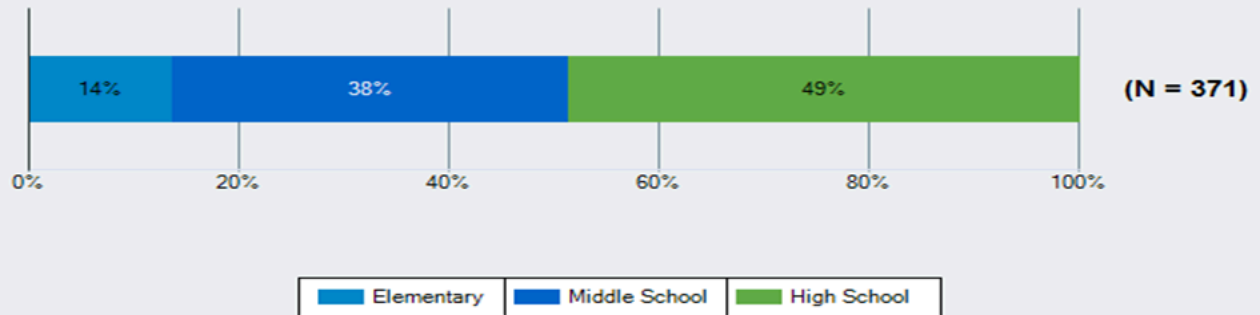
Rate your level of agreement with each statement about the Grading Policy IKDA specifically. The Grading Handbook guidelines and practices will be evaluated in later questions.



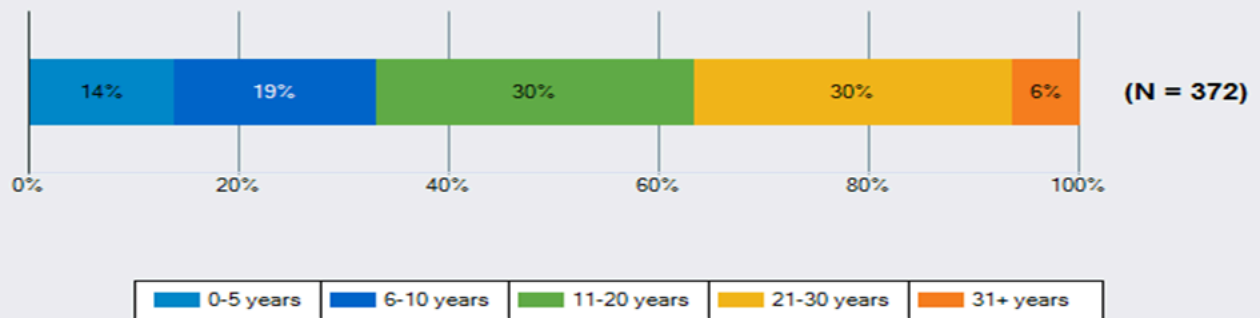
Rate your level of agreement with each statement about the Grading Handbook as a whole. You will evaluate individual guidelines and practices in later questions.



At what level do you primarily teach?



How many years have you worked as a teacher?



What is your primary content area?

