EdWorkingPaper No. 26-1380

The Chronic(les) of Absenteeism Measurement: Unpacking the Many Measures of Attendance and Evidence for a Lower Chronic Absenteeism Threshold

Tiffany Wu
University of Michigan

Christina Weiland
University of Michigan

Thomas Staines
University of Michigan

Chronic absenteeism has surged in recent years, drawing growing policy and research attention. However, a complicating factor often overlooked is that the measurement of absenteeism is inconsistent, with substantial researcher degrees of freedom. This study investigates how researchers' measurement choices shape predictions of academic risk and how absenteeism can be more effectively operationalized as an early warning signal. Using a sample of 8,891 students followed from Pre-K through eighth grade in Boston Public Schools, we (1) describe developmental patterns in absenteeism; (2) apply ROC curve analyses to evaluate the diagnostic accuracy of multiple absence measures for predicting scores on state standardized tests; (3) use Youden's J to derive empirical thresholds for chronic absenteeism; and (4) assess predictive validity over time. We find that measurement choices matter. By middle school, total and unexcused absences are more predictive of low academic performance than excused absences. Additionally, empirically derived thresholds for identifying students at risk consistently fall below the widely used chronic absenteeism benchmark of missing 10% of school days (≈18 days). We discuss implications for research, policy, and early warning systems.

VERSION: January 2026

Suggested citation: Wu, Tiffany, Christina Weiland, and Thomas Staines. (2026). The Chronic(les) of Absenteeism Measurement: Unpacking the Many Measures of Attendance and Evidence for a Lower Chronic Absenteeism Threshold. (EdWorkingPaper: 26-1380). Retrieved from Annenberg Institute at Brown University: https://doi.org/10.26300/1zvw-qw93

The Chronic(les) of Absenteeism Measurement: Unpacking the Many Measures of Attendance and Evidence for a Lower Chronic Absenteeism Threshold

Tiffany Wu
University of Michigan

Christina Weiland
University of Michigan

Thomas Staines
University of Michigan

Authors' note: This study was supported by grants from the Institute of Education Sciences, U.S. Department of Education: R305A220036 and R305B200011. The authors thank the Boston Public Schools (BPS), TeeAra Dias, the BPS Department of Early Childhood coaches and staff, the BPS Office of Data and Accountability (especially Megan Toney), and the Massachusetts Department of Elementary and Secondary Education (especially Tori Puglia and Matt Deninger) for their support and partnership.
Correspondence concerning this article should be addressed to Tiffany Wu, School of Education, 610 E. University Ave, Ann Arbor, MI 48104, or via e-mail at wutiffa@umich.edu.

Abstract

Chronic absenteeism has surged in recent years, drawing growing policy and research attention. However, a complicating factor often overlooked is that the *measurement* of absenteeism is inconsistent, with substantial researcher degrees of freedom. This study investigates how researchers' measurement choices shape predictions of academic risk and how absenteeism can be more effectively operationalized as an early warning signal. Using a sample of 8,891 students followed from Pre-K through eighth grade in Boston Public Schools, we (1) describe developmental patterns in absenteeism; (2) apply ROC curve analyses to evaluate the diagnostic accuracy of multiple absence measures for predicting scores on state standardized tests; (3) use Youden's J to derive empirical thresholds for chronic absenteeism; and (4) assess predictive validity over time. We find that measurement choices matter. By middle school, total and unexcused absences are more predictive of low academic performance than excused absences. Additionally, empirically derived thresholds for identifying students at risk consistently fall below the widely used chronic absenteeism benchmark of missing 10% of school days (≈18 days). We discuss implications for research, policy, and early warning systems.

Keywords: chronic absenteeism, student attendance, receiver operating characteristic (ROC) curves, signal detection theory, Youden's J

The Chronic(les) of Absenteeism Measurement: Unpacking the Many Measures of Attendance and Evidence for a Lower Chronic Absenteeism Threshold

In the 2013-2014 school year, the United States Department of Education reported nationwide rates of chronic absenteeism for the first time through the Civil Rights Data Collection survey. That year, approximately 14% of school-age children, or 6.8 million students nationwide, were chronically absent (U.S. Department of Education, 2016a). In less than a decade, the proportion of chronically absent students has doubled to 28% in 2022-23 (U.S. Department of Education, 2025), an alarming trend with serious implications for educational equity and policy.

Most commonly defined as missing 10% or more of school days, chronic absenteeism has been consistently correlated with lower academic achievement, weaker socioemotional skills, and reduced likelihood of high school graduation and postsecondary enrollment (Allensworth et al., 2021; Romero & Lee, 2007). Using quasi-experimental methods, emerging research suggests that regular school attendance in the early grades is not only strongly associated with, but may causally impact, long-term academic outcomes (Goodman, 2014; Gottfried, 2013). These patterns are not evenly distributed across students: children from historically marginalized communities disproportionately face structural barriers to consistent attendance, resulting in an "attendance gap" that parallels broader disparities in educational opportunity (Henderson & Fantuzzo, 2023). Because time spent in school is a primary means through which students access instruction and school-based supports, unequal attendance patterns may function as a mechanism through which opportunity gaps contribute to disparities in academic achievement (Chang & Romero, 2008; Henderson & Fantuzzo, 2023). Given this evidence, it is unsurprising that national attention has increasingly focused on improving attendance as one strategy for narrowing the opportunity gap.

One result of the growing national focus on attendance has been the adoption of chronic absenteeism as an accountability indicator and a broader shift toward using absenteeism-based measures in longitudinal research and policy (Every Student Succeeds Act, 2015). While the use of absenteeism outcomes and predictors has become increasingly common in education research, there has been relatively little examination of how different measurement choices influence key relationships. As Dougherty (2018) observed, although there are exceptions, absenteeism measures vary broadly across studies, with little testing of the sensitivity of results to measurement decisions. Moreover, Henderson and Fantuzzo (2023) demonstrated diagnostic-accuracy differences between using unexcused and excused absences, calling into question the common practice of defining chronic absenteeism by total absences and the use of the conventional 10% of total absences threshold.

In this study, we respond to the growing demand for greater clarity and rigor in how chronic absenteeism is measured and applied in educational research and policy. In particular, using a demographically diverse sample of students from Boston Public Schools, we investigate descriptive patterns of absenteeism from Pre-K to 8th grade, use Signal Detection Theory and Receiver Operating Characteristic (ROC) curve methodology to evaluate the diagnostic accuracy of different absence measures for predicting later academic achievement, empirically test candidate thresholds for defining chronic absenteeism, and assess predictive validity of multiple absence measures over time. We find that absence type and operationalization meaningfully shape diagnostic conclusions, with unexcused and total absences becoming increasingly more predictive of standardized test scores compared to excused absences in later grades. We also find that empirically derived thresholds based on ROC curve methodology are consistently lower than the widely used 10%, or 18 day, cutoff, with optimal thresholds in the 3-7% range, potentially offering a more accurate and timely signal for identifying students at risk. Our findings underscore how chronic absenteeism is not a fixed construct but a series of definitional choices and adding to the literature on how student attendance can be more rigorously and consistently operationalized to strengthen both research and policy applications.

The Measurement of Student Absenteeism

Calculating Student Absences

At its core, absenteeism is the inverse of attendance. For researchers working with school, district, or longitudinal state data, measuring student absences often involves a series of key decisions. These decisions can fall into three main categories:

1. Absence type: Whether to use excused absences, unexcused absences, or total absences (a combination of both excused and unexcused absences).
2. Variable type: Whether to measure absenteeism as a continuous variable (e.g., number of days absent) or as a binary indicator of chronic absenteeism (e.g., whether a student meets or exceeds a defined threshold).
3. Scale of measurement: Whether to use the raw number of days absent or to calculate an absence rate (e.g., days absent divided by days enrolled).

Among the options for absence type, excused absences are often perceived as less detrimental because they indicate that communication between a student’s family and the school has occurred (Gottfried, 2009). These absences are typically associated with reasons considered legitimate by schools, such as illness, and may reflect a higher level of parental engagement. In contrast, unexcused absences, also referred to as truancy, are viewed as more concerning, as they may reflect disengagement with school or insufficient adult supervision. Several studies support these distinctions. Gershenson et al. (2017) reported that unexcused absences were twice as harmful to academic achievement as excused ones among fourth and fifth graders in North Carolina. Gottfried (2009) found that a higher proportion of excused absences relative to total absences was actually associated with positive effects on reading and math test scores. More recently, Henderson and Fantuzzo (2023) concluded that only unexcused absences provided diagnostic accuracy in predicting later academic performance in early elementary grades.

However, the widely-used chronic absenteeism binary indicator does not distinguish between absence types. It is most commonly defined as having a *total* absence rate of greater than or equal to 10% (Allison et al., 2019; Faria et al., 2017). In most U.S. schools, this translates to missing 18 or more days out of a typical 180-day year. Implicit in this definition is the assumption that all absences, regardless of type or underlying cause, have equivalent consequences for student learning. Indeed, some evidence also supports this assumption. Balfanz and Byrnes (2012) argued that the total amount of instructional time lost is what matters most. Furthermore, in the Gershenson et al. (2017) study cited above, the authors found no statistically significant difference between excused and unexcused absences when analyzing academic outcomes using nationally representative ECLS-K data, contrasting with their North Carolina findings and highlighting the potential importance of context and sample characteristics.

In addition to the ambiguity surrounding absence type, there is no universal agreement on the appropriate cutoff for how much school a student has to miss before being chronically absent. Dougherty (2018) reviewed the literature and found that definitions vary widely, from not specifying a threshold (Gottfried, 2009) to using fixed day counts such as missing 11, 15, 18, or 20 days (Gottfried, 2015; Morrissey et al., 2014; Gershenson, 2016; Sheldon & Epstein, 2004), to using percentage-based thresholds, like 10% of a student's total enrollment days (Aucejo & Romano, 2016). When deciding whether to use raw counts of absent days or to use an absence rate, researchers must consider enrollment variation. Absence rates, calculated as the number of days absent divided by the number of days enrolled, adjust for differences in enrollment length caused by student mobility or by calendar variations across school types (charter, public, private) and states. Because 31 states and the District of Columbia require a minimum of 180 instructional days, most schools operate on a standard 180-day calendar. However, notable variation exists between states. Colorado, for example, mandates a minimum of 160 days. Thirty-five states also differentiate requirements by grade level, either in terms of days or instructional hours (Silva-Padron & McCann, 2023). As such, using rates may offer a more comparable metric for identifying chronic absenteeism across diverse student populations. However, day counts align with how attendance is typically tracked and communicated by educators and families.

In making measurement decisions, researchers may want to take into account that the causes and implications of absenteeism vary across developmental stages. In early grades, school attendance is largely dependent on parents or caregivers. As a result, unexcused absences may be less frequent, and even when present, they may reflect logistical challenges or communication breakdowns at the family level (e.g., forgetting to call the school) rather than student intent to miss school (Chang & Romero, 2008; Epstein & Sheldon, 2002). In such cases, researchers may be more likely to use total absences to better capture the full picture of lost instructional time. Absences in the early grades may also be more heavily influenced by common childhood illnesses, especially as children adjust to new social environments (Ready, 2010). In contrast, in later elementary and secondary grades, absenteeism is more likely reflective of student-driven factors, such as academic disengagement or school avoidance (Kearney, 2008). As students gain autonomy over their routines, patterns of unexcused absences may become more meaningful.

In all, the 3 absence types, 2 variable types, and 2 scales of measurement already yield 12 different combinations of possible absenteeism measures. Adding in variation in how chronic absenteeism is defined (at minimum, a choice between two different thresholds such as 15 days versus 18 days) doubles those options to 24. Factoring in developmental differences across grade spans—at minimum, three broad bands of elementary, middle, and high school—multiplies this further to at least 72. Taken together, these choices provide researchers with more than 70 degrees of freedom in how absenteeism is operationalized, each with the potential to influence findings, interpretations, and policy implications.

Chronic Absenteeism as a Signal

The variability in how absenteeism is measured becomes especially consequential in the context of chronic absenteeism, which has increasingly been used as both a policy and research indicator. Following the passage of the Every Student Succeeds Act (ESSA) of 2015, 36 states and the District of Columbia adopted school-level chronic absenteeism as an official indicator of school performance (Swaak, 2018). Around the same time, chronic absenteeism became a central feature of many early warning systems (EWSs), a data-driven tool which aims to proactively identify students at risk of negative outcomes (U.S. Department of Education, 2016b). Within EWS frameworks, chronic absenteeism is typically used in two ways: as a standalone threshold indicator (e.g., flagging students once they miss a certain number of school days) or as a predictor variable in regression-based models (Canbolat, 2024). Although most EWSs have focused on preventing high school dropout, limited recent research has suggested that chronic absenteeism may also be a valuable outcome and risk signal for EWS algorithms in earlier grades, where indicators such as standardized test scores or disciplinary referrals may not be as salient or not yet available (Wu & Weiland, 2025).

However, despite the growing use of chronic absenteeism in both research and practice, there remains limited empirical evidence validating specific absence thresholds as effective triggers for intervention. In one of the few studies addressing this issue, Canbolat (2024) used a multiple-cutoff regression discontinuity design and found that threshold-based EWSs reduced chronic absenteeism, defined using a 10% threshold, only among socioeconomically advantaged students. The same system showed no statistically significant effect for socioeconomically disadvantaged students or those identified using a 4% threshold. While these disparities may stem from factors such as the EWS algorithm or local implementation practices, they also raise important questions about the appropriateness of the 10% threshold. It is possible that the cutoff itself may need to be adapted or changed to be more effective across diverse student populations.

This concern parallels earlier critiques of arbitrary threshold-setting in accountability systems. Ho (2008) critiqued the reliance on proficiency cutoffs in test-based accountability frameworks, noting how the arbitrary location of the threshold can distort inferences about school or student performance. Dougherty (2018) extended this critique to chronic absenteeism, noting that conclusions drawn from attendance data can be highly sensitive to how chronic absenteeism is defined. Specifically, Dougherty argued that a universal 10% threshold may mask meaningful variation in who is identified as chronically absent, particularly across grade levels and socioeconomic subgroups. These insights underscore the importance of measurement sensitivity and contextual nuance in research based on absenteeism data.

Signal Detection Theory & the ROC Curve

Given the concerns raised about the validity and appropriateness of various chronic absenteeism thresholds, it is useful to consider alternative frameworks for evaluating how well absence measures identify students at risk. One such framework is Signal Detection Theory (SDT), which provides a structured way to assess decision-making under uncertainty. Originating in engineering and the computer sciences with radar researchers and now widely applied in medicine and cognitive psychology, SDT helps distinguish between meaningful signals (e.g., the number of absences that truly indicate academic risk) and background noise (random variation or absences unrelated to risk) (Heeger, 2007). Though underutilized in education research, SDT offers promising tools for addressing the chronic absenteeism threshold problem. SDT focuses on whether a measure has sufficient diagnostic accuracy to correctly identify students as at risk of a binary outcome, such as failing a class. This allows researchers to evaluate how well different absence measures distinguish between students who are and are not at risk of experiencing later academic difficulties.

A widely used and well-tested method within the SDT framework is the Receiver Operating Characteristic (ROC) curve. The ROC curve has been commonly applied to determine the classification accuracy of a model or indicator. Although its use in education is still emerging, the ROC curve has gained traction in several subfields, including educational measurement (Keller-Margulis, Shapiro, & Hintze, 2008), learning analytics (Hutt et al., 2019), and predictive modelling for machine-learned EWSs (Wu & Weiland, 2025). However, to our knowledge, only one study has applied the ROC framework to evaluate the diagnostic accuracy of a chronic absenteeism indicator: Henderson & Fantuzzo (2023) employed ROC curve analyses using administrative data from a large urban school district and found that only unexcused absences from kindergarten to second grade provided diagnostic accuracy in predicting future academic achievement in third grade. They called for replication of ROC analyses across additional school districts and over time to better assess the longitudinal validity of different attendance metrics. Most notably, they underscored the lack of an empirical basis for the widely adopted chronic absenteeism threshold and recommended the use of ROC methodology to establish a more evidence-based, context-sensitive cutoff for future studies.

Current Study

This study contributes to a growing body of research that interrogates how absenteeism is measured and used using ROC curve methodology. Building on Henderson and Fantuzzo's insights in the early elementary years, we broaden the empirical scope of ROC curve applications by examining the diagnostic accuracy of absenteeism measures from early elementary through middle school within a different large, urban district. In doing so, we test the generalizability of this analytic framework across diverse grade spans and developmental contexts. In addition, we demonstrate one way to empirically identify a chronic absenteeism threshold, discussing its strengths and limitations relative to conventional cutoffs. In particular, we address the following research questions:

RQ1. How do absenteeism descriptive patterns differ across Pre-K through 8th grade depending on how absence is measured?

RQ2. Using ROC curve analysis, what is the diagnostic accuracy of different absenteeism measures across different grade levels in predicting students *Not Meeting Expectations* for math and ELA standardized tests in 8th grade?

RQ3. Can ROC curves be used to identify an empirically justifiable threshold for chronic absenteeism?

RQ4. How does the diagnostic accuracy and optimal threshold of absenteeism vary across grades prior to 8^{th} grade, and does temporal proximity between absenteeism and assessment improve diagnostic accuracy?

Our primary analyses (RQ2-RQ3) focus on 8^{th} grade standardized test outcomes because they provide a key benchmark of middle school achievement and readiness for high school. To maintain clarity and brevity, we restrict our detailed reporting to 8^{th} grade in the main text; in RQ4, we examine temporal sensitivity by extending our analyses to grades 3-7 but summarize cross-grade patterns in the main text and provide full grade-specific results in the Appendix.

Methods

Data

Our sample for this paper was the population of students who enrolled in the Boston Public Schools (BPS) Pre-K program for four-year-olds between the 2007-2008 and 2009-2010 school years. We defined our sample by Pre-K entry because over half of U.S. public schools now have a Pre-K program (Little et al., 2025) and because recent research shows that absenteeism patterns established in Pre-K not only persist into the early elementary grades but also can be used to predict later academic and attendance outcomes (Wei, 2024; Wu & Weiland, 2025). We followed students from their Pre-K application to eighth grade (school years 2016-17 to 2018-19). Using each student's unique identifier, we merged on district and state administrative records, including information on demographics, students' attendance records, and ELA and math standardized test scores from the Massachusetts Department of Elementary and Secondary Education. There was a total of 8,891 students across these cohorts.

For RQ1, we included all students with non-missing attendance data for a given grade and verified evidence of enrollment and attendance (i.e., both days enrolled and days attended greater than zero). Sample sizes ranged from $N = 6{,}226$ to $N = 8{,}014$ across grade levels, as shown at the bottom of Figure 2. For RQ2-RQ3, we restricted the sample to students with non-missing state standardized test scores in eighth grade, resulting in a final analytic sample of $N = 7{,}145$. For RQ4, we restricted the sample to students with non-missing state standardized test scores in each respective grade from grades 3-7, with sample sizes ranging from $N = 6{,}813$ to $N = 7{,}343$. We explored whether findings were driven by different students joining or leaving the sample across the Pre-K through 8th grade years. To do so, for RQ1-RQ4, we refit our models using a common sample of students with no missing data in attendance or standardized test scores in all grades. We found our broad conclusions for all RQs were robust to a broader versus common sample approach. All findings for the common sample approach can be found in Appendix S9.

Table 1 reports student demographic characteristics for the state standardized test-defined analytic samples for RQ2-RQ3 (8^{th} grade) and RQ4 (3^{rd} through 7^{th} grades). In 8^{th} grade, the analytic sample was 48% female and racially/ethnically diverse (45% Hispanic/Latino, 29% Black, 15% White, 9% Asian, 3% multiracial/other). Roughly 49% of students spoke English as their home language, 30% spoke Spanish, and 21% spoke another language other than English or Spanish. About 94% of the sample listed the United States as their country of origin, and approximately 58% of students were eligible for free or reduced-price lunch. The samples for grades 3-8 were comparable for all student characteristics except for eligibility for free or reduced-price lunch, where the 8^{th} grade sample was less likely to be eligible.

Measures

Student Absences

For absenteeism, we calculated student absence rate by dividing each student's number of absence days by the total days the student was enrolled in their majority-enrolled school during a given year. Using a rate instead of a raw count of days attended accounts for variation in enrollment length, both between public and charter schools and for students who transferred schools mid-year.

Academic Achievement

To measure academic achievement, we used children's Massachusetts state standardized test performance in both math and ELA from 3rd to 8th grade. In Massachusetts, students are required to complete state assessments in grades 3-8, and the results serve as core indicators for state accountability systems, informing resource allocation and compliance with ESSA requirements (Massachusetts Department of Elementary & Secondary Education, 2018; Massachusetts Department of Elementary & Secondary Education, 2025; Rennie Center for Education Research & Policy, 2018). Across cohorts, the state used three different tests in our study years. Up until 2014, students were administered a version of the MCAS now called "Legacy MCAS." This test reported four achievement levels: Warning/Failing, Needs Improvement, Proficient, and Advanced. In 2015 and 2016, districts were able to choose to administer either Legacy MCAS or the Partnership for Assessment of Readiness for College and Careers (PARCC) assessment, an exam based on Common Core standards. The PARCC exam reported five achievement levels: Did Not Yet Meet Expectations, Partially Met Expectations, Approached Met Expectations, Met Expectations, and Exceeded Expectations. In 2016, 67% of districts in Massachusetts administered PARCC, and within the three largest school districts in the state—Boston, Worcester, and Springfield—individual schools chose which test to administer (Massachusetts Department of Elementary & Secondary Education, 2016). Beginning in 2017, all districts administered a new version of MCAS ("Next-Gen MCAS"). The Next-Gen MCAS reported four achievement levels: Not Meeting Expectations, Partially Meeting Expectations, Meeting Expectations, and Exceeding Expectations.

Because the state transitioned across multiple assessments, we aligned achievement levels pragmatically for our ROC analyses. Consistent with state definitions of not meeting expectations, we grouped together the following categories that fall short of demonstrating at least a partial understanding of the subject matter and would need additional supports to meet expectations: Legacy MCAS's *Warning/Needs Improvement*, Next-Generation MCAS's *Not Meeting Expectations*, and PARCC's *Did Not Yet Meet Expectations*. We use the Next-Generation MCAS terminology and refer to this group collectively as those *Not Meeting Expectations* throughout the rest of this paper. This alignment reflects an analytic choice rather than an official crosswalk. Additional details on the differences between the different assessments and what it means to be in these categories are in Appendix S1.

Analytical Approach

Signal Detection Theory & ROC Curves: A Primer

The central insight of SDT is that even when there is a binary reality (e.g., a condition is either present or not), the information used to make that determination often contains noise—ambiguity, variability, or measurement error—which complicates decision-making in detecting that binary reality. To illustrate the logic of SDT, consider a hypothetical scenario adapted from Professor David Heeger's teaching materials (Heeger, 2007): Consider a radiologist reviewing a CT scan for signs of a tumor. There is an underlying objective "truth": the tumor is either present or absent. However, the radiologist's interpretation of the CT scan, which may already be unclear or difficult to interpret, may or may not correspond with this objective reality. This creates four possible outcomes: a) A hit (true positive): A tumor is present, and the radiologist correctly identifies it; b) A correct rejection (true negative): A tumor is absent, and the radiologist correctly confirms that; c) A false alarm (false positive): A tumor is absent, but the radiologist mistakenly identifies one; or d) A miss (false negative): A tumor is present, but the radiologist incorrectly says there is none. Hits and correct rejections represent accurate classifications, while misses and false alarms constitute diagnostic errors.

The same logic can be applied to absenteeism, for example, when using absence rates to identify which students are at risk of scoring in the *Not Meeting Expectations* range on a standardized test. Just as in the medical example, there is an underlying "truth" (e.g., a student will either score high enough to meet expectations or not), but our measure (e.g., a student's absence rate) may or may not correctly capture that risk. This leads to the same four aforementioned outcomes: hits/true positives (correctly identifying students who will not meet expectations), correct rejections/true negatives (correctly identifying students who will meet expectations), false alarms/false positives (failing to identify students who will meet expectations), and misses/false negatives (failing to identify students who will not meet expectations). The central goal of SDT is to evaluate how well a test distinguishes between true cases and diagnostic errors, especially across different decision thresholds.

A commonly used tool in SDT, ROC analysis can help assess the degree to which a continuous measure (e.g., a student's total absence rate) can accurately determine a binary outcome (e.g., meeting or not meeting expectations on a standardized test). In Figure 1A, we illustrate the "ideal" ROC scenario, where the distributions of students who meet expectations (in blue in Figure 1A) and those who do not meet expectations (in red in Figure 1A) would be entirely distinct on the absence rate spectrum. In this case, a single threshold on the absence rate (dotted line in Figure 1A) would perfectly separate the two groups, enabling a flawless classification: All students with an absence rate lower than the absence threshold would be correctly classified as *Meeting Expectations,* and all the students with an absence rate higher than the absence threshold would be correctly classified as *Not Meeting Expectations.* However, in reality, the actual distribution of these two groups likely overlap, as shown in Figure 1B. Notice that because the distributions overlap, there will inevitably be both misses (false negatives) and false alarms (false positives) no matter what absence threshold we hypothetically choose. For example, setting a low threshold (Figure 1C) captures most students who are likely to not meet expectations, ensuring few students at risk of failing slip through the cracks, but also increases the number of false alarms, which can lead to unnecessary interventions. In contrast, a high threshold (Figure 1D) minimizes false alarms, meaning fewer students are incorrectly flagged. However, the lower false alarm rate comes at the cost of more misses, potentially overlooking students who would benefit from additional supports. In either case, the overlapping distributions underscore a central challenge in SDT: no single threshold is perfect, and each decision point involves important trade-offs in accuracy.

The full range of possible thresholds and their consequences can be visualized using a single graph known as the ROC curve. Rather than focusing on a single cutoff, the ROC curve plots the true positive rate (TPR, also known as sensitivity) on the y-axis against the false positive rate (FPR, also known as 1 – specificity, or 1 – the true negative rate) on the x-axis across all possible threshold values of a given indicator or classifier. TPR reflects the proportion of actual positive cases (e.g., students who do not meet expectations) that are correctly identified by the indicator. It is calculated as:

$$TPR \ = \ Sensitivity = \frac{True\ Positives}{True\ Positives + False\ Negatives} = \frac{Hits}{Hits + Misses}$$

FPR reflects the proportion of actual negative cases (e.g., students who meet expectations) that are incorrectly identified as at risk of failing. It is calculated as:

$$FPR = 1 - Specificity = \frac{False\ Positives}{False\ Positives + True\ Negatives} = \frac{False\ Alarms}{False\ Alarms + Correct\ Rejections}$$

A strong indicator would have a high TPR rate and a low FPR rate, hugging the top-left corner of the ROC plot.

Figure 2 illustrates two hypothetical ROC curves. Each point along both Curves A and B corresponds to one possible threshold cutoff, such as a particular absence level. At very high thresholds (far bottom-left), only the most extreme cases are flagged, yielding low rates of both true positives and false positives. At very low thresholds (far top-right), nearly all students are flagged, producing high true positive rates but also high false positive rates. Curve A (yellow) represents a stronger classifier: it rises quickly toward the top-left corner, showing that it captures many true positives with relatively few false alarms. Curve B (purple), by contrast, performs less effectively, lying closer to the diagonal line (i.e., when TPR = FPR or y = x), which represents random guessing.

The overall accuracy of each classifier can be summarized by the area under the ROC curve (AUC), which ranges from 0.5 (random guessing) to 1.0 (perfect classification). A higher AUC indicates that the classifier more reliably distinguishes between students who need support and those who do not. The DeLong test (DeLong, DeLong, & Clarke-Pearson, 1988) is a nonparametric method for statistically comparing the areas under two or more correlated ROC curves. Because AUC values are estimated from sample data, they are subject to sampling variability; the DeLong test provides a way to test whether the observed difference in AUCs between classifiers is greater than would be expected by chance. It does so by estimating the covariance of the AUCs using U-statistics, which accounts for the fact that the ROC curves are calculated on the same sample of students. A significant DeLong test result indicates that one predictor has significantly greater diagnostic accuracy than another. This test is widely used in medical diagnostics, as it allows researchers to rigorously assess whether improvements in predictive indicators are meaningful rather than due to random variation, and Henderson & Fantuzzo (2023) used it for testing whether there were statistically significant differences between the AUCs for excused, unexcused, and total absences.

While AUC provides a global summary of a classifier's accuracy across all possible thresholds, applied contexts often require selecting a single cutoff. A commonly used index for this purpose is Youden's J (Youden, 1950), defined as:

$$J = Sensitivity + Specificity - 1 = TPR - FPR$$

Youden's J ranges from 0 to 1, where higher values indicate a better balance between sensitivity (correctly identifying students who do not meet expectations and need additional support) and specificity (correctly identifying students who meet expectations). A value of 1 indicates a perfect signal with no false positives or false negatives, while a value of 0 indicates a signal that is no better than random chance. As shown in Figure 3, the optimal Youden's J corresponds to the red point on the ROC curve that lies farthest above the random guessing diagonal. The TPR and FPR at this point can be translated back into an underlying absence rate or number of days absent, yielding a concrete cutoff. This "optimal" threshold is often chosen because it identifies the point where the tradeoff between sensitivity and specificity is best balanced. As such, this may provide a data-driven way to compare candidate chronic absenteeism thresholds.

Analytical Strategies by Research Question

RQ1: Absenteeism Descriptive Patterns. To examine how absenteeism patterns descriptively differed by measurement choice, we generated descriptive statistics for excused, unexcused, and total absences, both in terms of absolute days missed and rates of absence (days absent divided by days enrolled). We visualized longitudinal absence trajectories using line graphs and provide the table of summary statistics used to create the line graph, disaggregated by key demographic subgroups for each grade from Pre-K to 8, in Appendix S2.

RQ2: ROC Curve Analysis. We used ROC curve analysis to evaluate the diagnostic accuracy of absenteeism measures in predicting whether a student would *Not Meet Expectations* on standardized math and ELA assessments in 8th grade. For each absence measure (days excused, days unexcused, days total, excused absence rate, unexcused absence rate, total absence rate), we estimated the ROC curve and computed the AUC. To test whether observed AUC differences across measures were statistically significant, we applied the DeLong test.

RQ3: Identifying Empirical Thresholds with Youden's J. We calculated the optimal Youden's J index for each of the ROC curves in RQ2 to derive the cutoff that maximized balanced classification accuracy for each grade.

RQ4: Temporal Proximity. To assess how predictive validity varied over time, we extended the ROC curve and Youden's J analyses to standardized math and ELA assessments from grades 3-7. We used absence measures from the same academic year or earlier than the standardized assessment to ensure temporal alignment between predictor and outcome. We then compared cross-grade patterns in diagnostic accuracy and thresholds to the 8th grade results presented in RQ2 and RQ3.

To support transparency and reproducibility, all R code used to estimate ROC curves, compute DeLong tests for AUC differences, and identify optimal cutoffs using Youden's J is available on the first author's GitHub page at github.com/author/ROC.

Results

RQ1: Absenteeism Descriptive Patterns

Figure 3 presents the mean absence rates for excused, unexcused, and total absences from Pre-K to 8th grade, with the exact percentages overall for each grade and by student subgroup provided in Appendix S2. As seen in Figure 3, the total absence rate in Pre-K is driven primarily by excused absences. However, the average excused absence rate declines sharply from the early grades, leveling off after 2nd grade, while unexcused absence rate steadily increases across grades. Around 2nd grade, the lines for excused (turquoise line) and unexcused (purple line) absence rates intersect, marking a shift where unexcused absences begin to surpass excused absences, a pattern that persists and gradually widens through middle school. The average total absence rate (yellow line) decreases after Pre-K and starts rising again in the middle school years; however, Figure 1 shows how total absences can obscure important differences in the underlying types of absences. As a sensitivity check, Appendix S3 recreates Figure 1 using the mean number of days absent, which yields the same overall pattern.

RQ2: ROC Curve Analysis

Figures 4 and 5 display ROC curves for both math and ELA outcomes, comparing absence rate and number-of-days operationalizations of excused, unexcused, and total absences from Pre-K through 8th grade. These curves evaluate how well each indicator classifies students as *Not Meeting Expectations* on the 8th grade MCAS for math and ELA, respectively. The corresponding AUCs and DeLong test comparisons are shown in Table 2. Specifically, Table 2 reports results from pairwise comparisons of the AUCs for different operationalizations of absenteeism, using the DeLong test. Each row compares two absence measures within a given grade level cohort (e.g., days absent vs. absence rate). The first two columns for each outcome show the AUCs for the two measures being compared, followed by the difference in AUC and the corresponding p-value. A positive AUC difference indicates that the first absence measure yielded a higher predictive accuracy than the second; negative differences indicate the reverse. Statistically significant p-values suggest that one measure provides a significantly stronger signal for identifying students at risk on 8th grade math or ELA performance.

Several consistent patterns emerge for both the math and ELA ROC curve figures. In Pre-K (first plot in Figures 4 and 5), absences show virtually no predictive power for later achievement, with ROC curves aligning closely to the diagonal reference line (AUC = 0.5), indicating performance no better than chance. Unexcused absences show the smallest AUC (AUC ≈ 0.50 in Table 2), and the AUC for total and excused absences are only modestly higher (AUC = 0.54-0.57 across math and ELA in Table 2). From 1st to 3rd grade, the accuracy of excused absences declines while unexcused absences become increasingly informative. By 3rd grade, unexcused and total absence curves overlap, with both unexcused and total absences surpassing excused absences in predictive value. For example, in 3rd grade for math, total absence rate has an AUC of 0.646, total number of absence days an AUC of 0.643, unexcused absence rate an AUC of 0.635, and number of unexcused absence days an AUC of 0.634, compared to excused absences which has a lower AUC of 0.543 for both the excused absence rate and number of excused absence days. This divergence is confirmed in the DeLong comparisons, where the differences between both total or unexcused absences and excused absences are statistically significant ($p < 0.0001$).

Predictive accuracy strengthens progressively as students approach middle school, with AUCs climbing across all absence indicators. By 8th grade, both total absence rate (AUC = 0.670) and unexcused absence rate (AUC = 0.673) emerge as the strongest predictors of math performance, and their difference is statistically indistinguishable ($p = 0.686$). Similarly, for

ELA, the 8th grade total absence rate (AUC = 0.639) and unexcused absence rate (AUC = 0.642) show the greatest diagnostic accuracy, again with no statistically significant difference between them ($p = 0.696$). These results suggest that, by adolescence, both total and unexcused absences carry equivalent and meaningful information about students' academic risk, while excused absences provide a weaker signal.

When comparing absence rate versus number of days absent, the two operationalizations yield nearly identical ROC curves across grades. Nonetheless, DeLong tests reveal that absence rate consistently produces slightly higher AUCs than number-of-days measures, with many differences reaching statistical significance even when effect sizes are small. For example, in 8th grade math, the total absence rate (AUC = 0.670) marginally outperforms the total number of days absent (AUC = 0.661), a difference that is statistically significant ($p < 0.0001$). This pattern indicates that absence rates, though only modestly more predictive, offer a consistently more reliable operationalization.

Overall, these results highlight three key insights. First, the predictive value of absences in early childhood is weaker and largely driven by excused absences, but by the upper elementary grades, unexcused absences become more salient. Second, predictive strength steadily increases across grades, peaking in middle school, where absence measures achieve their highest AUCs. Third, absence rates are generally superior to number-of-days measures, albeit by small margins, suggesting that rate-based indicators may be preferable in predictive frameworks such as early warning systems.

RQ3: Identifying Empirical Thresholds with Youden's J

Table 3 reports the optimal absence cutoffs derived from Youden's J, using both the total number of days absent and total absence rate. Based on RQ2, total absences emerged as the most consistent and diagnostically useful signal, so we calculated Youden's J for total absences in particular. We examined both total number of days absent and total absence rate because the number of days absent aligns directly with policy benchmarks like the 18-day cutoff, while absence rate accounts for differences in the length of the school year and offers a standardized measure that facilitates comparisons across grades, schools, or districts. Looking at both days and rates together provides a fuller picture, allowing us to examine whether results are consistent across how absence is operationalized. The slight differences in Youden's J between days and rates reflect the fact that absence rate normalizes absences by the number of enrollment days, while raw day counts can shift thresholds upward or downward depending on variation in school year length and differences in the beginning of student enrollment.

Across grades and subjects, the empirically derived thresholds consistently fall well below the conventional benchmark of 18 days or the widely used 10% policy threshold, with optimal thresholds most consistently clustering in the 3-7% range. For the total number of days absent, the empirically optimal cutoffs range from six days to 17 days, depending on the grade and subject. The earlier grades tend to have a higher optimal threshold. For example, in kindergarten, the Youden's J threshold corresponds to 17 days for math and 16 days for ELA. This drops to around 6 days in 5th grade before climbing back to 11-12 days in 8th grade. The optimal cutoff based on total absence rate supports this pattern. While the absence rate cutoffs tend to suggest a slightly lower number of days within the same grade and subject when translated from a percentage to a number of days, the overall pattern remains consistent.

These results have two important implications. First, although the thresholds for math and ELA are very similar, the differences across grades are more pronounced, suggesting developmental variation in how absenteeism signals academic risk. While implementing grade-specific chronic absenteeism cutoffs would be impractical for schools and confusing for policymakers and families, these findings do highlight how a single, fixed threshold could misclassify risk across developmental stages. Second, if chronic absenteeism is intended to serve as an early warning signal for the need for additional supports, the current policy definition of 10% of school days (about 18 days per year) is likely too high. Our findings imply that substantially lower thresholds would better identify students at risk of poor academic performance in a more timely way while balancing sensitivity and specificity.

At the same time, the values of Youden's J themselves are modest, ranging from 0.08 to 0.26. In medical diagnostics, where ROC methods are most commonly applied, a higher Youden's J indicates better test discrimination, but there is not a threshold universally considered clinically useful. For instance, a study diagnosing respiratory complications in preterm infants reported Youden's J values ranging from 0.207 to 0.421 (Cao et al., 2023), while a study detecting a subtype of stroke reported values from 0.277 to 0.614 (You et al., 2019). Compared with these clinical benchmarks, the values observed here are lower, which is not necessarily unexpected: standardized test performance reflects not only attendance but also prior achievement, instructional quality, socioeconomic conditions, and other factors. Just as effect sizes in education research are typically smaller than those in clinical medicine (Kraft, 2020), even modest predictive signals may still be practically meaningful when they help schools allocate resources and identify students who may otherwise be overlooked. However, more research is needed to benchmark Youden's J in absenteeism research against other educational samples and contexts to provide additional robustness.

These findings highlight both the promise and limitations of using absence thresholds for early warning purposes. On the one hand, Youden's J offers a data-driven method for identifying empirically grounded thresholds. On the other, the relatively low J values suggest that no single absence cutoff is likely to serve as a strong standalone predictor of academic risk, underscoring the importance of integrating attendance with other measures in EWSs.

RQ4: Temporal Proximity

Table 4 shows the top three absence predictors based on largest AUC and corresponding Youden's J thresholds for predicting math and ELA standardized test performance in grades 3-7. Detailed results on all the absence predictors for each grade, mirroring Table 2, can be found in Appendices S4-S8. Across both ELA and math, absences measured closer in time to the grade of the standardized test provide a more reliable diagnostic signal. For example, the best predictors of 3^{rd} grade ELA test performance are 3^{rd} grade unexcused absence rate, number of days unexcused, and total absence rate; the best predictors of 3^{rd} grade math test performance are 3^{rd} grade total absence rate and total number of days absent, followed by 2^{nd} grade total absence rate. The detailed results in Appendices S4-S8 provide further evidence of temporal proximity strengthening predictive validity. The best absence predictors are all either total absences or unexcused absences, and the AUCs of the best absence measures are around 0.60-0.65 across all grades. Both findings are consistent with the 8^{th} grade findings for RQ2. This suggests the timing of measurement matters: absences recorded in the same or immediately preceding school year are consistently more predictive of academic performance than those measured much earlier.

Interestingly, unexcused absences emerged more frequently as top predictors of ELA performance, whereas total absences were more often the strongest predictors of math performance. This divergence may reflect underlying differences in how attendance relates to the types of skills each subject requires. ELA performance may depend more heavily on sustained engagement with classroom discourse, reading, and writing, which are activities that are more sensitive to behavioral disengagement and better captured by unexcused absences. In contrast, mathematics tends to be more cumulative and content-driven, so total instructional time missed, regardless of the reason for absence, may be the more relevant indicator. This pattern aligns with prior research suggesting that cumulative absences have a larger negative effect on math achievement (Goodman, 2014).

The Youden's J thresholds varied across absence type, standardized test, and grade level but were generally consistent with those reported in RQ3, with values centered around 0.20. As expected, thresholds for unexcused absences were lower than those for total absences, indicating that fewer unexcused days are needed to signal academic risk. Notably, all Youden's J thresholds, particularly those based on the total number of days absent, were substantially lower than the current policy benchmarks of 18 days or 10% of the school year, with optimal thresholds clustering in the 3-7% range. This finding reinforces the 8th grade results from RQ3 and provides additional evidence that if chronic absenteeism is used as an early warning indicator, the operational cutoff should be lowered.

Discussion

This study provides new evidence on how chronic absenteeism should be measured and applied in educational research and policy. Overall, we find that measurement choices matter. Excused and unexcused absences carry different weight across developmental stages, and total absence rate is the most reliable and consistent way to operationalize chronic absenteeism across all grades. Because the conventional 10% benchmark misses children at risk of poor academic performance in grades 3-8, lower thresholds should be considered if chronic absenteeism is to function as an effective early warning signal.

Our results differ in important ways from Henderson and Fantuzzo (2023), who concluded that “only unexcused absences provided diagnostic accuracy in predicting later achievement in the early elementary grades” (p. 259). In contrast, we find that unexcused absences meaningfully predict outcomes starting in the later grades, and that total absence measures retain better diagnostic accuracy throughout the Pre-K to 8th grade span. We find that total absences are especially predictive for math performance compared to excused and unexcused absences, consistent with evidence that cumulative instructional time loss has a larger negative effect on math achievement and likely reflects the sequential nature of math learning where new concepts build directly on previously mastered material (Goodman, 2014). The divergence in our results may reflect differences in the age span in our samples, in local attendance policies (e.g., what counts as an excused absence), or in the composition of student populations across districts. These differences underscore the importance of examining attendance measures across multiple contexts before drawing generalizable conclusions.

Additionally, our empirical thresholds derived using Youden’s J suggest that the widely used benchmark of 10% is set too high if the goal of using chronic absenteeism is to identify students in need of support. At the same time, the exact optimal cutoff in our analyses varied across grades and outcomes. We therefore interpret our results as support for a lower, evidence-informed range—approximately 3-7%—rather than a single threshold, and we view replication across multiple districts and student populations as essential before proposing a universal cutoff.

Our findings also highlight the importance of temporal sensitivity. Absences recorded closer in time to standardized test administration were consistently more predictive of performance than those measured several years earlier, suggesting that attendance functions primarily as a proximal behavioral indicator. This conclusion aligns with emerging evidence emphasizing the predictive value of timely attendance indicators, both in EWSs using machine learning algorithms (Wu & Weiland, 2025) and compared to broader student behavior composites for predicting same-year test scores (Wu et al., 2025). This temporal pattern suggest that EWSs may be more effective when they rely on recent data, and integrating this temporal dimension into risk monitoring algorithms could help schools respond more dynamically to emerging patterns of absenteeism.

Taken together, our findings have several potential implications for research, policy, and practice. For research, these findings point to a need for more attention to measurement in studies that examine absenteeism. At a minimum, researchers should probe the sensitivity of their findings to plausible alternative constructs used in the literature and consider ROC-based methods for identifying empirically grounded thresholds. For policy and practice, attention to lower levels of absenteeism and developmental periods may be warranted when designing attendance interventions and EWSs.

Despite these contributions, several limitations bound our conclusions. First, the study is situated in one large, urban district, and findings may not generalize to suburban or rural settings, or to states with different reporting conventions. Future work should replicate these analyses across multiple districts and post-pandemic cohorts. Second, our ROC-based thresholds rely on Youden's J, which balances sensitivity (avoiding missed students who need help) and specificity (avoiding unnecessary intervention for students not at risk). Youden's J is most useful as a cost-neutral, balanced threshold for classification, offering a valuable "first-pass" cutoff before factoring in more specific consequences of misclassification. At the same time, it does not account for the relative costs and benefits of false positives and false negatives, which may be especially relevant in applied contexts such as school interventions. In practice, the cost of a false negative, or overlooking a student truly at risk, may be far greater than that of a false positive, especially if interventions are light-touch supports (e.g., attendance calls). In contrast, for higher-touch or resource-intensive interventions (e.g., targeted mentoring), schools may want to prioritize another statistic such as positive predictive value (PPV) to provide greater assurance that students detected as at risk are truly those at risk of chronic absenteeism. Future research should explicitly model these trade-offs.

Third, our outcome measure was limited to standardized test performance. Although academic achievement is a central educational benchmark, test scores capture only one dimension of student success. As our results for RQ4 suggest, different types of absences may signal risk for different domains, and those predictive of standardized test performance may differ from those associated with socioemotional development or engagement. Our results therefore capture only one facet of the broader insights that attendance patterns can provide. Overall, we echo Henderson & Fantuzzo's (2023) call for additional research in this space.

Finally, it is important to recognize that even the most precise thresholds for chronic absenteeism will have limited utility unless schools have the infrastructure to respond effectively. For an indicator to move from diagnosis to treatment, districts must be equipped with resources to address the root causes of absences. This requires asking not only who is absent, but also why (Chang & Romero, 2008; Ready, 2010). Without ways to identify the underlying causes and targeted supports that align with those causes, attendance indicators risk either perpetuating structural inequities or stigmatizing students flagged as at risk (Ginsburg et al., 2014). Thus, while empirically derived thresholds can sharpen the accuracy of EWSs, their effectiveness depends on being able to translate these early signals into meaningful interventions.

Conclusion

This study underscores the importance of treating chronic absenteeism not as a fixed construct but as a measurement choice with real consequences for research and policy. By showing that total absence rate provides a reliable indicator of future academic achievement level across grades, while empirically derived thresholds suggest the conventional 10% benchmark is too high, we demonstrate both the utility and the limitations of the current chronic absenteeism operationalization as an early warning signal. Our findings point to the need for more nuanced definitions of absenteeism and for future research that connects diagnostic accuracy to the educational supports and resources required to act on these signals effectively.

References

Allensworth, E., Balfanz, R., Rogers, T., & Demarzi, J. (2021). *Absent from school: Understanding and addressing student absenteeism*. Harvard Education Press.

Allison, M. A., Attisha, E., Lerner, M., De Pinto, C. D., Beers, N. S., Gibson, E. J., ... & Weiss-Harrison, A. (2019). The link between school attendance and good health. *Pediatrics*, *143*(2), 1-13.

Aucejo, E. M., & Romano, T. F. (2016). Assessing the effect of school days and absences on test score performance. *Economics of Education Review*, *55*, 70-87.

Balfanz, R., & Byrnes, V. (2012). The importance of being in school: A report on absenteeism in the nation's public schools. *The Education Digest*, *78*(2), 4-9.

Canbolat, Y. (2025). Early warning for whom? Regression discontinuity evidence from the effect of early warning system on student absence. *Educational Evaluation and Policy Analysis*, *47*(2), 477-502.

Chang, H. N., & Romero, M. (2008). Present, Engaged, and Accounted for: The Critical Importance of Addressing Chronic Absence in the Early Grades. Report. *National Center for Children in Poverty*.

DeLong, E. R., DeLong, D. M., & Clarke-Pearson, D. L. (1988). Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics*, 837-845.

Dougherty, S. M. (2018). How measurement and modeling of attendance matter to assessing dimensions of inequality. *Journal of Education for Students Placed at Risk (JESPAR)*, *23*(1-2), 9-23.

Epstein, J. L., & Sheldon, S. B. (2002). Present and accounted for: Improving student attendance through family and community involvement. *The Journal of Educational Research*, *95*(5), 308-318.

Every Student Succeeds Act, 20 U.S.C. § 6301 et seq. (2015).

Faria, A. M., Sorensen, N., Heppen, J., Bowdon, J., Taylor, S., Eisner, R., & Foster, S. (2017). Getting students on track for graduation: Impacts of the early warning intervention and monitoring system after one year. REL 2017-272. *Regional Educational Laboratory Midwest*, 1-82.

Gershenson, S. (2016). Linking teacher quality, student attendance, and student achievement. *Education Finance and Policy*, *11*(2), 125-149.

Gershenson, S., Jacknowitz, A., & Brannegan, A. (2017). Are student absences worth the worry in U.S. primary schools?. *Education Finance and Policy*, *12*(2), 137-165.

Ginsburg, A., Jordan, P., & Chang, H. (2014). Absences add up: How school attendance influences student success. *Attendance Works*.

Goodman, J. (2014). *Flaking out: Student absences and snow days as disruptions of instructional time* (No. w20221). National Bureau of Economic Research.

Gottfried, M. A. (2009). Excused versus unexcused: How student absences in elementary school affect academic achievement. *Educational Evaluation and Policy Analysis*, *31*(4), 392-415.

Gottfried, M. A. (2015). Can center-based childcare reduce the odds of early chronic absenteeism?. *Early Childhood Research Quarterly*, *32*, 160-173.

Heeger, D. (2007). *Signal detection theory.* New York University. Retrieved October 19, 2025, from https://www.cns.nyu.edu/~david/handouts/sdt/sdt.html

Henderson, C. M., & Fantuzzo, J. W. (2023). Challenging the core assumption of chronic absenteeism: Are excused and unexcused absences equally useful in determining academic risk status?. *Journal of Education for Students Placed at Risk (JESPAR)*, *28*(3), 259-293.

Keller-Margulis, M. A., Shapiro, E. S., & Hintze, J. M. (2008). Long-term diagnostic accuracy of curriculum-based measures in reading and mathematics. *School Psychology Review*, *37*(3), 374-390.

Little, M., Drake, T., Cohen-Vogel, L., Gragson, A., & Cadilla, V. (2025). Principal leadership of pre-K programs in elementary schools: Evidence from North Carolina. *Educational Policy*, *39*(5), 984-1018.

Ho, A. D. (2008). The problem with “proficiency”: Limitations of statistics and policy under No Child Left Behind. *Educational Researcher*, *37*(6), 351-360.

Hutt, S., Gardner, M., Duckworth, A. L., & D'Mello, S. K. (2019). Evaluating fairness and generalizability in models predicting on-time graduation from college applications. *International Educational Data Mining Society*.

Kearney, C. A. (2008). School absenteeism and school refusal behavior in youth: A contemporary review. *Clinical Psychology Review*, *28*(3), 451-471.

Massachusetts Department of Elementary and Secondary Education. (2016). *State releases 2016 assessment and accountability results* [Press release]. https://www.doe.mass.edu/news/news.aspx?id=23922

Massachusetts Department of Elementary and Secondary Education. (2018). *Summary of the next-generation district and school accountability system.*

Massachusetts Department of Elementary and Secondary Education. (2025). *Massachusetts Comprehensive Assessment System (MCAS).*

Morrissey, T. W., Hutchison, L., & Winsler, A. (2014). Family income, school attendance, and academic achievement in elementary school. *Developmental psychology*, *50*(3), 741.

Ready, D. D. (2010). Socioeconomic disadvantage, school attendance, and early cognitive development: The differential effects of school exposure. *Sociology of Education*, *83*(4), 271-286.

Rennie Center for Education Research & Policy. (2018). *MCAS next generation: What do the latest MCAS scores mean for our schools?*

Romero, M., & Lee, Y. S. (2007). *A national portrait of chronic absenteeism in the early grades.* New York, NY: National Center for Children in Poverty.

Sheldon, S. B., & Epstein, J. L. (2004). Getting students to school: Using family and community involvement to reduce chronic absenteeism. *School Community Journal*, *14*(2), 39-56.

Silva-Padron, G., & McCann, M. (2023). *50-state comparison: Instructional time policies.* Education Commission of the States. Retrieved from https://www.ecs.org/50state-comparison-instructional-time-policies-2023/

Swaak, T. (2018, July 31). With Nearly 8 Million Students Chronically Absent From School Each Year, 36 States Set Out to Tackle the Problem in New Federal Education Plans. Will It Make a Difference? *The 74 Million*. https://www.the74million.org/article/chronic-absenteeism-36-states-essa-plans/

U.S. Department of Education, Office for Civil Rights. (2016a). *Civil Rights Data Collection (CRDC) for the 2013-14 School Year.* Retrieved from https://www.attendanceworks.org/civil-rights-data-collection-crdc-for-the-2013-14-school-year/

U.S. Department of Education. (2016b). *Issue brief: Early warning systems.* 1-13.

U.S. Department of Education. (2025). *Chronic absenteeism.* Retrieved from https://www.ed.gov/teaching-and-administration/supporting-students/chronic-absenteeism

The White House. (2023, September 13). *Chronic absenteeism and disrupted learning require an all-hands-on-deck approach.* https://www.whitehouse.gov/cea/writtenmaterials/2023/09/13/chronic-absenteeism-and-disrupted-learning-require-an-all-handson-deck-approach/

Wei, W. S. (2024). Exploring patterns of absenteeism from prekindergarten through early elementary school and their associations with children's academic outcomes. *AERA Open*, 10.

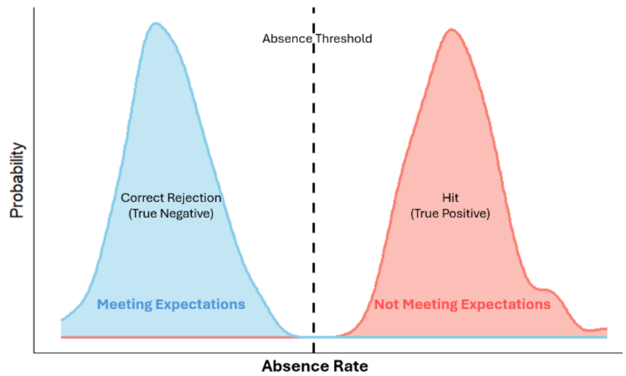
Wu, T., & Weiland, C. (2025). Leveraging modern machine learning to improve early warning systems and reduce chronic absenteeism in early childhood. EdWorkingPaper No. 24-1081. *Annenberg Institute for School Reform at Brown University*.

Wu, T., Weiland, C., Diemer, M. A., Unterman, R., Shapiro, A., & Staines, T. Measuring "Noncognitive" Skills at Scale: Building Longitudinal Student Behavior Composites Using Administrative Data. EdWorkingPaper No. 25-1250. *Annenberg Institute for School Reform at Brown University*.

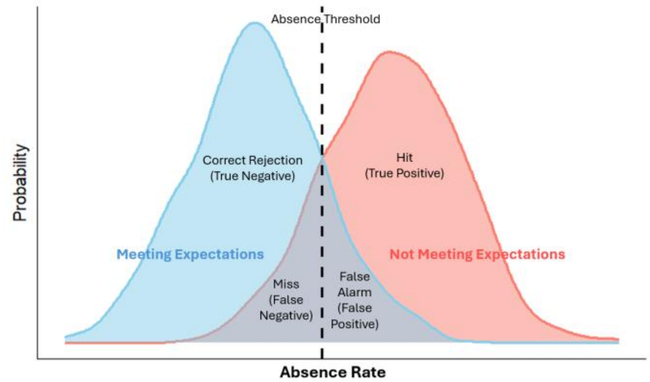
Youden, W. J. (1950). Index for rating diagnostic tests. *Cancer*, *3*(1), 32-35.

Figures & Tables

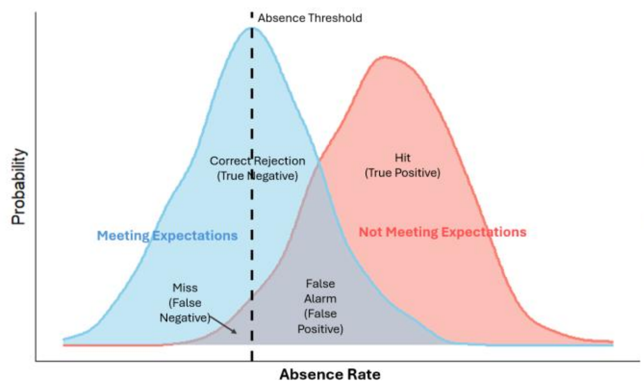
Figure 1. Hypothetical Probability of Occurrence Curves



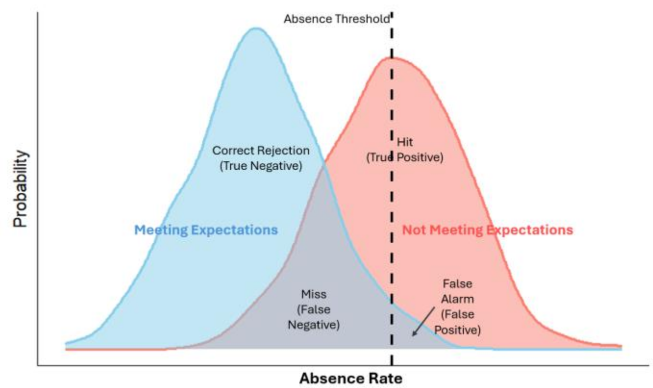
A. “Ideal” Distributions of Meeting and Not Meeting Expectations on Standardized Test



B. Actual Distributions Likely Overlap



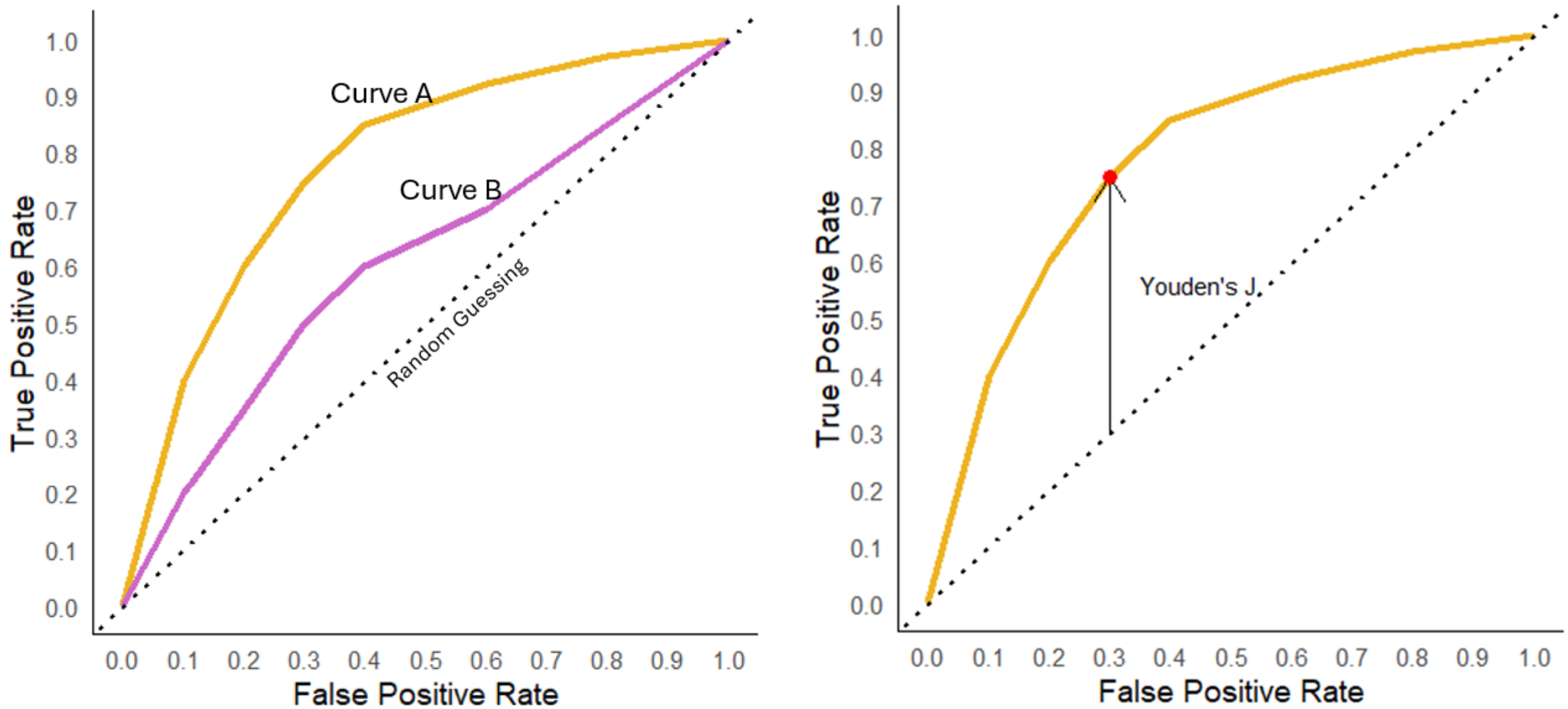
C. Absence Threshold with High Hit Rate and High False Alarm Rate



D. Absence Threshold with Lower False Alarm Rate but High Miss Rate

Note: In Figure 1A, we illustrate the “ideal” ROC scenario, where the distributions of students who meet expectations (in blue in Figure 1A) and those who do not meet expectations (in red in Figure 1A) would be entirely distinct on the absence rate spectrum. In this case, a single threshold on the absence rate (dotted line in Figure 1A) would perfectly separate the two groups, enabling a flawless classification: All students with an absence rate lower than the absence threshold would be correctly classified as Meeting Expectations, and all the students with an absence rate higher than the absence threshold would be correctly classified as Not Meeting Expectations. However, in reality, the actual distribution of these two groups likely overlap, as shown in Figure 1B. Notice that because the distributions overlap, there will inevitably be both misses (false negatives) and false alarms (false positives) no matter what absence threshold we hypothetically choose. For example, setting a low threshold (Figure 1C) captures most students who are likely to not meet expectations, ensuring few students at risk of failing slip through the cracks, but also increases the number of false alarms, which can lead to unnecessary interventions and stigmatization. In contrast, a high threshold (Figure 1D) minimizes false alarms, meaning fewer students are incorrectly flagged. However, the lower false alarm rate comes at the cost of more misses, potentially overlooking students who would benefit from additional supports. The code to create all these hypothetical probability of occurrence curves is on the first author’s Github page, at https://github.com/Author/ROC.

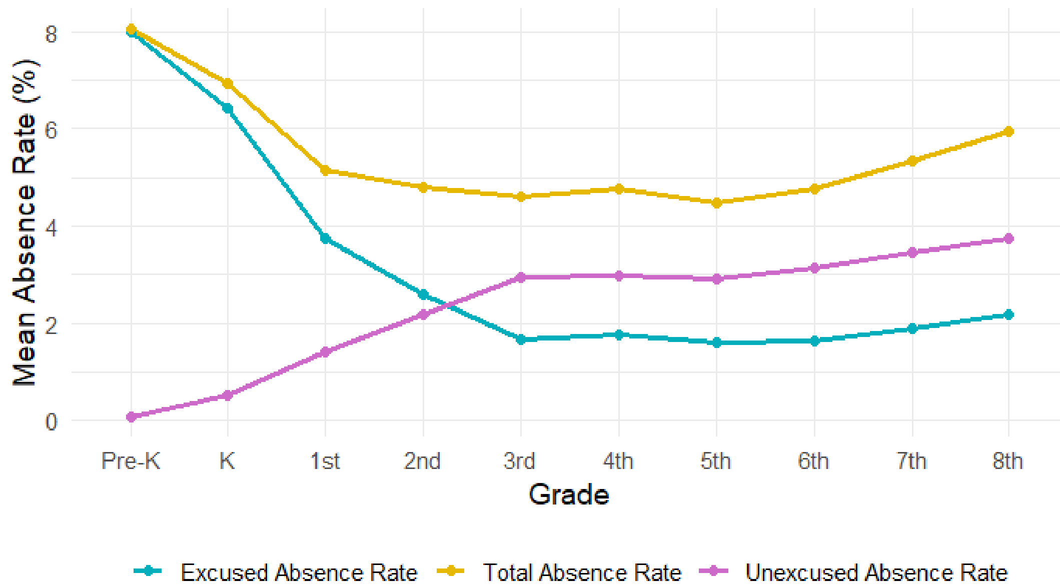
Figure 2. Hypothetical ROC Curves & Youden's J



A. Two Hypothetical ROC Curves **B. Youden's J Point on Curve A**

Note: In Figure 2A, each point along both Curves A and B corresponds to one possible threshold cutoff, such as a particular absence level. At very high thresholds (far bottom-left), only the most extreme cases are flagged, yielding low rates of both true positives and false positives. At very low thresholds (far top-right), nearly all students are flagged, producing high true positive rates but also high false positive rates. Curve A (yellow) represents a stronger classifier: it rises quickly toward the top-left corner, showing that it captures many true positives with relatively few false alarms. Curve B (purple), by contrast, performs less effectively, lying closer to the diagonal line (i.e., when TPR = FPR or y=x), which represents random guessing. Youden's J (shown for Curve A in Figure 2B) is the point on the ROC curve that is farthest from the Random Guessing line along the True Positive Rate axis. The code to create this ROC curve is on the first author's Github page, at https://github.com/Author/ROC.

Figure 3. Mean Absence Rate by Grade



Note: *N* = 6,226 for Pre-K, *N* = 8,014 for Kindergarten, *N* = 7,939 for 1st grade, *N* = 7,819 for 2nd grade, *N* = 7,743 for 3rd grade, *N* = 7,653 for 4th grade, *N* = 7,607 for 5th grade, *N* = 7,479 for 6th grade, *N* = 7,439 for 7th grade, *N* = 7,364 for 8th grade.

Figure 4. ROC Curves for Eighth Grade Math Outcome

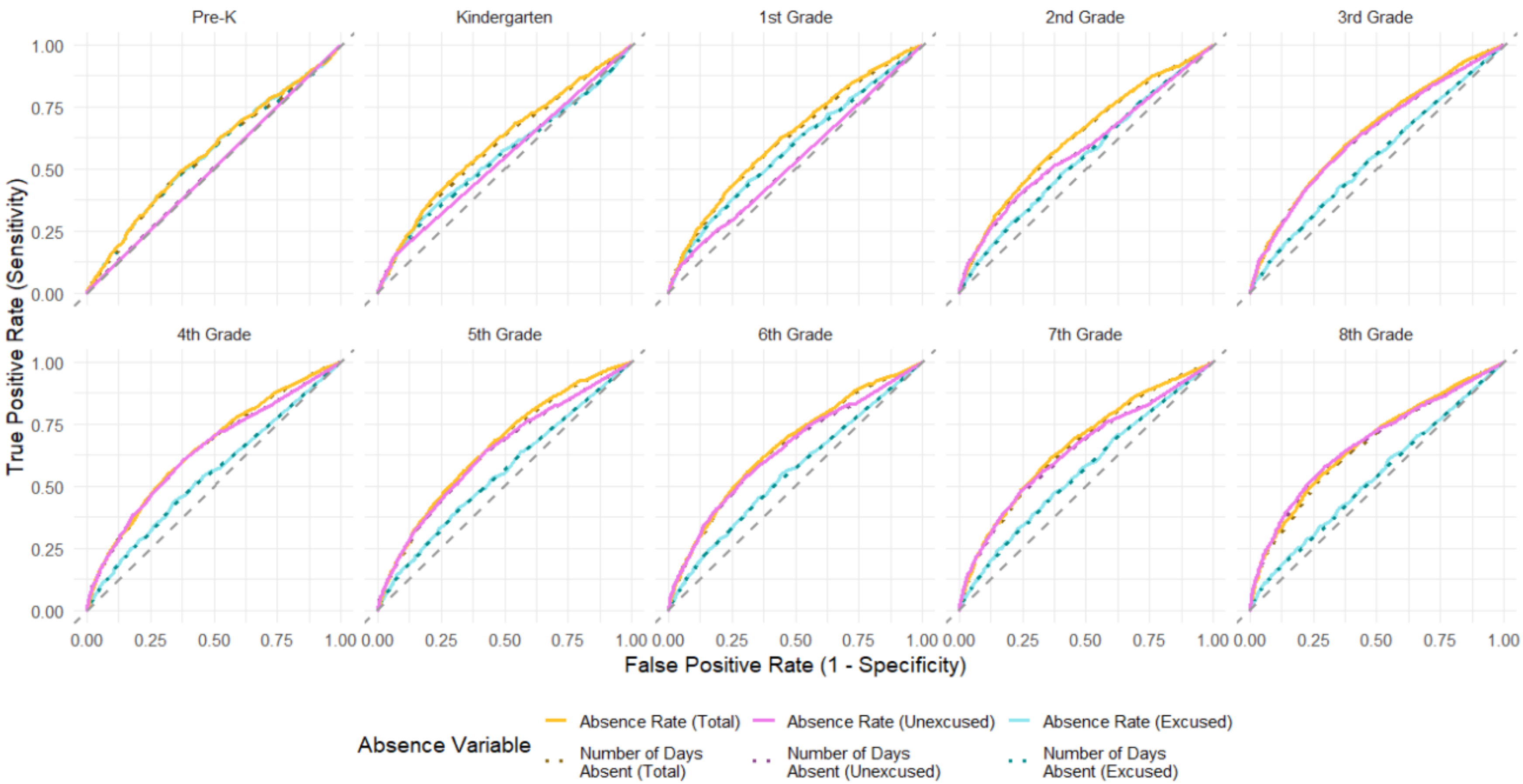


Figure 5. ROC Curves for Eighth Grade ELA Outcome

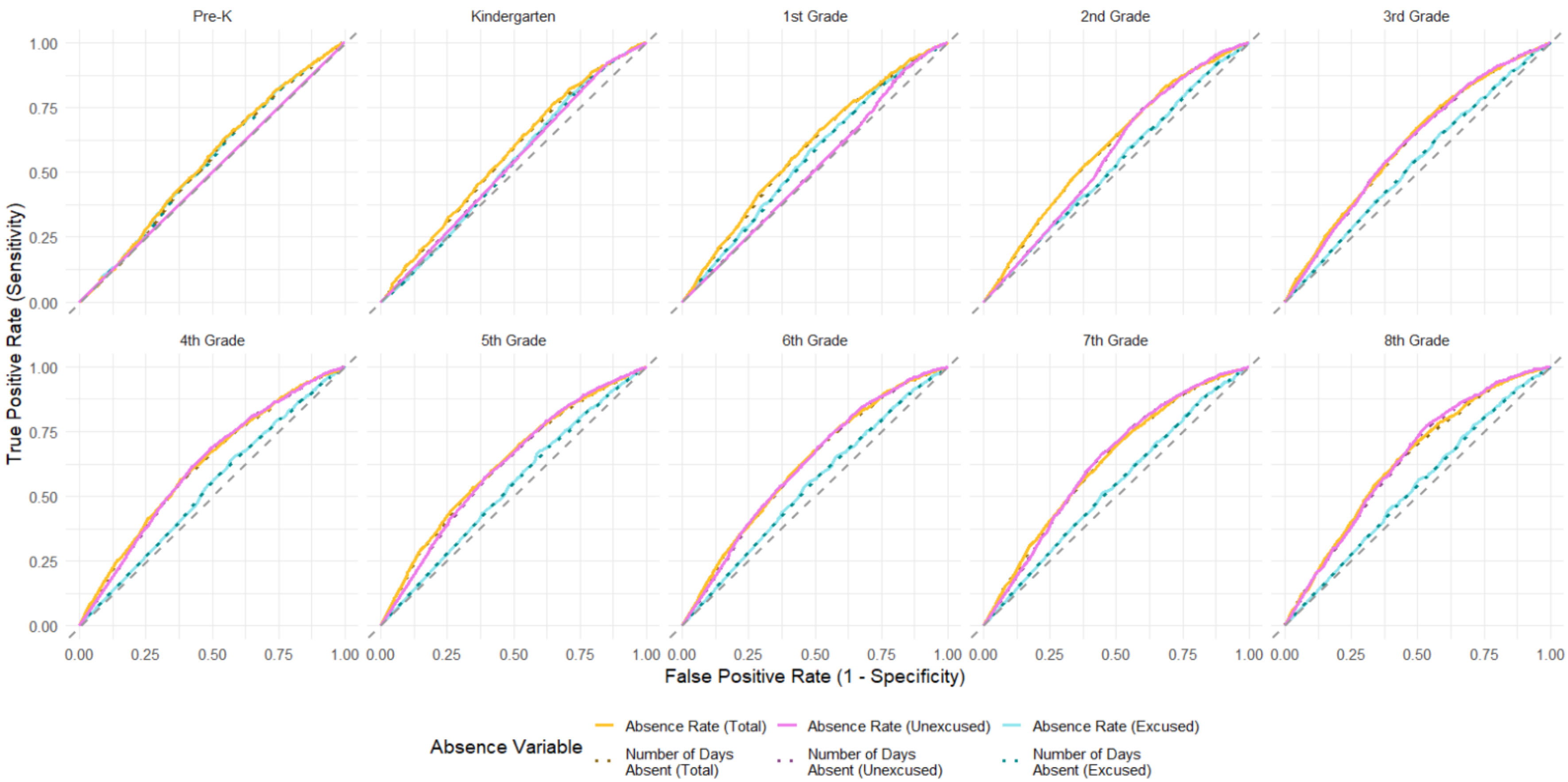


Table 1. Demographic Characteristics of Analytic Samples with Standardized Test Scores by Grade

	8th grade	7th grade	6th grade	5th grade	4th grade	3rd grade
Female (%)	48.37	48.46	48.10	47.98	47.80	48.73
Race/Ethnicity (%)						
Asian	8.65	8.67	8.45	8.43	8.38	8.89
Black	28.52	28.59	28.67	28.48	28.57	28.5
Hispanic	44.72	44.64	44.67	44.64	44.71	43.61
Mixed/Other	2.86	2.85	2.87	2.92	2.93	3.02
White	15.26	15.25	15.35	15.52	15.42	15.97
Home language (%)						
English	49.30	49.29	49.41	49.46	49.54	49.93
Spanish	29.67	29.56	29.71	29.86	29.85	28.82
Other	21.03	21.15	20.88	20.67	20.61	21.25
Country of origin (%)						
USA	93.98	93.84	93.92	93.98	93.84	93.92
Other	6.02	6.16	6.08	6.02	6.16	6.08
Eligible for free/ reduced price lunch (%)	53.92	57.45	60.97	66.41	69.94	69.10
N students	7,145	7,187	7,160	7,148	7,343	6,813

Table 2. DeLong Test Comparisons of AUCs for Various Absenteeism Measures in Predicting 8th Grade Math and ELA Outcomes

		8th Grade Math				8th Grade ELA			
Grade	**Absence Measures Compared**	**AUC for 1st Absence Measure**	**AUC for 2nd Absence Measure**	**AUC Difference**	**p-value**	**AUC for 1st Absence Measure**	**AUC for 2nd Absence Measure**	**AUC Difference**	**p-value**
Pre-K Absenteeism	Days excused vs Excused rate	0.555	0.565	-0.010	<0.05	0.534	0.543	-0.009	<0.01
	Days total vs Total rate	0.557	0.567	-0.010	<0.05	0.536	0.545	-0.009	<0.01
	Days unexcused vs Unexcused rate	0.502	0.502	0.000	0.086	0.500	0.500	0.000	0.310
	Total rate vs Excused rate	0.567	0.565	0.001	0.439	0.545	0.543	0.002	0.308
	Total rate vs Unexcused rate	0.567	0.502	0.065	<0.0001	0.545	0.500	0.045	<0.001
	Unexcused rate vs Excused rate	0.502	0.565	-0.063	<0.0001	0.500	0.543	-0.043	<0.001
Kindergarten Absenteeism	Days excused vs Excused rate	0.545	0.552	-0.007	<0.01	0.528	0.534	-0.006	<0.01
	Days total vs Total rate	0.595	0.603	-0.008	<0.001	0.565	0.571	-0.006	<0.01
	Days unexcused vs Unexcused rate	0.542	0.542	0.000	0.963	0.534	0.534	0.000	0.680
	Total rate vs Excused rate	0.603	0.552	0.051	<0.0001	0.571	0.534	0.037	<0.0001
	Total rate vs Unexcused rate	0.603	0.542	0.062	<0.0001	0.571	0.534	0.037	<0.001
	Unexcused rate vs Excused rate	0.542	0.552	-0.011	0.477	0.534	0.534	0.000	0.993
1st Grade Absenteeism	Days excused vs Excused rate	0.577	0.581	-0.005	<0.01	0.556	0.560	-0.005	<0.01
	Days total vs Total rate	0.621	0.627	-0.006	<0.01	0.589	0.595	-0.006	<0.01
	Days unexcused vs Unexcused rate	0.529	0.529	0.000	<0.05	0.516	0.517	0.000	<0.05
	Total rate vs Excused rate	0.627	0.581	0.046	<0.0001	0.595	0.560	0.034	<0.0001
	Total rate vs Unexcused rate	0.627	0.529	0.098	<0.0001	0.595	0.517	0.078	<0.0001
	Unexcused rate vs Excused rate	0.529	0.581	-0.052	<0.01	0.517	0.560	-0.044	<0.01
2nd Grade Absenteeism	Days excused vs Excused rate	0.548	0.551	-0.002	0.08	0.527	0.529	-0.002	<0.05
	Days total vs Total rate	0.628	0.632	-0.004	<0.05	0.600	0.602	-0.002	0.132
	Days unexcused vs Unexcused rate	0.582	0.583	-0.001	<0.05	0.571	0.571	0.000	0.462
	Total rate vs Excused rate	0.632	0.551	0.081	<0.0001	0.602	0.529	0.074	<0.0001
	Total rate vs Unexcused rate	0.632	0.583	0.049	<0.0001	0.602	0.571	0.031	<0.01

	Unexcused rate vs Excused rate	0.583	0.551	0.032	0.058	0.571	0.529	0.043	<0.01
3rd Grade Absenteeism	Days excused vs Excused rate	0.543	0.543	0.000	0.894	0.534	0.535	-0.001	0.414
	Days total vs Total rate	0.643	0.645	-0.002	0.152	0.609	0.612	-0.003	<0.05
	Days unexcused vs Unexcused rate	0.634	0.635	-0.001	0.08	0.606	0.609	-0.002	<0.01
	Total rate vs Excused rate	0.646	0.543	0.102	<0.0001	0.612	0.535	0.077	<0.0001
	Total rate vs Unexcused rate	0.646	0.635	0.011	0.18	0.612	0.609	0.004	0.623
	Unexcused rate vs Excused rate	0.635	0.543	0.092	<0.0001	0.609	0.535	0.074	<0.0001
4th Grade Absenteeism	Days excused vs Excused rate	0.545	0.546	-0.001	0.196	0.528	0.530	-0.002	0.086
	Days total vs Total rate	0.647	0.650	-0.003	<0.05	0.615	0.619	-0.004	<0.01
	Days unexcused vs Unexcused rate	0.638	0.640	-0.002	0.109	0.614	0.616	-0.002	<0.05
	Total rate vs Excused rate	0.650	0.546	0.104	<0.0001	0.619	0.530	0.089	<0.0001
	Total rate vs Unexcused rate	0.650	0.640	0.010	0.168	0.619	0.616	0.003	0.643
	Unexcused rate vs Excused rate	0.640	0.546	0.094	<0.0001	0.616	0.530	0.086	<0.0001
5th Grade Absenteeism	Days excused vs Excused rate	0.550	0.552	-0.002	<0.05	0.533	0.536	-0.003	<0.01
	Days total vs Total rate	0.650	0.656	-0.006	<0.001	0.621	0.627	-0.006	<0.001
	Days unexcused vs Unexcused rate	0.633	0.636	-0.004	<0.001	0.613	0.617	-0.004	<0.001
	Total rate vs Excused rate	0.656	0.552	0.104	<0.0001	0.627	0.536	0.091	<0.0001
	Total rate vs Unexcused rate	0.656	0.636	0.020	<0.01	0.627	0.617	0.010	0.147
	Unexcused rate vs Excused rate	0.636	0.552	0.085	<0.0001	0.617	0.536	0.081	<0.0001
6th Grade Absenteeism	Days excused vs Excused rate	0.553	0.556	-0.003	<0.01	0.538	0.540	-0.002	<0.05
	Days total vs Total rate	0.656	0.660	-0.005	<0.001	0.618	0.622	-0.004	<0.001
	Days unexcused vs Unexcused rate	0.641	0.645	-0.004	<0.001	0.617	0.620	-0.003	<0.01
	Total rate vs Excused rate	0.660	0.556	0.105	<0.0001	0.622	0.540	0.082	<0.0001
	Total rate vs Unexcused rate	0.660	0.645	0.015	<0.05	0.622	0.620	0.002	0.768
	Unexcused rate vs Excused rate	0.645	0.556	0.089	<0.0001	0.620	0.540	0.080	<0.0001
7th Grade Absenteeism	Days excused vs Excused rate	0.559	0.563	-0.004	<0.01	0.535	0.536	-0.001	0.189
	Days total vs Total rate	0.658	0.665	-0.007	<0.001	0.630	0.632	-0.002	0.065

	Days unexcused vs Unexcused rate	0.640	0.644	-0.004	<0.01	0.632	0.633	-0.001	0.177
	Total rate vs Excused rate	0.665	0.563	0.102	<0.0001	0.632	0.536	0.096	<0.0001
	Total rate vs Unexcused rate	0.665	0.644	0.021	<0.01	0.632	0.633	-0.001	0.881
	Unexcused rate vs Excused rate	0.644	0.563	0.082	<0.0001	0.633	0.536	0.097	<0.0001
8th Grade Absenteeism	Days excused vs Excused rate	0.532	0.538	-0.005	<0.0001	0.532	0.535	-0.003	<0.001
	Days total vs Total rate	0.661	0.670	-0.009	<0.0001	0.633	0.639	-0.006	<0.0001
	Days unexcused vs Unexcused rate	0.666	0.673	-0.007	<0.0001	0.638	0.642	-0.004	<0.01
	Total rate vs Excused rate	0.670	0.538	0.133	<0.0001	0.639	0.535	0.104	<0.0001
	Total rate vs Unexcused rate	0.670	0.673	-0.003	0.686	0.639	0.642	-0.003	0.696
	Unexcused rate vs Excused rate	0.673	0.538	0.136	<0.0001	0.642	0.535	0.107	<0.0001

Note: This table reports results from pairwise comparisons of the area under the ROC curve (AUC) for different operationalizations of absenteeism, using the DeLong test for correlated ROC curves. Each row compares two absence measures within a given grade level cohort (e.g., days absent vs. absence rate). The first two columns for each outcome show the AUCs for the two measures being compared, followed by the difference in AUC and the corresponding p-value. A positive AUC difference indicates that the first absence measure yielded a higher predictive accuracy than the second; negative differences indicate the reverse. Statistically significant p-values suggest that one measure provides a significantly stronger signal for identifying students at risk on 8th grade math or ELA performance.

Table 3. Youden's J and Optimal Total Absence Cutoffs for Predicting 8th Grade MCAS Math and ELA Performance

Grade	8th Grade MCAS Subject	Youden's J (Total Days Absent)	Youden's J Threshold (Total Days Absent)	Youden's J (Total Absence Rate)	Youden's J Threshold (Total Absence Rate)	Equivalent Days from Rate Threshold
Pre-K	Math	0.115	13 days	0.121	7.34%	13 days
	ELA	0.079	15 days	0.088	7.31%	13 days
Kindergarten	Math	0.147	17 days	0.168	7.44%	13 days
	ELA	0.102	16 days	0.116	8.55%	15 days
1st Grade	Math	0.177	10 days	0.187	5.01%	9 days
	ELA	0.140	10 days	0.148	5.04%	9 days
2nd Grade	Math	0.191	10 days	0.200	4.46%	8 days
	ELA	0.148	7 days	0.147	3.34%	6 days
3rd Grade	Math	0.217	8 days	0.221	4.41%	8 days
	ELA	0.169	9 days	0.174	4.48%	8 days
4th Grade	Math	0.226	9 days	0.230	4.75%	9 days
	ELA	0.175	7 days	0.184	3.86%	7 days
5th Grade	Math	0.212	6 days	0.223	3.32%	6 days
	ELA	0.179	6 days	0.189	3.32%	6 days
6th Grade	Math	0.239	9 days	0.245	4.52%	8 days
	ELA	0.170	9 days	0.180	3.88%	7 days
7th Grade	Math	0.244	9 days	0.252	4.75%	9 days
	ELA	0.195	10 days	0.198	5.45%	10 days
8th Grade	Math	0.250	12 days	0.264	6.18%	11 days
	ELA	0.203	11 days	0.218	5.54%	10 days

Note: This table reports Youden's J statistics and the corresponding optimal absenteeism cutoffs for predicting 8th grade math and ELA standardized testing achievement level of *Not Meeting Expectations*. Youden's J represents the maximum vertical distance between the ROC curve and the random chance line. For each grade level and subject, we report results for both total days absent and total absence rate. The optimal cutoff is the number of days or absence rate percentage at which Youden's J is maximized. To facilitate comparison, absence rate cutoffs are also expressed in terms of their equivalent number of days absent (final column).

Table 4. Top 3 Absence Measures and Optimal Thresholds for Predicting Math and ELA Standardized Test Performance in Grades 3-7

	Math				ELA			
Standardized Test Grade	**Top Absence Predictors**	**AUCs**	**Youden's J**	**Youden's J Threshold (Number of Days)**	**Top Absence Predictors**	**AUCs**	**Youden's J**	**Youden's J Threshold (Number of Days)**
Grade 3	Grade 3 Total Absence Rate	0.658	0.246	9	Grade 3 Unexcused Absence Rate	0.619	0.179	3
Grade 3	Grade 3 Total Days Absent	0.654	0.239	9	Grade 3 Unexcused Absence Days	0.618	0.178	5
Grade 3	Grade 2 Total Absence Rate	0.642	0.210	11	Grade 3 Total Absence Rate	0.614	0.170	10
Grade 4	Grade 4 Total Absence Rate	0.672	0.251	8	Grade 4 Unexcused Absence Rate	0.642	0.220	5
Grade 4	Grade 4 Total Days Absent	0.667	0.243	7	Grade 4 Unexcused Absence Days	0.639	0.216	5
Grade 4	Grade 4 Unexcused Absence Rate	0.658	0.250	4	Grade 4 Total Absence Rate	0.629	0.189	7
Grade 5	Grade 5 Total Absence Rate	0.644	0.208	7	Grade 5 Unexcused Absence Rate	0.630	0.204	4
Grade 5	Grade 4 Total Absence Rate	0.641	0.207	8	Grade 5 Unexcused Absence Days	0.627	0.202	5
Grade 5	Grade 5 Total Days Absent	0.637	0.196	7	Grade 4 Unexcused Absence Rate	0.624	0.181	5
Grade 6	Grade 5 Total Absence Rate	0.644	0.209	6	Grade 5 Total Absence Rate	0.616	0.165	7
Grade 6	Grade 5 Total Days Absent	0.636	0.198	5	Grade 4 Total Absence Rate	0.615	0.182	9
Grade 6	Grade 4 Total Absence Rate	0.635	0.203	9	Grade 6 Total Absence Rate	0.612	0.173	7
Grade 7	Grade 7 Total Absence Rate	0.644	0.212	8	Grade 7 Total Absence Rate	0.649	0.228	11
Grade 7	Grade 7 Total Days Absent	0.638	0.204	9	Grade 7 Unexcused Absence Days	0.645	0.223	11
Grade 7	Grade 6 Total Absence Rate	0.635	0.207	7	Grade 7 Unexcused Absence Rate	0.643	0.223	8

Note: This table reports the top 3 absence predictors for each grade's standardized test outcome based on largest AUC. Youden's J statistics and the corresponding optimal absenteeism cutoffs are for predicting each grade's math and ELA standardized testing achievement level of *Not Meeting Expectations*. Youden's J represents the maximum vertical distance between the ROC curve and the random chance line. The optimal Youden's J Threshold is the number of days at which Youden's J is maximized. Youden's J Thresholds for absence rates have been converted to the equivalent number of days missed based on a n180-day school year.

Appendix

Appendix S1. Technical Notes on MCAS Comparability

This appendix synthesizes information relevant to understanding the comparability of Massachusetts standardized assessments used in this study: Legacy MCAS, Next-Generation (Next-Gen) MCAS, and PARCC. While the state provides procedures for equating these tests, including mode-adjusted theta scores, differences in test domains, item formats, administration mode, and proficiency categories may affect their comparability. The following sections summarize the key distinctions most relevant for interpreting our analyses.

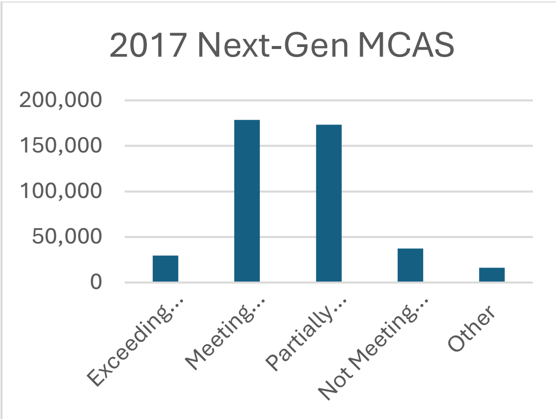
Main Differences

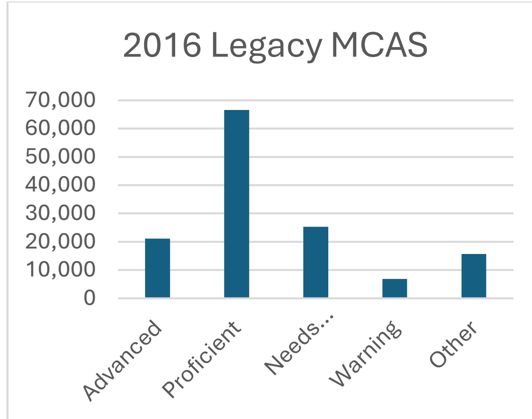
Next-Gen MCAS, introduced in 2017 for grades 3-8 (and later for grade 10), represented a fusion of Legacy MCAS with updates inspired by PARCC. While the underlying curriculum frameworks remained largely stable across the two tests, there were notable changes in test design. First, Next-Gen introduced a greater emphasis on writing, with essay components required in every grade, whereas Legacy MCAS included essays only in grades 4, 7, and 10. In mathematics, algebra and geometry content was given increased weight, especially in grade 10. Second, the test mode transitioned from paper-and-pencil to primarily computer-based delivery by 2019, with new online-only item types such as drag-and-drop, hot text, and multi-select responses. PARCC is even more difficult to compare, as its technical reports group questions differently. However, it has an even greater focus on writing in the ELA portion than either Legacy or Next-Gen. It does share with Next-Gen MCAS many of its new question types, but it is hard to determine the exact number of each type in a given test. These shifts suggest that while the same content standards and curriculum frameworks underlie both tests, the exact methods for testing that knowledge and assessment format of that content differ in ways that may influence student performance and proficiency classification.

Changes in Proficiency Levels

Alongside changes in test design, Next-Gen MCAS introduced revised performance categories intended to align more closely with expectations for college and career readiness. Legacy MCAS used four categories: Advanced, Proficient, Needs Improvement, and Warning/Failing. Next-Gen MCAS shifted to Exceeding Expectations, Meeting Expectations, Partially Meeting Expectations, and Not Meeting Expectations. PARCC, meanwhile, used a five-level scale. Although these categories aim to capture broadly comparable performance distinctions, they are not identical. For example, in order to get students college and career ready, the proficiency tiers for MCAS changed from Legacy MCAS to Next-Gen MCAS. Legacy MCAS was more lenient, bucketing a much higher proportion of students in the proficient category than Next-Gen MCAS does. (Note that 2016 Legacy MCAS had fewer students overall due to PARCC taking some students, but the proportion is similar throughout our Legacy MCAS years).

Figure S1.1. Changes in Distribution for 2016 and 2017 MCAS Categories





As seen in the above Figure S1.1., these changes highlight that shifts in category distributions may reflect redefinition rather than substantive changes in student performance.

Differences Between Legacy and Next-Gen MCAS
In 2017, MA DESE transitioned from the Legacy MCAS to the Next-Gen MCAS for grades 3-8, with grade 10 following in 2021. While the curriculum standards underlying the two assessments remained broadly similar, important changes were introduced in test design and emphasis. Key changes included: A greater emphasis on writing, with essay responses included for all grades (rather than only in grades 4, 7, and 10); A stronger focus on algebra and geometry in mathematics, with these domains comprising a larger share of the tested material; A shift from paper-and-pencil to online administration, with widespread adoption of technology-enhanced item types such as drag-and-drop, hot text, and multiple-select questions; A redefinition of proficiency levels and redistribution of students across these categories.

ELA assessments in Legacy MCAS were weighted heavily toward reading comprehension, especially in grades 3, 5, 6, and 8, where nearly all raw points came from reading. In Next-Gen MCAS, writing and language standards carry much greater weight: for example, writing accounts for 20% of raw points in grades 5-8, where it was previously not assessed The following Table S1.1 details how that changes score calculation, using the 2014 and 2023 exams. Each value is the target percentage of raw points scored in each domain. On an individual test the actual percentage may vary by a few percentage points. This chart shows that, for grades 3, 5, 6, and 8, the actual tested material is very different, with no writing ability tested. Also, the Next-Gen MCAS has less of an emphasis on reading comprehension – the for grades 3, 5, 6, and 8, the Legacy MCAS is almost entirely a reading comprehension test.

Table S1.1. ELA Assessment Score Calculation Changes

Grade	Legacy Reading	Next-Gen Reading	Legacy Language	Next-Gen Language	Legacy Writing	Next-Gen Writing
3	85	65	15	25	0	10
4	64	65	8	25	28	10
5	88	55	12	25	0	20
6	88	55	12	25	0	20
7	64	55	8	25	28	20
8	88	55	12	25	0	20

10	64	55	8	25	28	20

In mathematics, while the domains remained consistent with Massachusetts curriculum frameworks, algebra and geometry received increased emphasis, rising from 60% of the grade 10 test under Legacy to 70-80% under Next-Gen.

Table S1.2. Math Assessment Score Calculation Changes

Reporting Category	Legacy % of raw score points	Next-Gen % of raw score points
Number and Quantity	20	15
Algebra and Functions	30	35
Geometry	30	35
Statistics and Probability	20	15

Comparisons with PARCC

PARCC, administered in Massachusetts in 2015 and 2016, was based on the Common Core State Standards and differed notably from Legacy MCAS. PARCC placed a much greater emphasis on writing in ELA, allocating 30-45% of raw points to composition compared to Legacy MCAS, where some grades had no writing at all. The following charts compare how PARCC ELA domains compare to Legacy (2016) and Next-Gen (2023) MCAS.

Table S1.3. ELA Assessment Score Calculation Changes between PARCC and Legacy

Grade	PARCC Reading	Legacy Reading	PARCC Language	Legacy Language	PARCC Writing	Legacy Writing
3	41	81	27	19	45	0
4	41	85	24	15	48	0
5	38	88	25	12	37	0
6	41	88	26	12	31	0
7	40	90	28	10	31	0
8	41	87	27	13	31	0
10	45	65	26	7	31	28

Table S1.4. ELA Assessment Score Calculation Changes between PARCC and Next-Gen

Grade	PARCC Reading	Next-Gen Reading	PARCC Language	Next-Gen Language	PARCC Writing	Next-Gen Writing
3	41	65	27	25	45	10
4	41	65	24	25	48	10
5	38	55	25	25	37	20
6	41	55	26	25	31	20
7	40	55	28	25	31	20
8	41	55	27	25	31	20
10	45	55	26	25	31	20

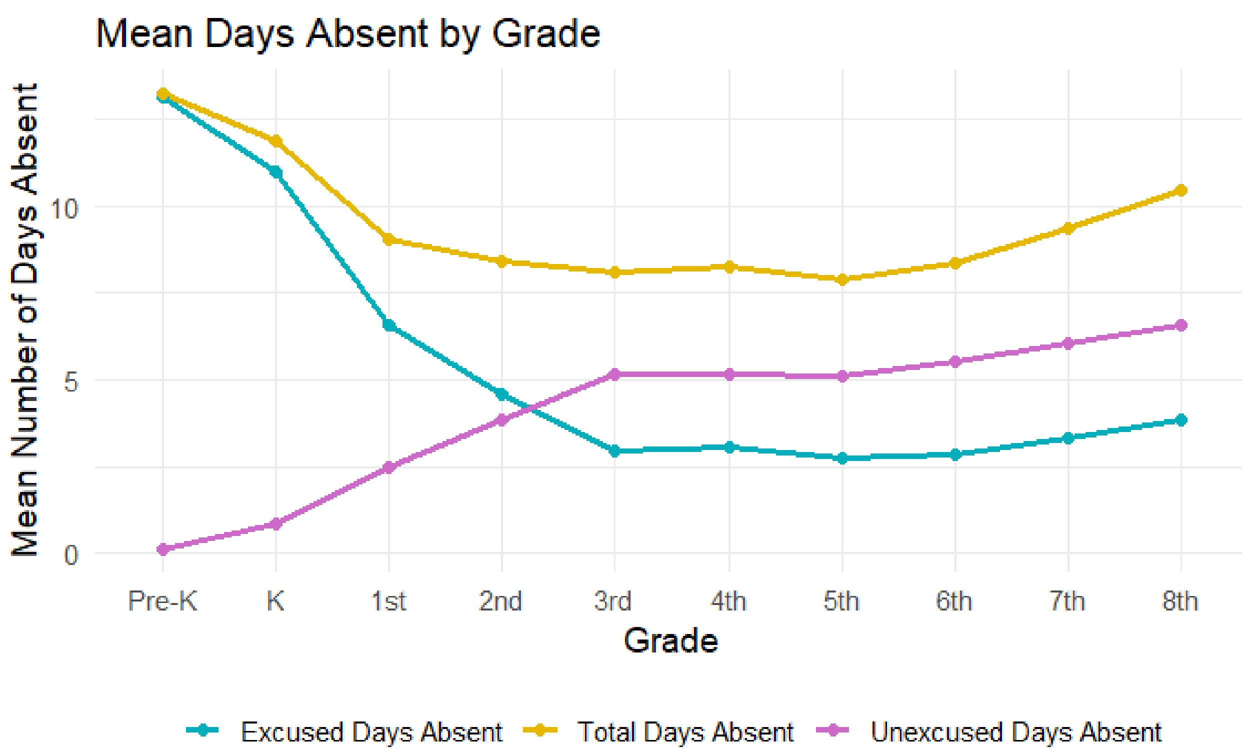
The math portion is more difficult to compare. Rather than specifying the different domains and how many points worth of questions each has, the PARCC technical notes split every grade into “Major Content”, “Additional & Supporting Content”, “Expressing Mathematical Reasoning”, and “Modeling and Applications”. The curriculum frameworks suggest that what falls under individual domains is similar

Appendix S2. Absent Rates for Each Grade Level Overall and by Student Subgroup

		Total	Sex	Race					Home Language			FRL
		Overall	Male	Asian	Black	Hispanic	Mixed/ Other	White	Home Language: English	Home Language: Spanish	Home Language: Other	Eligible for Free/ Reduced Price Lunch
Pre-K Absence Rates	Total	8.06 (8.29)	8.17 (8.55)	6.9 (7.19)	8.13 (8.88)	8.83 (8.63)	8 (8.66)	6.71 (6.4)	8.21 (8.85)	8.62 (8.13)	6.94 (7.19)	8.64 (8.49)
	Excused	7.99 (8.31)	8.1 (8.56)	6.89 (7.2)	8.02 (8.92)	8.78 (8.66)	7.8 (8.66)	6.7 (6.4)	8.13 (8.88)	8.57 (8.16)	6.92 (7.2)	8.56 (8.53)
	Unexcused	0.06 (0.64)	0.07 (0.69)	0.02 (0.21)	0.11 (0.81)	0.05 (0.55)	0.2 (1.53)	0.01 (0.13)	0.08 (0.74)	0.05 (0.55)	0.02 (0.25)	0.09 (0.78)
Kindergarten Absence Rates	Total	6.92 (6.91)	6.9 (6.92)	5.24 (5.52)	6.87 (7.02)	7.7 (7.45)	7.33 (7.97)	5.58 (4.86)	7.05 (6.99)	7.7 (7.26)	5.37 (5.81)	7.54 (7.36)
	Excused	6.41 (6.66)	6.39 (6.69)	4.73 (5.26)	6.32 (6.81)	7.13 (7.13)	6.84 (7.9)	5.33 (4.84)	6.54 (6.77)	7.19 (6.99)	4.93 (5.54)	6.97 (7.13)
	Unexcused	0.51 (2.77)	0.51 (2.64)	0.51 (2.6)	0.55 (2.71)	0.56 (3.16)	0.5 (2.46)	0.25 (1.5)	0.51 (2.67)	0.52 (3.02)	0.44 (2.57)	0.57 (2.92)
1st Grade Absence Rates	Total	5.14 (4.95)	5.21 (5.08)	2.91 (3.33)	5.2 (5.12)	5.77 (5.23)	5.1 (5.9)	4.38 (3.69)	5.48 (5.17)	5.64 (5)	3.43 (3.78)	5.62 (5.29)
	Excused	3.74 (4.35)	3.79 (4.43)	2.27 (2.93)	3.68 (4.36)	4.16 (4.63)	4.04 (5.65)	3.35 (3.5)	4 (4.53)	4.01 (4.43)	2.57 (3.37)	4.07 (4.62)
	Unexcused	1.4 (3.36)	1.41 (3.47)	0.64 (2.02)	1.53 (3.69)	1.61 (3.61)	1.06 (2.97)	1.03 (2.36)	1.48 (3.56)	1.63 (3.61)	0.86 (2.34)	1.55 (3.71)
2nd Grade Absence Rates	Total	4.78 (4.73)	4.89 (4.85)	2.58 (3.13)	5.03 (5.21)	5.3 (4.84)	4.88 (4.79)	3.98 (3.54)	5.18 (5.13)	5.18 (4.48)	3.09 (3.44)	5.25 (5.09)
	Excused	2.6 (3.65)	2.7 (3.73)	1.53 (2.47)	2.65 (3.81)	2.78 (3.81)	2.87 (4.01)	2.51 (3.16)	2.86 (4.02)	2.64 (3.41)	1.77 (2.68)	2.79 (3.92)
	Unexcused	2.18 (3.69)	2.19 (3.71)	1.04 (2.32)	2.39 (4.22)	2.52 (3.82)	2 (3.61)	1.47 (2.36)	2.32 (3.95)	2.53 (3.76)	1.32 (2.59)	2.46 (4.02)
3rd Grade Absence Rates	Total	4.61 (4.85)	4.75 (5.02)	2.05 (2.48)	4.76 (4.96)	5.13 (4.97)	5.26 (6.75)	4.06 (4.33)	5.1 (5.2)	4.97 (4.81)	2.81 (3.42)	5 (4.94)
	Excused	1.67 (2.75)	1.73 (2.91)	0.91 (1.75)	1.72 (2.78)	1.66 (2.61)	2.1 (3.13)	1.94 (3.32)	1.93 (3.15)	1.56 (2.42)	1.11 (1.85)	1.7 (2.78)
	Unexcused	2.94 (4.02)	3.02 (4.08)	1.15 (1.83)	3.04 (4.17)	3.47 (4.22)	3.16 (6.2)	2.12 (2.93)	3.17 (4.22)	3.41 (4.17)	1.69 (2.95)	3.3 (4.07)
4th Grade Absence Rates	Total	4.75 (5.28)	4.83 (5.19)	2.41 (5.66)	4.98 (5.5)	5.29 (5.34)	4.25 (4.21)	4.12 (4.13)	5.12 (5.3)	5.15 (5.1)	3.14 (5.07)	5.21 (5.67)
	Excused	1.77 (3.49)	1.77 (3.38)	1.33 (4.24)	1.8 (3.65)	1.74 (3.42)	1.8 (3.23)	2.01 (2.92)	1.94 (3.42)	1.64 (3.29)	1.44 (3.84)	1.83 (3.87)
	Unexcused	2.98 (3.94)	3.06 (3.92)	1.08 (3.58)	3.18 (4.13)	3.55 (4)	2.45 (3.29)	2.11 (3.11)	3.19 (4.07)	3.5 (3.9)	1.7 (3.15)	3.38 (4.04)
5th Grade Absence Rates	Total	4.49 (5.23)	4.64 (5.47)	2.04 (4.1)	4.61 (5.29)	5.03 (5.57)	4.41 (5.76)	4.06 (3.99)	4.91 (5.36)	4.81 (4.98)	2.88 (4.72)	5 (5.75)
	Excused	1.59 (3.24)	1.67 (3.57)	1.02 (3.35)	1.63 (3.36)	1.6 (3.33)	1.44 (2.16)	1.85 (2.78)	1.8 (3.39)	1.49 (3.04)	1.16 (3)	1.65 (3.62)
	Unexcused	2.9 (3.96)	2.97 (3.98)	1.01 (1.89)	2.99 (4.16)	3.43 (4.18)	2.98 (5.54)	2.21 (2.88)	3.11 (4.05)	3.33 (3.78)	1.72 (3.52)	3.35 (4.26)
6th Grade Absence Rates	Total	4.76 (5.31)	4.91 (5.16)	1.74 (2.82)	5.04 (5.52)	5.25 (5.57)	5.12 (6.48)	4.39 (4.26)	5.34 (5.72)	5.08 (5.09)	2.79 (3.82)	5.24 (5.69)
	Excused	1.63 (2.83)	1.66 (2.7)	0.68 (1.44)	1.84 (3.37)	1.54 (2.57)	1.76 (3.22)	1.98 (2.78)	1.91 (3.07)	1.47 (2.47)	1.04 (2.17)	1.59 (2.77)
	Unexcused	3.14 (4.32)	3.25 (4.24)	1.06 (2.28)	3.2 (4.34)	3.72 (4.76)	3.36 (4.78)	2.41 (3.12)	3.43 (4.67)	3.61 (4.3)	1.75 (3.07)	3.65 (4.77)
7th Grade Absence Rates	Total	5.35 (6.04)	5.56 (6.18)	2.19 (4.21)	5.54 (6.02)	5.91 (6.22)	5.7 (6.66)	5.04 (5.71)	5.94 (6.45)	5.76 (5.94)	3.38 (4.6)	5.94 (6.36)
	Excused	1.89 (3.27)	1.91 (3.19)	0.84 (1.77)	1.97 (3.6)	1.88 (3.1)	1.92 (3.11)	2.36 (3.61)	2.19 (3.67)	1.73 (2.9)	1.35 (2.5)	1.8 (2.97)

	Unexcused	3.46 (4.87)	3.65 (5.11)	1.36 (3.33)	3.58 (4.78)	4.03 (5.17)	3.78 (5.44)	2.68 (4.29)	3.74 (5.16)	4.03 (5.04)	2.04 (3.53)	4.14 (5.33)
8th Grade Absence Rates	Total	5.93 (6.62)	6.05 (6.67)	2.58 (3.59)	6.12 (6.69)	6.6 (7.12)	5.63 (6.07)	5.53 (5.72)	6.46 (6.77)	6.45 (6.62)	3.9 (5.36)	6.75 (7.25)
	Excused	2.18 (3.73)	2.17 (3.5)	1.25 (2.5)	2.23 (3.58)	2.17 (4.04)	2.19 (3.47)	2.68 (3.58)	2.49 (3.83)	2.04 (3.41)	1.62 (3.08)	2.16 (3.66)
	Unexcused	3.74 (5.21)	3.88 (5.5)	1.33 (2.11)	3.89 (5.4)	4.44 (5.65)	3.44 (4.47)	2.85 (4.25)	3.97 (5.36)	4.41 (5.44)	2.28 (4.1)	4.59 (5.89)

Appendix S3. Mean Days Absent by Grade



Note: N = 6,226 for Pre-K, N = 8,014 for Kindergarten, N = 7,939 for 1st grade, N = 7,819 for 2nd grade, N = 7,743 for 3rd grade, N = 7,653 for 4th grade, N = 7,607 for 5th grade, N = 7,479 for 6th grade, N = 7,439 for 7th grade, N = 7,364 for 8th grade.

Appendix S4. DeLong Test Comparisons of AUCs for Various Absenteeism Measures in Predicting 3rd Grade Math and ELA Outcomes

		Math				ELA			
Grade	Absence Measures Compared	AUC for 1st Absence Measure	AUC for 2nd Absence Measure	AUC Difference	p-value	AUC for 1st Absence Measure	AUC for 2nd Absence Measure	AUC Difference	p-value
Pre-K Absenteeism	Days excused vs Excused rate	0.591	0.606	-0.015	<0.01	0.562	0.571	-0.009	<0.05
	Days total vs Total rate	0.589	0.604	-0.015	<0.01	0.559	0.569	-0.009	<0.05
	Days unexcused vs Unexcused rate	0.491	0.491	0.000	0.152	0.494	0.494	0.000	0.376
	Total rate vs Excused rate	0.604	0.606	-0.002	0.187	0.569	0.571	-0.002	0.185
	Total rate vs Unexcused rate	0.604	0.491	0.113	<0.0001	0.569	0.494	0.075	<0.0001
	Unexcused rate vs Excused rate	0.491	0.606	-0.115	<0.0001	0.494	0.571	-0.077	<0.0001
Kindergarten Absenteeism	Days excused vs Excused rate	0.594	0.597	-0.004	<0.05	0.593	0.598	-0.005	<0.05
	Days total vs Total rate	0.594	0.598	-0.004	<0.05	0.589	0.594	-0.005	<0.05
	Days unexcused vs Unexcused rate	0.494	0.494	0.000	0.542	0.489	0.489	0.000	0.4
	Total rate vs Excused rate	0.598	0.597	0.001	0.846	0.594	0.598	-0.004	0.485
	Total rate vs Unexcused rate	0.598	0.494	0.105	<0.0001	0.594	0.489	0.105	<0.0001
	Unexcused rate vs Excused rate	0.494	0.597	-0.104	<0.0001	0.489	0.598	-0.109	<0.0001
1st Grade Absenteeism	Days excused vs Excused rate	0.583	0.585	-0.002	0.095	0.535	0.537	-0.002	0.109
	Days total vs Total rate	0.623	0.626	-0.003	<0.05	0.590	0.593	-0.003	<0.05
	Days unexcused vs Unexcused rate	0.514	0.515	-0.001	<0.001	0.546	0.547	-0.001	<0.001
	Total rate vs Excused rate	0.626	0.585	0.041	<0.0001	0.593	0.537	0.056	<0.0001
	Total rate vs Unexcused rate	0.626	0.515	0.111	<0.0001	0.593	0.547	0.045	<0.01
	Unexcused rate vs Excused rate	0.515	0.585	-0.070	<0.001	0.547	0.537	0.011	0.603
2nd Grade Absenteeism	Days excused vs Excused rate	0.557	0.559	-0.002	0.052	0.518	0.520	-0.002	0.082
	Days total vs Total rate	0.639	0.642	-0.003	<0.05	0.593	0.593	-0.001	0.374
	Days unexcused vs Unexcused rate	0.596	0.598	-0.002	<0.01	0.586	0.587	-0.001	0.149
	Total rate vs Excused rate	0.642	0.559	0.083	<0.0001	0.593	0.520	0.074	<0.0001
	Total rate vs Unexcused rate	0.642	0.598	0.044	<0.0001	0.593	0.587	0.007	0.522

	Unexcused rate vs Excused rate	0.598	0.559	0.039	<0.05	0.587	0.520	0.067	<0.001
3rd Grade Absenteeism	Days excused vs Excused rate	0.558	0.559	-0.001	0.448	0.520	0.521	-0.001	0.606
	Days total vs Total rate	0.654	0.658	-0.004	<0.01	0.612	0.614	-0.002	0.246
	Days unexcused vs Unexcused rate	0.637	0.639	-0.002	<0.01	0.618	0.619	-0.001	0.127
	Total rate vs Excused rate	0.658	0.559	0.098	<0.0001	0.614	0.521	0.093	<0.0001
	Total rate vs Unexcused rate	0.658	0.639	0.018	<0.05	0.614	0.619	-0.005	0.573
	Unexcused rate vs Excused rate	0.639	0.559	0.080	<0.0001	0.619	0.521	0.098	<0.0001

Note: This table reports results from pairwise comparisons of the area under the ROC curve (AUC) for different operationalizations of absenteeism, using the DeLong test for correlated ROC curves. Each row compares two absence measures within a given grade level cohort (e.g., days absent vs. absence rate). The first two columns for each outcome show the AUCs for the two measures being compared, followed by the difference in AUC and the corresponding p-value. A positive AUC difference indicates that the first absence measure yielded a higher predictive accuracy than the second; negative differences indicate the reverse. Statistically significant p-values suggest that one measure provides a significantly stronger signal for identifying students at risk on this grade's math or ELA performance.

Appendix S5. DeLong Test Comparisons of AUCs for Various Absenteeism Measures in Predicting 4th Grade Math and ELA Outcomes

		Math				ELA			
Grade	**Absence Measures Compared**	**AUC for 1st Absence Measure**	**AUC for 2nd Absence Measure**	**AUC Difference**	**p-value**	**AUC for 1st Absence Measure**	**AUC for 2nd Absence Measure**	**AUC Difference**	**p-value**
Pre-K Absenteeism	Days excused vs Excused rate	0.581	0.592	-0.011	<0.01	0.550	0.554	-0.003	0.217
	Days total vs Total rate	0.579	0.590	-0.011	<0.01	0.549	0.552	-0.003	0.244
	Days unexcused vs Unexcused rate	0.495	0.495	0.000	0.089	0.496	0.496	0.000	0.162
	Total rate vs Excused rate	0.590	0.592	-0.002	0.082	0.552	0.554	-0.001	0.364
	Total rate vs Unexcused rate	0.590	0.495	0.095	<0.0001	0.552	0.496	0.056	<0.0001
	Unexcused rate vs Excused rate	0.495	0.592	-0.097	<0.0001	0.496	0.554	-0.057	<0.0001
Kindergarten Absenteeism	Days excused vs Excused rate	0.609	0.615	-0.006	<0.01	0.579	0.585	-0.006	<0.001
	Days total vs Total rate	0.610	0.616	-0.006	<0.01	0.578	0.584	-0.006	<0.001
	Days unexcused vs Unexcused rate	0.493	0.493	0.000	0.257	0.492	0.492	0.000	0.188
	Total rate vs Excused rate	0.616	0.615	0.001	0.775	0.584	0.585	-0.001	0.902
	Total rate vs Unexcused rate	0.616	0.493	0.123	<0.0001	0.584	0.492	0.092	<0.0001
	Unexcused rate vs Excused rate	0.493	0.615	-0.122	<0.0001	0.492	0.585	-0.093	<0.0001
1st Grade Absenteeism	Days excused vs Excused rate	0.596	0.599	-0.003	<0.05	0.558	0.561	-0.003	<0.01
	Days total vs Total rate	0.640	0.644	-0.004	<0.05	0.595	0.598	-0.004	<0.01
	Days unexcused vs Unexcused rate	0.516	0.517	-0.001	<0.01	0.513	0.513	-0.001	<0.01
	Total rate vs Excused rate	0.644	0.599	0.045	<0.0001	0.598	0.561	0.037	<0.0001
	Total rate vs Unexcused rate	0.644	0.517	0.127	<0.0001	0.598	0.513	0.085	<0.0001
	Unexcused rate vs Excused rate	0.517	0.599	-0.081	<0.0001	0.513	0.561	-0.048	<0.001
2nd Grade Absenteeism	Days excused vs Excused rate	0.550	0.552	-0.002	<0.05	0.535	0.537	-0.002	<0.05
	Days total vs Total rate	0.635	0.640	-0.005	<0.01	0.593	0.597	-0.004	<0.01
	Days unexcused vs Unexcused rate	0.574	0.576	-0.002	<0.001	0.548	0.549	-0.001	<0.01
	Total rate vs Excused rate	0.640	0.552	0.088	<0.0001	0.597	0.537	0.060	<0.0001
	Total rate vs Unexcused rate	0.640	0.576	0.064	<0.0001	0.597	0.549	0.048	<0.0001

	Unexcused rate vs Excused rate	0.576	0.552	0.024	0.169	0.549	0.537	0.012	0.431
3rd Grade Absenteeism	Days excused vs Excused rate	0.555	0.556	-0.001	0.324	0.521	0.521	-0.001	0.488
	Days total vs Total rate	0.645	0.648	-0.003	<0.01	0.609	0.613	-0.003	<0.01
	Days unexcused vs Unexcused rate	0.626	0.628	-0.002	<0.01	0.621	0.623	-0.002	<0.01
	Total rate vs Excused rate	0.648	0.556	0.092	<0.0001	0.613	0.521	0.091	<0.0001
	Total rate vs Unexcused rate	0.648	0.628	0.019	<0.05	0.613	0.623	-0.011	0.084
	Unexcused rate vs Excused rate	0.628	0.556	0.073	<0.0001	0.623	0.521	0.102	<0.0001
4th Grade Absenteeism	Days excused vs Excused rate	0.548	0.550	-0.002	<0.05	0.521	0.520	0.000	0.769
	Days total vs Total rate	0.667	0.672	-0.005	<0.0001	0.626	0.629	-0.003	<0.01
	Days unexcused vs Unexcused rate	0.655	0.657	-0.002	<0.01	0.639	0.642	-0.002	<0.01
	Total rate vs Excused rate	0.672	0.550	0.122	<0.0001	0.629	0.520	0.109	<0.0001
	Total rate vs Unexcused rate	0.672	0.657	0.015	0.051	0.629	0.642	-0.013	<0.05
	Unexcused rate vs Excused rate	0.657	0.550	0.107	<0.0001	0.642	0.520	0.121	<0.0001

Note: This table reports results from pairwise comparisons of the area under the ROC curve (AUC) for different operationalizations of absenteeism, using the DeLong test for correlated ROC curves. Each row compares two absence measures within a given grade level cohort (e.g., days absent vs. absence rate). The first two columns for each outcome show the AUCs for the two measures being compared, followed by the difference in AUC and the corresponding p-value. A positive AUC difference indicates that the first absence measure yielded a higher predictive accuracy than the second; negative differences indicate the reverse. Statistically significant p-values suggest that one measure provides a significantly stronger signal for identifying students at risk on this grade's math or ELA performance.

Appendix S6. DeLong Test Comparisons of AUCs for Various Absenteeism Measures in Predicting 5th Grade Math and ELA Outcomes

		Math				ELA			
Grade	**Absence Measures Compared**	**AUC for 1st Absence Measure**	**AUC for 2nd Absence Measure**	**AUC Difference**	**p-value**	**AUC for 1st Absence Measure**	**AUC for 2nd Absence Measure**	**AUC Difference**	**p-value**
Pre-K Absenteeism	Days excused vs Excused rate	0.564	0.574	-0.010	<0.01	0.559	0.564	-0.004	0.197
	Days total vs Total rate	0.566	0.576	-0.010	<0.01	0.560	0.564	-0.004	0.21
	Days unexcused vs Unexcused rate	0.499	0.499	0.000	0.085	0.497	0.497	0.000	0.209
	Total rate vs Excused rate	0.576	0.574	0.002	0.311	0.564	0.564	0.000	0.936
	Total rate vs Unexcused rate	0.576	0.499	0.078	<0.0001	0.564	0.497	0.066	<0.0001
	Unexcused rate vs Excused rate	0.499	0.574	-0.076	<0.0001	0.497	0.564	-0.066	<0.0001
Kindergarten Absenteeism	Days excused vs Excused rate	0.597	0.602	-0.005	<0.05	0.590	0.595	-0.005	<0.05
	Days total vs Total rate	0.600	0.605	-0.005	<0.05	0.592	0.597	-0.005	<0.05
	Days unexcused vs Unexcused rate	0.495	0.495	0.000	0.13	0.496	0.496	0.000	0.579
	Total rate vs Excused rate	0.605	0.602	0.003	0.529	0.597	0.595	0.002	0.653
	Total rate vs Unexcused rate	0.605	0.495	0.109	<0.0001	0.597	0.496	0.101	<0.0001
	Unexcused rate vs Excused rate	0.495	0.602	-0.106	<0.0001	0.496	0.595	-0.099	<0.0001
1st Grade Absenteeism	Days excused vs Excused rate	0.592	0.596	-0.004	<0.01	0.566	0.568	-0.002	0.092
	Days total vs Total rate	0.613	0.619	-0.006	<0.001	0.593	0.596	-0.003	0.06
	Days unexcused vs Unexcused rate	0.496	0.497	-0.001	<0.01	0.504	0.505	-0.001	<0.05
	Total rate vs Excused rate	0.619	0.596	0.023	<0.01	0.596	0.568	0.028	<0.01
	Total rate vs Unexcused rate	0.619	0.497	0.122	<0.0001	0.596	0.505	0.091	<0.0001
	Unexcused rate vs Excused rate	0.497	0.596	-0.099	<0.0001	0.505	0.568	-0.064	<0.001
2nd Grade Absenteeism	Days excused vs Excused rate	0.562	0.564	-0.002	0.088	0.536	0.538	-0.002	0.073
	Days total vs Total rate	0.601	0.604	-0.003	<0.05	0.581	0.584	-0.003	0.072
	Days unexcused vs Unexcused rate	0.526	0.527	-0.002	<0.01	0.539	0.541	-0.002	0.056
	Total rate vs Excused rate	0.604	0.564	0.040	<0.0001	0.584	0.538	0.046	<0.0001
	Total rate vs Unexcused rate	0.604	0.527	0.077	<0.0001	0.584	0.541	0.043	<0.001

	Unexcused rate vs Excused rate	0.527	0.564	-0.036	<0.05	0.541	0.538	0.003	0.884
3rd Grade Absenteeism	Days excused vs Excused rate	0.553	0.554	-0.001	0.353	0.532	0.532	0.000	0.81
	Days total vs Total rate	0.619	0.622	-0.003	<0.05	0.601	0.604	-0.003	<0.05
	Days unexcused vs Unexcused rate	0.603	0.605	-0.002	<0.01	0.599	0.601	-0.003	<0.05
	Total rate vs Excused rate	0.622	0.554	0.068	<0.0001	0.604	0.532	0.071	<0.0001
	Total rate vs Unexcused rate	0.622	0.605	0.017	<0.05	0.604	0.601	0.002	0.777
	Unexcused rate vs Excused rate	0.605	0.554	0.051	<0.001	0.601	0.532	0.069	<0.0001
4th Grade Absenteeism	Days excused vs Excused rate	0.550	0.551	0.000	0.72	0.527	0.528	0.000	0.746
	Days total vs Total rate	0.637	0.641	-0.004	<0.01	0.611	0.615	-0.005	<0.05
	Days unexcused vs Unexcused rate	0.623	0.625	-0.002	<0.05	0.621	0.624	-0.003	<0.05
	Total rate vs Excused rate	0.641	0.551	0.091	<0.0001	0.615	0.528	0.088	<0.0001
	Total rate vs Unexcused rate	0.641	0.625	0.016	<0.05	0.615	0.624	-0.009	0.243
	Unexcused rate vs Excused rate	0.625	0.551	0.075	<0.0001	0.624	0.528	0.096	<0.0001
5th Grade Absenteeism	Days excused vs Excused rate	0.551	0.552	-0.001	0.11	0.501	0.503	-0.002	<0.05
	Days total vs Total rate	0.637	0.644	-0.007	<0.0001	0.614	0.619	-0.005	<0.01
	Days unexcused vs Unexcused rate	0.613	0.617	-0.004	<0.0001	0.627	0.630	-0.003	<0.01
	Total rate vs Excused rate	0.644	0.552	0.092	<0.0001	0.619	0.503	0.116	<0.0001
	Total rate vs Unexcused rate	0.644	0.617	0.027	<0.001	0.619	0.630	-0.011	0.139
	Unexcused rate vs Excused rate	0.617	0.552	0.065	<0.0001	0.630	0.503	0.126	<0.0001

Note: This table reports results from pairwise comparisons of the area under the ROC curve (AUC) for different operationalizations of absenteeism, using the DeLong test for correlated ROC curves. Each row compares two absence measures within a given grade level cohort (e.g., days absent vs. absence rate). The first two columns for each outcome show the AUCs for the two measures being compared, followed by the difference in AUC and the corresponding p-value. A positive AUC difference indicates that the first absence measure yielded a higher predictive accuracy than the second; negative differences indicate the reverse. Statistically significant p-values suggest that one measure provides a significantly stronger signal for identifying students at risk on this grade's math or ELA performance.

Appendix S7. DeLong Test Comparisons of AUCs for Various Absenteeism Measures in Predicting 6th Grade Math and ELA Outcomes

		Math				ELA			
Grade	Absence Measures Compared	AUC for 1st Absence Measure	AUC for 2nd Absence Measure	AUC Difference	p-value	AUC for 1st Absence Measure	AUC for 2nd Absence Measure	AUC Difference	p-value
Pre-K Absenteeism	Days excused vs Excused rate	0.562	0.571	-0.010	<0.01	0.526	0.534	-0.008	<0.05
	Days total vs Total rate	0.563	0.573	-0.010	<0.01	0.527	0.535	-0.008	<0.05
	Days unexcused vs Unexcused rate	0.500	0.500	0.000	0.21	0.501	0.501	0.000	0.161
	Total rate vs Excused rate	0.573	0.571	0.001	0.384	0.535	0.534	0.001	0.607
	Total rate vs Unexcused rate	0.573	0.500	0.073	<0.0001	0.535	0.501	0.034	<0.05
	Unexcused rate vs Excused rate	0.500	0.571	-0.071	<0.0001	0.501	0.534	-0.033	<0.05
Kindergarten Absenteeism	Days excused vs Excused rate	0.594	0.598	-0.004	<0.05	0.572	0.576	-0.004	0.052
	Days total vs Total rate	0.595	0.599	-0.004	<0.05	0.574	0.579	-0.004	<0.05
	Days unexcused vs Unexcused rate	0.494	0.494	0.000	0.174	0.498	0.498	0.000	0.495
	Total rate vs Excused rate	0.599	0.598	0.001	0.856	0.579	0.576	0.003	0.585
	Total rate vs Unexcused rate	0.599	0.494	0.105	<0.0001	0.579	0.498	0.081	<0.0001
	Unexcused rate vs Excused rate	0.494	0.598	-0.104	<0.0001	0.498	0.576	-0.078	<0.0001
1st Grade Absenteeism	Days excused vs Excused rate	0.609	0.612	-0.003	<0.05	0.550	0.552	-0.002	0.119
	Days total vs Total rate	0.618	0.622	-0.004	<0.05	0.594	0.597	-0.003	<0.05
	Days unexcused vs Unexcused rate	0.481	0.482	-0.001	<0.05	0.525	0.525	-0.001	<0.05
	Total rate vs Excused rate	0.622	0.612	0.010	0.174	0.597	0.552	0.045	<0.0001
	Total rate vs Unexcused rate	0.622	0.482	0.139	<0.0001	0.597	0.525	0.072	<0.0001
	Unexcused rate vs Excused rate	0.482	0.612	-0.130	<0.0001	0.525	0.552	-0.027	0.13
2nd Grade Absenteeism	Days excused vs Excused rate	0.569	0.570	-0.001	0.16	0.524	0.524	-0.001	0.569
	Days total vs Total rate	0.610	0.612	-0.002	0.09	0.593	0.594	-0.001	0.343
	Days unexcused vs Unexcused rate	0.527	0.528	-0.001	<0.05	0.565	0.566	-0.001	0.278
	Total rate vs Excused rate	0.612	0.570	0.041	<0.0001	0.594	0.524	0.070	<0.0001
	Total rate vs Unexcused rate	0.612	0.528	0.084	<0.0001	0.594	0.566	0.028	<0.05

	Unexcused rate vs Excused rate	0.528	0.570	-0.042	<0.01	0.566	0.524	0.042	<0.05
3rd Grade Absenteeism	Days excused vs Excused rate	0.555	0.557	-0.001	0.137	0.531	0.532	-0.001	0.345
	Days total vs Total rate	0.613	0.615	-0.002	0.074	0.594	0.596	-0.002	0.146
	Days unexcused vs Unexcused rate	0.586	0.587	-0.001	<0.05	0.594	0.595	-0.001	0.12
	Total rate vs Excused rate	0.615	0.557	0.058	<0.0001	0.596	0.532	0.064	<0.0001
	Total rate vs Unexcused rate	0.615	0.587	0.028	<0.001	0.596	0.595	0.001	0.949
	Unexcused rate vs Excused rate	0.587	0.557	0.030	<0.05	0.595	0.532	0.063	<0.0001
4th Grade Absenteeism	Days excused vs Excused rate	0.552	0.552	0.000	0.878	0.530	0.530	0.000	0.906
	Days total vs Total rate	0.632	0.635	-0.002	0.06	0.610	0.615	-0.005	<0.05
	Days unexcused vs Unexcused rate	0.610	0.611	-0.001	0.134	0.606	0.609	-0.003	<0.05
	Total rate vs Excused rate	0.635	0.552	0.083	<0.0001	0.615	0.530	0.085	<0.0001
	Total rate vs Unexcused rate	0.635	0.611	0.024	<0.01	0.615	0.609	0.006	0.464
	Unexcused rate vs Excused rate	0.611	0.552	0.059	<0.0001	0.609	0.530	0.079	<0.0001
5th Grade Absenteeism	Days excused vs Excused rate	0.554	0.557	-0.003	<0.01	0.524	0.527	-0.003	<0.05
	Days total vs Total rate	0.636	0.644	-0.008	<0.0001	0.609	0.616	-0.007	<0.001
	Days unexcused vs Unexcused rate	0.612	0.616	-0.005	<0.0001	0.602	0.606	-0.004	<0.0001
	Total rate vs Excused rate	0.644	0.557	0.087	<0.0001	0.616	0.527	0.089	<0.0001
	Total rate vs Unexcused rate	0.644	0.616	0.028	<0.0001	0.616	0.606	0.010	0.216
	Unexcused rate vs Excused rate	0.616	0.557	0.060	<0.0001	0.606	0.527	0.079	<0.0001
6th Grade Absenteeism	Days excused vs Excused rate	0.544	0.547	-0.002	<0.01	0.531	0.535	-0.004	<0.001
	Days total vs Total rate	0.628	0.634	-0.006	<0.0001	0.607	0.612	-0.005	<0.001
	Days unexcused vs Unexcused rate	0.615	0.619	-0.004	<0.0001	0.604	0.607	-0.004	<0.01
	Total rate vs Excused rate	0.634	0.547	0.087	<0.0001	0.612	0.535	0.077	<0.0001
	Total rate vs Unexcused rate	0.634	0.619	0.015	<0.05	0.612	0.607	0.005	0.504
	Unexcused rate vs Excused rate	0.619	0.547	0.072	<0.0001	0.607	0.535	0.072	<0.0001

Note: This table reports results from pairwise comparisons of the area under the ROC curve (AUC) for different operationalizations of absenteeism, using the DeLong test for correlated ROC curves. Each row compares two absence measures within a given grade level cohort (e.g., days absent vs. absence rate). The first two columns for each outcome show the AUCs for the two measures being compared, followed by the difference in AUC and the corresponding p-value. A positive AUC difference indicates that the first absence measure yielded a higher predictive accuracy than the second; negative differences indicate the reverse. Statistically significant p-values suggest that one measure provides a significantly stronger signal for identifying students at risk on this grade's math or ELA performance.

Appendix S8. DeLong Test Comparisons of AUCs for Various Absenteeism Measures in Predicting 7th Grade Math and ELA Outcomes

		Math				ELA			
Grade	**Absence Measures Compared**	**AUC for 1st Absence Measure**	**AUC for 2nd Absence Measure**	**AUC Difference**	**p-value**	**AUC for 1st Absence Measure**	**AUC for 2nd Absence Measure**	**AUC Difference**	**p-value**
Pre-K Absenteeism	Days excused vs Excused rate	0.558	0.567	-0.009	<0.01	0.561	0.565	-0.004	0.22
	Days total vs Total rate	0.560	0.569	-0.009	<0.01	0.563	0.567	-0.004	0.224
	Days unexcused vs Unexcused rate	0.500	0.500	0.000	<0.05	0.501	0.501	0.000	0.152
	Total rate vs Excused rate	0.569	0.567	0.001	0.414	0.567	0.565	0.003	0.218
	Total rate vs Unexcused rate	0.569	0.500	0.068	<0.0001	0.567	0.501	0.066	<0.0001
	Unexcused rate vs Excused rate	0.500	0.567	-0.067	<0.0001	0.501	0.565	-0.063	<0.0001
Kindergarten Absenteeism	Days excused vs Excused rate	0.585	0.589	-0.004	<0.05	0.571	0.578	-0.007	<0.01
	Days total vs Total rate	0.593	0.597	-0.004	<0.05	0.580	0.587	-0.007	<0.01
	Days unexcused vs Unexcused rate	0.500	0.500	0.000	0.095	0.499	0.499	0.000	0.249
	Total rate vs Excused rate	0.597	0.589	0.008	0.091	0.587	0.578	0.009	0.118
	Total rate vs Unexcused rate	0.597	0.500	0.097	<0.0001	0.587	0.499	0.088	<0.0001
	Unexcused rate vs Excused rate	0.500	0.589	-0.089	<0.0001	0.499	0.578	-0.079	<0.0001
1st Grade Absenteeism	Days excused vs Excused rate	0.582	0.584	-0.002	0.051	0.544	0.546	-0.002	0.078
	Days total vs Total rate	0.606	0.609	-0.003	<0.05	0.598	0.601	-0.003	<0.05
	Days unexcused vs Unexcused rate	0.501	0.501	0.000	0.21	0.539	0.540	-0.001	<0.01
	Total rate vs Excused rate	0.609	0.584	0.025	<0.001	0.601	0.546	0.055	<0.0001
	Total rate vs Unexcused rate	0.609	0.501	0.108	<0.0001	0.601	0.540	0.061	<0.0001
	Unexcused rate vs Excused rate	0.501	0.584	-0.083	<0.0001	0.540	0.546	-0.006	0.728
2nd Grade Absenteeism	Days excused vs Excused rate	0.574	0.576	-0.002	<0.05	0.506	0.507	-0.001	0.259
	Days total vs Total rate	0.608	0.609	-0.001	0.255	0.600	0.602	-0.002	0.115
	Days unexcused vs Unexcused rate	0.514	0.515	0.000	0.36	0.598	0.598	0.000	0.435
	Total rate vs Excused rate	0.609	0.576	0.033	<0.0001	0.602	0.507	0.095	<0.0001
	Total rate vs Unexcused rate	0.609	0.515	0.095	<0.0001	0.602	0.598	0.004	0.741

	Unexcused rate vs Excused rate	0.515	0.576	-0.061	<0.0001	0.598	0.507	0.091	<0.0001
3rd Grade Absenteeism	Days excused vs Excused rate	0.543	0.543	0.000	0.942	0.536	0.537	-0.001	0.308
	Days total vs Total rate	0.622	0.622	-0.001	0.553	0.626	0.628	-0.002	0.212
	Days unexcused vs Unexcused rate	0.602	0.603	-0.001	0.124	0.618	0.620	-0.002	0.084
	Total rate vs Excused rate	0.622	0.543	0.080	<0.0001	0.628	0.537	0.092	<0.0001
	Total rate vs Unexcused rate	0.622	0.603	0.019	<0.01	0.628	0.620	0.008	0.283
	Unexcused rate vs Excused rate	0.603	0.543	0.061	<0.0001	0.620	0.537	0.083	<0.0001
4th Grade Absenteeism	Days excused vs Excused rate	0.543	0.543	0.001	0.503	0.536	0.536	0.000	0.924
	Days total vs Total rate	0.624	0.627	-0.002	0.057	0.629	0.632	-0.003	<0.05
	Days unexcused vs Unexcused rate	0.611	0.613	-0.002	<0.05	0.631	0.634	-0.002	<0.05
	Total rate vs Excused rate	0.627	0.543	0.084	<0.0001	0.632	0.536	0.096	<0.0001
	Total rate vs Unexcused rate	0.627	0.613	0.014	<0.05	0.632	0.634	-0.002	0.79
	Unexcused rate vs Excused rate	0.613	0.543	0.070	<0.0001	0.634	0.536	0.098	<0.0001
5th Grade Absenteeism	Days excused vs Excused rate	0.547	0.548	-0.001	0.522	0.532	0.534	-0.001	0.227
	Days total vs Total rate	0.625	0.631	-0.005	<0.001	0.625	0.631	-0.006	<0.01
	Days unexcused vs Unexcused rate	0.604	0.607	-0.003	<0.001	0.624	0.628	-0.004	<0.01
	Total rate vs Excused rate	0.631	0.548	0.083	<0.0001	0.631	0.534	0.097	<0.0001
	Total rate vs Unexcused rate	0.631	0.607	0.023	<0.001	0.631	0.628	0.003	0.635
	Unexcused rate vs Excused rate	0.607	0.548	0.059	<0.0001	0.628	0.534	0.094	<0.0001
6th Grade Absenteeism	Days excused vs Excused rate	0.546	0.548	-0.002	<0.05	0.534	0.536	-0.003	<0.01
	Days total vs Total rate	0.631	0.634	-0.004	<0.01	0.622	0.627	-0.005	<0.001
	Days unexcused vs Unexcused rate	0.611	0.613	-0.002	<0.01	0.619	0.623	-0.004	<0.001
	Total rate vs Excused rate	0.634	0.548	0.086	<0.0001	0.627	0.536	0.091	<0.0001
	Total rate vs Unexcused rate	0.634	0.613	0.021	<0.01	0.627	0.623	0.004	0.512
	Unexcused rate vs Excused rate	0.613	0.548	0.066	<0.0001	0.623	0.536	0.086	<0.0001
7th Grade Absenteeism	Days excused vs Excused rate	0.541	0.544	-0.003	<0.01	0.543	0.544	-0.002	<0.05
	Days total vs Total rate	0.637	0.643	-0.006	<0.0001	0.645	0.649	-0.003	<0.01

Days unexcused vs Unexcused rate	0.628	0.632	-0.004	<0.0001	0.640	0.643	-0.002	<0.01
Total rate vs Excused rate	0.643	0.544	0.099	<0.0001	0.649	0.544	0.104	<0.0001
Total rate vs Unexcused rate	0.643	0.632	0.012	0.072	0.649	0.643	0.006	0.42
Unexcused rate vs Excused rate	0.632	0.544	0.088	<0.0001	0.643	0.544	0.099	<0.0001

Note: This table reports results from pairwise comparisons of the area under the ROC curve (AUC) for different operationalizations of absenteeism, using the DeLong test for correlated ROC curves. Each row compares two absence measures within a given grade level cohort (e.g., days absent vs. absence rate). The first two columns for each outcome show the AUCs for the two measures being compared, followed by the difference in AUC and the corresponding p-value. A positive AUC difference indicates that the first absence measure yielded a higher predictive accuracy than the second; negative differences indicate the reverse. Statistically significant p-values suggest that one measure provides a significantly stronger signal for identifying students at risk on this grade's math or ELA performance.

Appendix S9. Robustness Checks using a Common Sample

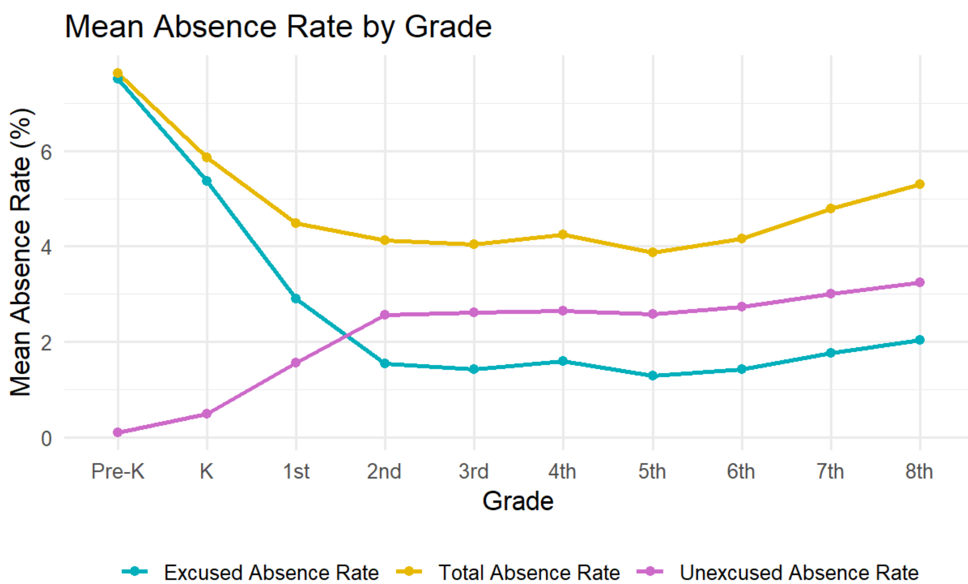
We explored whether findings were driven by different students joining or leaving the sample across the Pre-K through 8th grade years. To do so, for RQ1-RQ4, we refit our models using a common sample of students with no missing data in attendance or standardized test scores in all grades. We found our broad conclusions for all RQs were robust to a broader versus common sample approach. All findings for the common sample approach can be found here.

The analytic sample for this approach only includes students with valid attendance value from Pre-K to 8th grade (thus excluding students who did not attend Pre-K) and both a non-missing ELA and a non-missing math standardized test score for every year from 3rd-8th grade. In total, the sample size is N = 2,810 for all tables and figures below.

RQ1: Absenteeism Descriptive Patterns
To examine how absenteeism patterns descriptively differed by measurement choice, we generated descriptive statistics for excused, unexcused, and total absences, both in terms of absolute days missed and rates of absence (days absent divided by days enrolled). We visualized longitudinal absence trajectories using line graphs.

As seen in Figure S9.1 below, the patterns for mean absence rate by grade closely mirror those of Figure 3. The only difference is that unexcused absence rate exceeds excused absence rate starting in 2nd grade instead of 3rd grade. Importantly, the overall shape, magnitude, and relative ordering of excused and unexcused absence trajectories remain nearly identical across samples. Even with this restricted analytic sample, the developmental pattern of absenteeism, with excused absences driving absenteeism in the earlier years and unexcused absences exceeding them later, remains similar. We believe this supports the robustness of our main RQ1 findings.

Figure S9.1 Mean Absence Rate by Grade



RQ2: ROC Curve Analysis. We used ROC curve analysis to evaluate the diagnostic accuracy of absenteeism measures in predicting whether a student would *Not Meet Expectations* on standardized math and ELA assessments in 8^{th} grade. For each absence measure (days excused, days unexcused, days total, excused absence rate, unexcused absence rate, total absence rate), we estimated the ROC curve and computed the AUC. To test whether observed AUC differences across measures were statistically significant, we applied the DeLong test.

To test robustness using the common sample, we replicated Table 2 for 8^{th} grade below. We find very similar findings to Table 2 in the main paper. We see that unexcused absences show the smallest AUC in the earlier grades, with the AUC for total and excused absences marginally higher. The AUC, particularly for total and unexcused absences, increases as students approach middle school, signaling a strengthening of predictive accuracy.

Like in the main paper these results highlight the following insights: First, the predictive value of absences is weaker and largely driven by excused absences in early childhood, but by the upper elementary grades, unexcused absences become more salient. Second, predictive strength steadily increases across grades, peaking in middle school, where absence measures achieve their highest AUCs. However, we do not see that absence rates are superior to number-of-days measures. This is likely because this sample is a select group of students who remained in the Massachusetts Department of Secondary and Elementary Education dataset for ten consecutive years. Because of this, these students likely exhibit less mobility and more stable enrollment histories, so the variability in total days absent is artificially constrained, attenuating differences between rates and counts. In real-world settings where student mobility and enrollment length may vary more, absence rate remains more generalizable.

Table S9.1. DeLong Test Comparisons of AUCs for Various Absenteeism Measures in Predicting 8th Grade Math and ELA Outcomes

		8th Grade Math				8th Grade ELA			
Grade	**Absence Measures Compared**	**AUC for 1st Absence Measure**	**AUC for 2nd Absence Measure**	**AUC Difference**	**p-value**	**AUC for 1st Absence Measure**	**AUC for 2nd Absence Measure**	**AUC Difference**	**p-value**
Pre-K Absenteeism	Days excused vs Excused rate	0.579	0.584	-0.005	0.35	0.535	0.541	-0.006	0.227
	Days total vs Total rate	0.580	0.585	-0.004	0.399	0.540	0.546	-0.006	0.226
	Days unexcused vs Unexcused rate	0.502	0.502	0.000	0.096	0.500	0.500	0.000	0.478
	Total rate vs Excused rate	0.585	0.584	0.001	0.717	0.546	0.541	0.005	0.303
	Total rate vs Unexcused rate	0.585	0.502	0.082	<0.0001	0.546	0.500	0.046	<0.05
	Unexcused rate vs Excused rate	0.502	0.584	-0.081	<0.0001	0.500	0.541	-0.041	<0.05
Kindergarten Absenteeism	Days excused vs Excused rate	0.582	0.586	-0.005	0.085	0.557	0.558	-0.001	0.503
	Days total vs Total rate	0.583	0.588	-0.005	0.061	0.547	0.550	-0.002	0.302
	Days unexcused vs Unexcused rate	0.493	0.493	0.000	0.376	0.489	0.489	0.000	0.281

	Total rate vs Excused rate	0.588	0.586	0.002	0.866	0.550	0.558	-0.009	0.251
	Total rate vs Unexcused rate	0.588	0.493	0.094	<0.0001	0.550	0.489	0.061	<0.001
	Unexcused rate vs Excused rate	0.493	0.586	-0.093	<0.0001	0.489	0.558	-0.069	<0.001
1st Grade Absenteeism	Days excused vs Excused rate	0.568	0.568	0.000	0.963	0.552	0.554	-0.002	0.185
	Days total vs Total rate	0.615	0.617	-0.002	0.16	0.584	0.588	-0.004	<0.05
	Days unexcused vs Unexcused rate	0.537	0.538	0.000	0.287	0.516	0.517	-0.001	<0.05
	Total rate vs Excused rate	0.617	0.568	0.049	<0.01	0.588	0.554	0.034	<0.05
	Total rate vs Unexcused rate	0.617	0.538	0.079	<0.001	0.588	0.517	0.071	<0.001
	Unexcused rate vs Excused rate	0.538	0.568	-0.030	0.352	0.517	0.554	-0.037	0.21
2nd Grade Absenteeism	Days excused vs Excused rate	0.585	0.588	-0.003	0.192	0.540	0.545	-0.005	<0.05
	Days total vs Total rate	0.649	0.652	-0.003	0.217	0.580	0.584	-0.003	0.172
	Days unexcused vs Unexcused rate	0.608	0.610	-0.002	0.105	0.576	0.578	-0.002	0.058
	Total rate vs Excused rate	0.652	0.588	0.064	<0.001	0.584	0.545	0.039	<0.05
	Total rate vs Unexcused rate	0.652	0.610	0.042	<0.01	0.584	0.578	0.006	0.666
	Unexcused rate vs Excused rate	0.610	0.588	0.021	0.445	0.578	0.545	0.033	0.177
3rd Grade Absenteeism	Days excused vs Excused rate	0.539	0.539	0.000	0.835	0.530	0.529	0.001	0.709
	Days total vs Total rate	0.642	0.642	0.000	0.921	0.572	0.572	0.000	0.902
	Days unexcused vs Unexcused rate	0.635	0.635	0.000	0.928	0.565	0.565	0.000	0.778
	Total rate vs Excused rate	0.642	0.539	0.104	<0.0001	0.572	0.529	0.043	<0.01
	Total rate vs Unexcused rate	0.642	0.635	0.007	0.595	0.572	0.565	0.007	0.602
	Unexcused rate vs Excused rate	0.635	0.539	0.096	<0.001	0.565	0.529	0.036	0.15
4th Grade Absenteeism	Days excused vs Excused rate	0.554	0.553	0.001	0.642	0.541	0.541	0.000	0.986
	Days total vs Total rate	0.666	0.665	0.001	0.273	0.618	0.617	0.001	0.349
	Days unexcused vs Unexcused rate	0.656	0.655	0.001	0.29	0.608	0.609	-0.001	0.508
	Total rate vs Excused rate	0.665	0.553	0.112	<0.0001	0.617	0.541	0.076	<0.0001
	Total rate vs Unexcused rate	0.665	0.655	0.010	0.444	0.617	0.609	0.008	0.527
	Unexcused rate vs Excused rate	0.655	0.553	0.102	<0.001	0.609	0.541	0.068	<0.01

5th Grade Absenteeism	Days excused vs Excused rate	0.531	0.533	-0.002	0.204	0.525	0.526	-0.001	0.295
	Days total vs Total rate	0.649	0.649	0.000	0.992	0.594	0.596	-0.002	0.206
	Days unexcused vs Unexcused rate	0.652	0.652	0.000	0.969	0.584	0.585	-0.001	0.164
	Total rate vs Excused rate	0.649	0.533	0.116	<0.0001	0.596	0.526	0.070	<0.0001
	Total rate vs Unexcused rate	0.649	0.652	-0.003	0.779	0.596	0.585	0.011	0.389
	Unexcused rate vs Excused rate	0.652	0.533	0.119	<0.0001	0.585	0.526	0.059	<0.05
6th Grade Absenteeism	Days excused vs Excused rate	0.569	0.570	-0.001	0.416	0.548	0.548	0.000	0.839
	Days total vs Total rate	0.653	0.655	-0.002	0.128	0.595	0.597	-0.002	0.187
	Days unexcused vs Unexcused rate	0.633	0.634	-0.001	0.323	0.585	0.586	-0.001	0.209
	Total rate vs Excused rate	0.655	0.570	0.085	<0.0001	0.597	0.548	0.049	<0.01
	Total rate vs Unexcused rate	0.655	0.634	0.021	0.114	0.597	0.586	0.011	0.356
	Unexcused rate vs Excused rate	0.634	0.570	0.064	<0.05	0.586	0.548	0.038	0.105
7th Grade Absenteeism	Days excused vs Excused rate	0.546	0.544	0.002	0.171	0.525	0.525	0.000	0.824
	Days total vs Total rate	0.651	0.652	-0.001	0.561	0.605	0.606	-0.001	0.411
	Days unexcused vs Unexcused rate	0.644	0.644	0.000	0.844	0.620	0.620	0.000	0.913
	Total rate vs Excused rate	0.652	0.544	0.108	<0.0001	0.606	0.525	0.081	<0.0001
	Total rate vs Unexcused rate	0.652	0.644	0.008	0.58	0.606	0.620	-0.014	0.224
	Unexcused rate vs Excused rate	0.644	0.544	0.100	<0.001	0.620	0.525	0.095	<0.0001
8th Grade Absenteeism	Days excused vs Excused rate	0.514	0.513	0.001	0.559	0.511	0.513	-0.002	0.255
	Days total vs Total rate	0.677	0.680	-0.004	0.174	0.617	0.621	-0.004	0.122
	Days unexcused vs Unexcused rate	0.691	0.694	-0.003	0.249	0.639	0.641	-0.002	0.353
	Total rate vs Excused rate	0.680	0.513	0.167	<0.0001	0.621	0.513	0.108	<0.0001
	Total rate vs Unexcused rate	0.680	0.694	-0.014	0.303	0.621	0.641	-0.019	0.1
	Unexcused rate vs Excused rate	0.694	0.513	0.181	<0.0001	0.641	0.513	0.128	<0.0001

Note: This table reports results from pairwise comparisons of the area under the ROC curve (AUC) for different operationalizations of absenteeism, using the DeLong test for correlated ROC curves. Each row compares two absence measures within a given grade level cohort (e.g., days absent vs. absence rate). The first two columns for each outcome show the AUCs for the two measures being compared, followed by the difference in AUC and the corresponding p-value. A positive AUC difference indicates that the first absence measure yielded a higher predictive accuracy than the second; negative differences indicate the reverse. Statistically significant p-values suggest that one measure provides a significantly stronger signal for identifying students at risk on 8th grade math or ELA performance.

RQ3: Identifying Empirical Thresholds with Youden's J
We calculated the optimal Youden's J index for each of the ROC curves in RQ2 to derive the cutoff that maximized balanced classification accuracy for each grade. To test the robustness of this, we replicated the Total Days Absent thresholds in Table 3 in the main text with our common sample, seen in Table S9.2 below. As with the main text results, all Youden's J recommended thresholds are less than the commonly used 18 day cutoff, with developmental variation.

Table S9.2. Youden's J and Optimal Total Absence Cutoffs for Predicting 8th Grade MCAS Math and ELA Performance

Grade	8th Grade MCAS Subject	Youden's J (Total Days Absent)	Youden's J Threshold (Total Days Absent)
Pre-K	Math	0.140	14 days
	ELA	0.103	14 days
Kindergarten	Math	0.139	8 days
	ELA	0.064	11 days
1st Grade	Math	0.171	8 days
	ELA	0.136	6 days
2nd Grade	Math	0.199	6 days
	ELA	0.114	11 days
3rd Grade	Math	0.217	6 days
	ELA	0.124	9 days
4th Grade	Math	0.252	7 days
	ELA	0.179	5 days
5th Grade	Math	0.214	5 days
	ELA	0.155	5 days
6th Grade	Math	0.255	9 days
	ELA	0.151	6 days
7th Grade	Math	0.252	9 days
	ELA	0.151	10 days
8th Grade	Math	0.267	10 days
	ELA	0.187	10 days

Note: This table reports Youden's J statistics and the corresponding optimal absenteeism cutoffs for predicting 8th grade math and ELA standardized testing achievement level of *Not Meeting Expectations*. Youden's J represents the maximum vertical distance between the ROC curve and the random chance line.

RQ4: Temporal Proximity

To assess how predictive validity varied over time, we extended the ROC curve and Youden's J analyses to standardized math and ELA assessments from grades 3-7. We used absence measures from the same academic year or earlier than the standardized assessment to ensure temporal alignment between predictor and outcome. We then compared cross-grade patterns in diagnostic accuracy and thresholds to the 8th grade results presented in RQ2 and RQ3.

To test the robustness of this, we replicated Table 4 in the main text with our common sample below in Table S9.3. As seen below, all Youden's J recommended cutoffs are less than 18 days or 10% of a 180-day school year.

Table S9.3. Top 3 Absence Measures and Optimal Thresholds for Predicting Math and ELA Standardized Test Performance in Grades 3-7

	Math				ELA			
Standardized Test Grade	Top Absence Predictors	AUCs	Youden's J	Youden's J Threshold (Number of Days)	Top Absence Predictors	AUCs	Youden's J	Youden's J Threshold (Number of Days)
Grade 3	Grade 2 Absence Rate (Total)	0.643	0.217	6	Grade 3 Days Absent (Total)	0.596	0.152	7
Grade 3	Grade 2 Days Absent (Total)	0.641	0.210	7	Grade 3 Absence Rate (Total)	0.595	0.153	7
Grade 3	Grade 3 Absence Rate (Total)	0.639	0.228	9	Grade 2 Absence Rate (Unexcused)	0.594	0.173	4
Grade 4	Grade 4 Absence Rate (Total)	0.700	0.297	10	Grade 4 Absence Rate (Unexcused)	0.638	0.221	5
Grade 4	Grade 4 Days Absent (Total)	0.698	0.292	7	Grade 4 Days Absent (Unexcused)	0.636	0.218	6
Grade 4	Grade 4 Absence Rate (Unexcused)	0.680	0.291	5	Grade 4 Absence Rate (Total)	0.630	0.192	6
Grade 5	Grade 4 Days Absent (Total)	0.664	0.241	7	Grade 4 Absence Rate (Total)	0.616	0.179	8
Grade 5	Grade 4 Absence Rate (Total)	0.664	0.240	7	Grade 4 Days Absent (Total)	0.616	0.176	8
Grade 5	Grade 5 Absence Rate (Total)	0.651	0.219	7	Grade 4 Absence Rate (Unexcused)	0.611	0.162	9
Grade 6	Grade 4 Absence Rate (Total)	0.658	0.238	9	Grade 4 Absence Rate (Total)	0.631	0.197	9
Grade 6	Grade 4 Days Absent (Total)	0.658	0.230	10	Grade 4 Days Absent (Total)	0.630	0.196	9
Grade 6	Grade 5 Absence Rate (Total)	0.651	0.218	4	Grade 5 Absence Rate (Unexcused)	0.626	0.185	7
Grade 7	Grade 7 Absence Rate (Unexcused)	0.660	0.279	5	Grade 7 Absence Rate (Unexcused)	0.641	0.227	4

Grade 7	Grade 7 Days Absent (Unexcused)	0.660	0.274	6	Grade 7 Days Absent (Unexcused)	0.641	0.224	4
Grade 7	Grade 7 Absence Rate (Total)	0.658	0.262	7	Grade 7 Absence Rate (Total)	0.639	0.206	11

Note: This table reports the top 3 absence predictors for each grade's standardized test outcome based on largest AUC. Youden's J statistics and the corresponding optimal absenteeism cutoffs are for predicting each grade's math and ELA standardized testing achievement level of *Not Meeting Expectations*. Youden's J represents the maximum vertical distance between the ROC curve and the random chance line. The optimal Youden's J Threshold is the number of days at which Youden's J is maximized. Youden's J Thresholds for absence rates have been converted to the equivalent number of days missed based on a 180-day school year.