The College Readiness Scorecard: Research Foundations for a Ten-Dimension Composite Model

A Comprehensive Research Compendium

General Academic, Inc.

2026

Table of Contents

Introduction: Why a Ten-Dimension Scorecard

College readiness has never been reducible to a single number. For decades, admissions offices at institutions across the country have grappled with the same fundamental question: how do you fairly and accurately judge whether a student is prepared to succeed in college? Their answer has been holistic review – a structured but contextual evaluation of academic achievement, course rigor, standardized testing, personal character, extracurricular engagement, and a cluster of noncognitive traits that bear directly on whether a student will persist through a degree.

The College Readiness Scorecard operationalizes this same philosophy. Rather than collapsing a student's preparation into a single GPA or test score, the scorecard evaluates students across ten dimensions – from Academic Achievement (D01) to Financial Preparedness (D10) – aggregating them into a single composite score on a 1-9 scale. Each dimension carries a weight reflecting its evidence-based importance to long-term college success. And sitting at the top of that hierarchy, carrying the highest individual weight at 20%, is D01: Academic Achievement.

The scorecard's ten-dimension structure is not an invention. It is a formalization of what selective colleges have practiced for generations. NACAC's national surveys of admissions officers consistently show that no single factor dominates admissions decisions. Colleges are not simply seeking the students with the most A's or the highest test scorers. They are evaluating a profile.

This compendium presents the full research justification for each of the scorecard's ten dimensions. Drawing on a body of research spanning national admissions surveys, longitudinal outcome studies, and institutional data from actual high school profiles, each chapter marshals the evidence for why that dimension belongs in the framework, why its weight is calibrated the way it is, how it is measured, and how it relates to the broader composite.

The following table summarizes the scorecard's architecture:

Dimension	Weight	What It Measures
D01: Academic Achievement	20%	GPA, class rank, grade trend
D02: Standardized Testing	15%	SAT/ACT scores (test-optional design)
D03: Course Rigor	10%	AP, IB, dual-credit, honors enrollment
D04: Personal Qualities	12%	Character, resilience, empathy
D05: Extracurricular Depth	10%	Duration, leadership, hours, achievements
D06: College Knowledge	8%	Campus visits, counselor meetings, college list, essay prep
D07: Self-Awareness & Mindset	8%	Growth mindset, emotional intelligence, tenacity
D08: Academic Skills	7%	Academic ownership, active learning habits

Dimension	Weight	What It Measures
D09: College Motivation	5%	Eagerness for college, school enthusiasm, balance
D10: Financial Preparedness	5%	FAFSA awareness, scholarship research, family cost discussions

Academic dimensions (D01 + D02 + D03) collectively carry 45% of the total composite weight – consistent with NACAC analysis showing academic factors account for roughly 50-70% of decision weight at selective colleges. The holistic remainder – personal qualities, activities, college knowledge, mindset, skills, motivation, and financial preparedness – accounts for 55%. Every dimension earns its weight through research.

Chapter 1: Academic Achievement (D01) – 20% Weight

Dimension D01 | Weight: 20% | Composition: GPA Unweighted (60%) + Class Rank (40%) + Grade Trend Modifier (+/-0.5)

Introduction: The Case for a Multidimensional Scorecard

College readiness has never been reducible to a single number. For decades, admissions offices at institutions across the country have grappled with the same fundamental question: how do you fairly and accurately judge whether a student is prepared to succeed in college? Their answer has been holistic review – a structured but contextual evaluation of academic achievement, course rigor, standardized testing, personal character, extracurricular engagement, and a cluster of noncognitive traits that bear directly on whether a student will persist through a degree.

The College Readiness Scorecard operationalizes this same philosophy. Rather than collapsing a student's preparation into a single GPA or test score, the scorecard evaluates students across ten dimensions – from Academic Achievement (D01) to Financial Preparedness (D10) – aggregating them into a single composite score on a 1-9 scale. Each dimension carries a weight reflecting its evidence-based importance to long-term college success. And sitting at the top of that hierarchy, carrying the highest individual weight at 20%, is D01: Academic Achievement.

This article explains why. Drawing on a body of research spanning national admissions surveys, longitudinal outcome studies, and institutional data from actual high school profiles, we make the case that GPA and class rank – the two pillars of D01 – are not merely convenient metrics. They are the most empirically validated predictors of college success available, and they deserve the prominence the scorecard assigns them.

Part I: Why Multidimensional Assessment Is the Right Framework

Colleges Have Always Used Multiple Factors – and That Is Not Changing

The scorecard's ten-dimension structure is not an invention. It is a formalization of what selective colleges have practiced for generations. NACAC's national surveys of admissions

officers consistently show that no single factor dominates admissions decisions. In 2018, 75% of colleges rated overall GPA as considerably important, 73% rated grades in college-prep courses as considerably important, and 62% rated curriculum strength as considerably important – but only 46% rated standardized test scores at the same level of importance, down from 59% in 2007. Colleges are not simply seeking the students with the most A's. They are evaluating a profile.

As NACAC itself has stated: "There is no definite plan or specific combination of factors that will guarantee a student admission to their preferred institution." The scorecard takes this reality seriously. Its weighted composite approach mirrors what admission offices actually do – assign varying importance to a range of factors and aggregate them into an overall judgment. By making those weights explicit and transparent, the scorecard reduces the subjectivity and potential bias that can creep into opaque holistic review processes.

The theoretical foundation for this approach is deep. NACAC's Fundamentals of College Admission Counseling describes holistic review as evaluating "the whole person" using multiple criteria, noting that transparent holistic review produces fairer outcomes than opaque processes relying on unstated criteria. "Holistic admission is an institutional value or priority communicated to those who review applications, for whom it forms the basis for weighing a range of factors to evaluate the applicant." The scorecard embeds this value into its architecture.

Academic Factors Are the Core – But Not the Whole

Across multiple years of NACAC survey data, the pattern is consistent: academic factors rank first. In 2013, 78% of colleges rated grades in college-prep courses as considerably important – the top factor, ahead of curriculum strength (60%), overall GPA (57%), and test scores (51%). In 2010, the figure was 82%. "Grades in college preparatory courses remain the most highly rated factor in the admission decision across all institution types," NACAC concluded.

Yet the same surveys make clear that academic factors alone do not tell the whole story. Teacher recommendations, counselor letters, essays, extracurricular activities, and demonstrated interest all registered as moderately or considerably important across a meaningful share of institutions – especially selective ones. At colleges accepting fewer than half their applicants, essays were rated considerably important by 51% of respondents and teacher recommendations by 40%. "Academic performance in high school has been the most important consideration in freshman admission decisions for decades" – but the rest of the application matters too.

The scorecard reflects this evidence precisely. Academic dimensions (D01 + D02 + D03) collectively carry 45% of the total composite weight – consistent with NACAC analysis showing that academic factors receive in actual admissions decisions at selective colleges. The remaining 55% is distributed across personal qualities (D04, 12%), extracurricular depth (D05, 10%), college knowledge (D06, 8%), self-awareness and mindset (D07, 8%), academic skills (D08, 7%), college motivation (D09, 5%), and financial preparedness (D10, 5%). Every dimension earns its weight through research.

Part II: Why GPA Is the Cornerstone of D01

GPA Is the Strongest Individual Predictor of College Outcomes

The case for placing GPA at the center of the Academic Achievement dimension is, by any reasonable standard, overwhelming.

Allensworth and Clark's landmark 2020 study in Educational Researcher found that high school GPA is the strongest single predictor of first-year college GPA across multiple meta-analyses – more predictive than standardized test scores, class rank, or any other commonly available metric. Students with a high school GPA above 3.5 have first-year college retention rates 15-20 percentage points higher than students with a GPA below 3.0. Predictive correlations for 6-year graduation rates range from 0.40 to 0.50, holding across institutional selectivity levels and racial/ethnic groups. Critically, GPA is more predictive for students from lower-income backgrounds than standardized test scores – a finding with significant equity implications.

Bowen, Chingos, and McPherson's Crossing the Finish Line (Princeton University Press, 2009), one of the most comprehensive longitudinal studies of American higher education, corroborated these findings. High school GPA predicts first-year college GPA with correlations of 0.50-0.60, 4-year graduation with correlations of 0.35-0.45, and 6-year graduation with correlations of 0.40-0.50. These are not modest effects. They reflect GPA's status as a multi-year, multi-teacher, multi-course summary of academic performance – a signal more stable and comprehensive than any single standardized test can provide.

A longitudinal study by Swanson and Cole (2022) tracking 793 low-income, high-achieving students confirmed that an average high school GPA of 3.6 significantly predicted college cumulative GPA outcomes through the first three years of college ($p<0.05$ to $p<0.1$). The study used college GPA as the dependent variable, a methodological choice that itself reflects the research community's consensus: GPA represents academic achievement and college readiness as both cause and effect.

GPA as a Prerequisite – Not Just a Predictor

Klasik's 2012 analysis of the ELS:2002 dataset, spanning nearly 9,000 students, identified "attaining minimal academic qualifications" – measured primarily through GPA – as one of five critical steps to four-year college enrollment. The statistic is striking: 95% of students who ultimately enroll in four-year colleges complete this step. Academic preparation is not merely predictive of success in college; it is constitutive of access to college in the first place.

Klasik also found that racial and family income gaps in college enrollment are largely explained by differences in academic preparation, and that Black students are 19 percentage points less likely than White students to attain minimal academic qualifications. This underscores why D01, the academic achievement dimension, cannot be underweighted or treated as interchangeable with softer dimensions. For students from historically underserved backgrounds, academic preparation gaps are real and consequential – and a

scorecard that obscures them through excessive weight on harder-to-measure noncognitive factors does students a disservice by failing to surface where intervention is most needed.

Why Unweighted GPA – And Not Weighted

The scorecard uses unweighted GPA as the primary component of D01 (accounting for 60% of the dimension score). This choice reflects both fairness and psychometric rigor.

Pattison, Grodsky, and Muller's research in the NASSP Bulletin (2014) documents the chaotic variability in weighted GPA systems across American high schools. Some schools add 1.0 for AP courses; others add 0.25; others add nothing. Schools in affluent districts offering 30+ AP courses give their students more opportunities to boost weighted GPAs than students at under-resourced schools where the most rigorous available option might be a single honors class. The result: weighted GPA systematically advantages students from well-resourced schools without necessarily reflecting greater academic achievement.

This is not theoretical. The Herzstein School – an independent school using a non-standard 8.0-point grading scale with +0.25 for AP and +0.1 for Honors – produces GPAs that are mathematically incomparable to a suburban public school using a 4.0 scale with +1.0 for AP. Northern Valley Regional High School reports weighted GPAs ranging from 1.47 to 5.03, with a median of 3.94. Sequoia High School, serving a predominantly Latino student body (57.4%) with 36% on free/reduced lunch, provides weighted, unweighted, and UC-specific A-G GPAs on the same transcript – three different numbers for the same academic record.

College admissions officers navigate this complexity by recalculating GPA using their own standards, typically focusing on unweighted core-course GPA. Rice University admissions officer Dr. Accrocco put it plainly: "Average GPA is a totally meaningless number without the high school context." The scorecard's use of unweighted GPA acknowledges this reality. Course rigor – the dimension that rewards students who challenge themselves with AP, IB, and honors coursework – is assessed separately in D03, where it can be evaluated on its own terms rather than conflated with raw grade points.

Part III: Why Class Rank Belongs in D01

Rank Provides Context That GPA Alone Cannot

A GPA of 3.7 means something very different at a school where 60% of students have GPAs above 3.5 than at a school where fewer than 15% do. Class rank – or its functional equivalents: percentiles, quintiles, deciles – provides the school-specific context that makes GPA interpretable as a signal of genuine academic achievement rather than a reflection of grade inflation.

NACAC's national survey data supports the scorecard's decision to include rank at 40% of D01 while subordinating it to GPA at 60%. The 2018 survey found that 75% of colleges rated overall GPA as considerably important, while only 9% rated class rank at the same level – down from 23% in 2007. "Class rank importance dropped from 23% (2007) to 9% (2018)

[rating it as considerably important]," the report noted. Yet 31% of colleges still rated rank as moderately important, and its contextual function in interpreting GPA across diverse school environments has not diminished even as its formal use in admissions has.

At highly selective colleges, 90-95% of admitted students rank in the top 10% of their high school class when rank data is available. At selective colleges accepting 20-50% of applicants, 70-85% of admitted students are in the top quartile. These distributions provide empirical grounding for the scorecard's class rank score mapping: Top 10% earns a score of 9; Top 25% earns a 7; Top 50% earns a 5; lower quintiles earn 3 or 1, respectively.

The Ivy League Academic Index as Precedent

The practice of combining GPA and class rank into a composite academic metric has well-established institutional precedent. The Ivy League schools historically used an Academic Index (AI) that combined GPA, class rank, and test scores using explicit weighting formulas – typically giving GPA and rank roughly equal weight, with test scores as a third component. CollegeVine and Solomon Admissions document that some institutions converted class rank to a numerical score and averaged it with converted GPA and test scores. This institutional precedent supports the scorecard's 60/40 GPA-to-rank split within D01, with GPA receiving slightly greater weight given its superior predictive validity and universal availability.

In practice, as training materials from Rice University's Center for College Readiness document, admissions officers always evaluate GPA and rank together as a paired measure. Real application examples confirm this: Wendy Weiss's 92.65 unweighted GPA in the 2nd decile, Maggie McMurty's 9.29 weighted GPA (non-standard scale) in the 1st quintile, Bernice Baker's 88.75 weighted GPA in the 3rd decile – all evaluated jointly, never in isolation. "A high GPA with low rank suggests grade inflation, while a moderate GPA with top rank suggests a competitive school."

Handling Schools That Don't Report Rank

The declining availability of class rank is a real challenge. About 50% of high schools still calculate formal numerical rank; the other half have eliminated it, often due to concerns about student stress or because minimal differentiation at the top makes ordinal ranking feel arbitrary. The scorecard's response to this reality is built into its design: when rank data is unavailable, the class rank component uses available quintile or percentile information where provided. The 60/40 GPA-to-rank weighting ensures the dimension retains validity even when rank is partially unavailable, since the majority weight rests on the universally available unweighted GPA.

Part IV: The Grade Trend Modifier – Small Signal, Real Meaning

Why Trajectory Matters in Holistic Review

The D01 dimension includes a Grade Trend Modifier of +/-0.5 – a small adjustment to the composite dimension score that reflects whether a student's academic performance improved, held steady, or declined across high school.

This modifier is intentionally modest: +/-0.5 on a 9-point dimension score represents a swing of roughly 5.6% in either direction. It cannot compensate for a weak GPA or low class rank, nor does it attempt to. Its function is narrower: to distinguish between two students with identical GPAs who arrived at that GPA via very different trajectories.

Admissions officers consistently report that upward grade trends – a student who earned B's in 9th grade but A's by 11th grade – are viewed favorably, even when cumulative GPA is lower than a student who peaked early. IvyWise and CollegeReadyPlan, drawing on practitioner experience, note that upward trends signal resilience, adaptation, and academic maturation – traits that predict continued growth in the more demanding academic environment of college. Conversely, downward trends – particularly the "senior slump" or sharp grade decline in 12th grade – raise genuine concerns about motivation and readiness for college-level work.

The modifier formalizes what holistic reviewers have always done informally: give modest credit for trajectory, without allowing it to override cumulative record.

Part V: The GPA Score Mapping – Empirically Grounded Thresholds

Not Arbitrary Cutoffs – Calibrated to Real Admissions Outcomes

The scorecard maps unweighted GPA to a 1-9 dimension score using five thresholds: 3.8+ earns a 9; 3.5-3.79 earns a 7; 3.0-3.49 earns a 5; 2.5-2.99 earns a 3; below 2.5 earns a 1.

These thresholds are calibrated to actual admissions outcomes across the institutional selectivity spectrum, as documented in NACAC's national survey data and Common Data Set filings:

GPA 3.8+ (Score = 9): Median admitted student GPA at highly selective institutions (acceptance rate <20%) is typically 3.8-4.0 unweighted. A score of 9 correctly signals eligibility for the most competitive tier.

GPA 3.5-3.79 (Score = 7): Median admitted student GPA at selective institutions (acceptance rate 20-50%) falls in this range. A score of 7 signals strong preparation for selective college admission.

GPA 3.0-3.49 (Score = 5): Median admitted student GPA at moderately selective institutions (acceptance rate 50-75%). A score of 5 – the midpoint on the 1-9 scale – reflects adequate but not exceptional preparation.

GPA 2.5-2.99 (Score = 3): Students in this range are primarily admitted at less selective or open-access institutions. A score of 3 signals developing preparation requiring significant support.

GPA below 2.5 (Score = 1): A minimum GPA of 2.0 is required for NCAA eligibility and for admission to most four-year institutions. Scores below 2.5 signal foundational readiness gaps.

This mapping means the scorecard's GPA thresholds carry real-world meaning. A student scoring a 7 on D01 is not simply "above average" in some abstract statistical sense – they are demonstrating the academic achievement profile associated with admission to selective four-year institutions. A student scoring a 9 is competitive for highly selective institutions. The score is interpretable because it is anchored to something real.

Grade Inflation and Why Context Remains Essential

The national trend toward grade inflation – documented by Gershenson's 2018 Fordham Institute analysis showing the percentage of students earning A averages rose from 28% to 47% between 1990 and 2020 – reinforces both the need for class rank as a contextualizing factor within D01 and the separate assessment of course rigor in D03. As more students arrive at the same GPA destination, the path matters more. The scorecard's architecture – unweighted GPA in D01, course rigor assessed separately in D03 – handles this correctly.

Part VI: D01 in the Broader Scorecard Architecture

The 20% Weight – Highest, But Not Dominant

D01 carries the highest individual weight in the scorecard at 20%. This is appropriate. GPA is the strongest predictor of college success across 50+ years of predictive validity research. NACAC surveys consistently show it is the factor rated most important by the highest share of college admissions offices in every major national survey going back decades. It functions as a gating prerequisite for college enrollment for 95% of students who ultimately attend four-year institutions.

But 20% is not 100%. It is not even a majority. The combined weight of academic dimensions D01, D02, and D03 is 45% – consistent with NACAC analysis showing academic factors account for roughly 50-70% of decision weight at selective colleges. The holistic remainder – personal qualities, activities, college knowledge, mindset, skills, motivation, and financial preparedness – accounts for 55%.

This architecture serves students well. A student with a lower GPA who demonstrates exceptional growth mindset (D07), strong self-regulated learning skills (D08), and genuine college motivation (D09) can still earn a meaningful composite score. A first-generation student whose GPA reflects the limits of an under-resourced school, not their intellectual capacity, is not defined by that single number. As NACAC's equity framework notes, holistic review "can be seen as an institutional commitment to enrolling students who have demonstrated potential through both traditional and non-traditional means."

An Equity Note on Academic Achievement

The scorecard's equity commitment is visible precisely in how it handles D01. It uses unweighted GPA – the most comparable metric across wealthy and under-resourced schools – rather than weighted GPA, which systematically advantages students from AP-rich environments. It includes class rank as a contextualizing factor. It separates course rigor into its own dimension (D03), so students at schools with limited AP offerings are not penalized in D01 for what their schools did not offer.

At the same time, the scorecard does not pretend academic achievement gaps do not exist. Klasik's research found that Black students are 19 percentage points less likely than White students to attain minimal academic qualifications for four-year college enrollment. This gap is real, it is consequential, and it is exactly the kind of information a college readiness tool should surface – not to limit students, but to identify where intervention, preparation, and support are most needed.

Conclusion: Evidence, Transparency, and the Purpose of D01

The case for D01 as the highest-weighted dimension in the College Readiness Scorecard rests on a convergence of evidence that is unusually robust for the field of educational assessment. GPA is the strongest single predictor of first-year college GPA, retention, and graduation. It is the factor rated most important by the largest share of college admissions offices in every major national survey going back decades. It functions as a gating prerequisite for college enrollment for 95% of students who ultimately attend four-year institutions. Its unweighted form provides the fairest cross-school comparison available. And its combination with class rank, in the 60/40 split used by D01, mirrors the paired evaluation that trained admissions officers have always performed.

The grade trend modifier adds a final, carefully constrained acknowledgment that trajectory matters – that a student who is still growing deserves credit for that growth, and a student who is declining should be understood in that context.

D01 earns its 20% weight. And the scorecard's broader architecture ensures that a strong D01 score is meaningful – anchored in real admissions outcomes – without that score alone being decisive. That is precisely the balance that a rigorous, equitable, evidence-based college readiness tool should strike.

Chapter 2: Standardized Testing (D02) – 15% Weight

Part of the College Readiness Scorecard Research Series – D02: Standardized Testing (15% Weight)

Introduction: The Test Score Question

Few questions in contemporary education spark more debate than the role of standardized tests in college readiness. In the wake of COVID-19, more than 1,800 colleges adopted test-

optional or test-blind policies, prompting widespread questions about whether the SAT and ACT remain meaningful measures of student preparation. The College Readiness Scorecard answers this question with nuance: yes, standardized testing belongs in any serious readiness framework – but at a carefully calibrated weight that reflects both its genuine predictive power and its well-documented limitations.

Dimension 2 (D02) of the College Readiness Scorecard assigns standardized testing a 15% weight – second among the scorecard's ten dimensions, but trailing Academic Achievement (D01) at 20%. This article explains why that allocation is precisely right: grounded in decades of psychometric validity research, national admission practice data, and a clear-eyed view of equity concerns that make test scores a valuable-but-limited signal of college readiness.

Part I: Why Standardized Testing Belongs in the Framework

1.1 Tests Predict College Performance – Meaningfully, If Not Decisively

The foundational case for including standardized testing in any college readiness assessment is empirical: test scores genuinely predict college outcomes. The evidence is extensive and consistent across more than two decades of research.

The College Board's 2013 validity study, drawing on 148,768 students across 150 institutions, found a correlation of r = 0.54 between SAT Total scores and first-year college GPA. The ACT's comparable research across 125,000 students found a Composite-to-GPA correlation of r = 0.53. These are not trivial numbers. They mean that, all else equal, a student's standardized test score provides a meaningful signal about how that student is likely to perform academically in their first college year.

Importantly, this predictive power does not fade after freshman year. College Board's landmark 2024 study – tracking SAT scores against GPA through all four years of college – found that the incremental validity of SAT scores above high school GPA remained consistent from Year 1 through Year 4 (Delta R-squared approx. 0.06 throughout). As the study concluded: "The added predictive value of SAT scores above HSGPA is stable through analyses of each yearly cumulative GPA."

Perhaps most compellingly, research shows that test-based readiness benchmarks predict degree completion, not just GPA. College Board's 2013 longitudinal study tracked 898,517 students through the National Student Clearinghouse and found that students who met the SAT College Readiness Benchmark were nearly twice as likely to graduate in four years (58% vs. 31%) and had substantially higher six-year graduation rates (69% vs. 45%). These advantages persisted across every demographic subgroup examined.

This is the bedrock justification for D02: standardized tests, whatever their limitations, are predictive of actual college outcomes at meaningful statistical magnitude.

1.2 Tests Add Information Beyond GPA Alone

A critical finding across the validity literature is that test scores and high school GPA measure overlapping but distinct constructs. When combined, they explain substantially more variance in college performance than either measure alone.

Westrick et al.'s 2019 College Board study (n = 180,000 students, 201 institutions) found that HSGPA alone explained 37% of first-year GPA variance (R-squared = 0.37), while SAT scores alone explained 33% (R-squared = 0.33). Combined, the two measures explained 43% – a meaningful gain that justifies treating them as separate dimensions rather than collapsing them into a single academic score. ACT research yields nearly identical results: combined ACT + HSGPA produces R-squared = 0.43 vs. HSGPA alone at 0.37.

The ACT's broader Multiple Measures Framework reinforces this logic. A study of 358,920 students found that academic and non-academic readiness measures captured meaningfully distinct aspects of student preparation, with a correlation of 0.76 between ACT Composite and career-readiness assessments – strong, but far from redundant. In the scorecard context, this principle justifies assigning D02 its own 15% weight rather than folding testing into D01.

1.3 Admission Officers Still Value Tests – Though Less Than Before

Empirical data on actual admission practice confirms that standardized testing remains a significant factor even in an era of expanding test-optional policies. NACAC's 2019 State of College Admission survey found that 46% of colleges rated SAT/ACT scores as "considerably important" in Fall 2018 admission decisions.

That 46% figure places test scores squarely in the top tier of admission factors, alongside grades in all courses (75%), grades in college prep courses (73%), and strength of curriculum (62%). As NACAC summarized: "Academic performance in high school has been the most important consideration in freshman admission decisions for decades. Seventy-five percent of colleges rated grades in all high school courses as considerably important… nearly half gave admission test scores (ACT/SAT) considerable weight."

Niche's 2022 national survey of 21,866 college-bound seniors confirms that testing remains widespread in practice: 75% of students took standardized tests despite widespread test-optional policies, and 66% believed submitting scores helps their admission chances even when not required. This is not a dying metric; it is a recalibrated one.

Part II: Why Testing Is Weighted at 15%, Not Higher

The case for inclusion is clear. The case for calibration – for why D02 sits at 15% rather than 25% or 30% – rests on an equally substantial body of evidence demonstrating test scores' limitations relative to GPA, their susceptibility to socioeconomic inflation, and the evolving professional consensus toward reduced test emphasis.

2.1 GPA Is the Stronger Predictor – By a Wide Margin

The most decisive evidence for calibrating D02 below D01 comes from Allensworth and Clark's landmark 2020 study published in Educational Researcher. Analyzing 55,084 Chicago Public Schools graduates tracked into four-year colleges, the study found a striking result: "Students' high school grade point averages are five times stronger than their ACT scores at predicting college graduation."

This is not a close call. GPAs predicted college graduation with consistency across all high schools in the study; ACT score predictive validity varied widely by high school and was, in some cases, negatively correlated with graduation among highest scorers at certain schools. The authors explain why: "GPAs measure a very wide variety of skills and behaviors that are needed for success in college, where students will encounter widely varying content and expectations. In contrast, standardized tests measure only a small set of the skills that students need to succeed in college."

This evidence alone justifies the 20% (D01) vs. 15% (D02) allocation. NACAC's longitudinal survey data confirms the same pattern in admission practice: GPA importance among admission officers rose from 52% to 75% between 2007 and 2018, while test score importance fell from 59% to 46% – a 29-percentage-point gap that mirrors the scorecard's weight differential.

2.2 Tests Explain Only a Fraction of College Success Variance

Peer-reviewed research reinforces this calibration. Maruyama's 2024 study in Analyses of Social Issues and Public Policy found that ACT Composite scores and background variables together predicted approximately 12% of variance in college GPA. Meaningful – but modest. Twelve percent explained variance in a complex, multifactorial outcome is precisely what a 15% dimension weight implies: significant enough to include as a discrete measure, but clearly subordinate to more comprehensive indicators.

The College Board's own research is candid about this: the incremental validity of the SAT is Delta R-squared approx. 0.06. Adding SAT scores to a prediction model already containing high school GPA moves from R-squared = 0.37 to R-squared = 0.43. That improvement in classification accuracy (4-6% for identifying at-risk students) is real and useful – it is also not the dominant signal in a multi-dimensional readiness model.

2.3 Test Scores Reflect Socioeconomic Privilege, Not Just Academic Preparation

Beyond predictive validity, standardized tests carry a well-documented equity concern: they partially reflect access to resources rather than academic potential. Reardon's 2011 study documented an average SAT score gap of approximately 250 points between students from top- and bottom-income quintiles – a gap that has widened by 40% over three decades. This gap persists even after controlling for academic preparation and course-taking.

A systematic review of 70+ empirical studies found that while SAT/ACT are statistically reliable, they are "not equally valid or predictive for all students across gender, race, and

socioeconomic demographics." Tests "systematically favor students with more access to high-quality schooling, stable socioeconomic conditions, and test prep opportunities" – some of which can cost thousands of dollars.

Nankervis's 2011 study added a specific gender dimension: male students score approximately 35 points higher on SAT Math on average (approximately one-third of a standard deviation), yet the SAT under-predicts female college success and over-predicts male success. Stereotype threat accounts for approximately 60% of this gap. Rigid score cutoffs create male-to-female admission eligibility ratios ranging from 1.05 to 1.65 at selective institutions.

These equity realities inform two specific design choices in D02: using continuous linear interpolation across score bands (rather than hard cutoffs) and weighting testing at 15% rather than 20-30%, ensuring that socioeconomic disparities in test access do not dominate the composite readiness score.

2.4 The Professional Consensus Has Shifted – But Not to Zero

Harvard's Making Caring Common project's "Turning the Tide" report (2016), endorsed by admissions deans from Harvard, MIT, Yale, Princeton, Stanford, and 70+ other selective institutions, recommended that admissions tests "be optional or clearly contextualized" and that "students be discouraged from taking standardized tests more than twice." This landmark document reflects the professional consensus: test scores are one signal among many, not a defining criterion.

The test-optional movement's growth to 1,800+ colleges confirms this shift at institutional scale. Research consistently finds that students who do not submit scores at test-optional institutions achieve nearly identical college outcomes: Syverson, Franks, and Hiss (2018), examining 955,000 students across 33 test-optional institutions over a decade, found that non-submitters had only 0.05 GPA point lower first-year averages than submitters – a negligible difference – and similar graduation rates.

The 15% weight for D02 reflects this professional recalibration. Testing is neither irrelevant (it adds real predictive information) nor paramount (it is now one of many factors, not the first filter).

2.5 Grade Inflation Makes Testing a Useful Complement

One underappreciated justification for maintaining D02 at a meaningful weight – even as test scores decline in relative importance – comes from the grade inflation literature. Schaeffer and Reeves (2020) documented that the percentage of high school students with A averages (3.7+ GPA) increased from 28% in 1998 to 47% in 2016, without commensurate increases in SAT scores. Grade inflation varies widely across high schools, making GPA comparability challenging across contexts.

Standardized tests, whatever their limitations, provide a common scale that is less susceptible to school-specific grading practices. This justifies D02's role as a complement to D01, not a redundant measure. Together, the two dimensions (academic achievement and

testing) capture what GPA-alone assessments cannot: both the school-contextualized signal of academic effort (D01) and the cross-contextual benchmark signal of academic skill (D02).

Part III: Why the Default Score of 3 Is the Right Design Choice

For students who do not submit standardized test scores, the College Readiness Scorecard assigns a default D02 score of 3 – a deliberate design choice with a clear empirical rationale.

3.1 Most Students Are Not Fully Test-Ready

ACT's longitudinal condition reports provide the statistical anchor for this default. In 2009, only 23% of high school graduates met all four ACT College Readiness Benchmarks; 77% fell below benchmark in at least one subject. A score of 3 on the 1-9 scorecard scale falls in the "Developing" range – reflecting, accurately, that most students without test data are likely to fall in the majority of their peers, not in the exceptional quartile.

Common App data from 2020-21 (740,000+ applicants) confirms the calibration: students who submitted test scores to test-optional institutions had admission rates of 54.8% vs. 51.2% for non-submitters – a 3.6 percentage point advantage. Non-submitters self-selected strategically (often skipping submission because scores were below institutional averages), which supports a default in the mid-to-low range.

Syverson et al.'s large-scale study reinforces the default-at-3 choice: non-submitters at test-optional colleges achieved first-year GPAs only 0.05 points lower than submitters. This finding supports treating non-submission as a signal of "developing but not deficient" readiness – which a score of 3 accurately represents.

3.2 The Default Protects Equity Without Distorting the Score

Niche's 2022 survey found that first-generation, underrepresented minority, and low-income students are significantly less likely to take or submit standardized tests. Klasik's 2013 research showed that mandatory state ACT testing increased college enrollment among first-generation and low-income students by 6-8 percentage points – suggesting that many of these students simply lack access to testing infrastructure, not academic ability.

A default score of 3 ensures that students from under-resourced backgrounds who do not test are not penalized with a score of 1-2 (which would imply deficiency) while also not inflating their readiness estimate to 5+ (which would misrepresent the absence of data). The default is honest, protective, and calibrated to the empirical distribution of student readiness.

Part IV: The max(SAT, ACT) Design – Fairness Through Student Choice

The scorecard calculates D02 using the higher of a student's SAT or ACT score. This design reflects both psychometric evidence and equity principles.

The College Board-ACT official concordance tables (2018) establish that SAT and ACT scores are percentile-equivalent at matched score levels: SAT 1200 is approximately ACT 24 (75th percentile); SAT 1400 is approximately ACT 31 (95th percentile). Using concordance-validated equivalence, the scorecard can treat SAT and ACT interchangeably without penalizing students for format preference. The correlation between SAT and ACT validity findings (r approx. 0.53-0.54 against FYGPA for both tests; Delta R-squared approx. 0.06-0.07 over HSGPA for both) confirms that both instruments measure sufficiently overlapping constructs to support equivalent treatment.

The max() approach serves students who perform differently across test formats – which is common, given the SAT's emphasis on evidence-based reading and the ACT's inclusion of a Science section. Taking the higher score respects student choice, aligns with college counselor recommendations, and avoids penalizing students for choosing one valid instrument over another.

Nankervis's gender equity research adds a specific justification: since the SAT Math section has documented differential validity by gender, allowing students to use ACT scores (which show smaller gender gaps on the science and math composites) provides an additional equity protection for female test-takers.

Part V: D02 Score Thresholds – Grounded in Validated Benchmarks

The scorecard maps SAT/ACT scores to D02 levels (1-9) using empirically anchored thresholds that correspond to college readiness benchmarks, institutional selectivity tiers, and documented graduation rate differentials.

D02 Score	SAT Range	ACT Range	Benchmark Context
9	1500+	34+	Top 1-3% nationally; highly selective admission range
8	1400-1499	31-33	~95th percentile; strong readiness across all majors
7	1300-1399	28-30	~88th percentile; competitive at most four-year institutions
6	1200-1299	26-27	~75th percentile; solid readiness
5	1100-1199	23-25	Near ACT benchmark average; adequately prepared
4	1000-1099	20-22	Below average; some readiness gaps
3	900-999	17-19	Near/below benchmarks; significant readiness gaps
2	800-899	14-16	Well below benchmarks; remediation likely
1	Below 800	Below 14	Severely below readiness thresholds

These thresholds are calibrated to ACT's College Readiness Benchmarks (English 18, Math 22, Reading 22, Science 23), College Board's SAT benchmark, and Bound et al.'s research showing that SAT score bands predict college access and six-year graduation rates at meaningfully differentiated levels. A score of 5 (adequate preparation) maps to the range where students approach – but have not fully cleared – all benchmark thresholds, consistent with the ACT's finding that only 38% of 2018 graduates met all four benchmarks.

The use of continuous linear interpolation within bands directly addresses the gender equity concern documented by Nankervis: rigid cutoffs systematically disadvantage female students whose SAT Math scores underpredict their college performance. Smooth interpolation ensures that a student scoring 1198 is not treated categorically differently from one scoring 1200.

Part VI: D02 Within the Larger Framework

6.1 The 10-Dimension Composite: Testing as One Signal Among Many

The College Readiness Scorecard's treatment of standardized testing is inseparable from its broader multi-dimensional architecture. The NACAC holistic review framework establishes that college readiness requires simultaneous evaluation of GPA, course rigor, testing, personal qualities, extracurriculars, and contextual factors. No single metric determines readiness; all contribute.

Soland's (2017) data-driven study using 1,000+ predictors and a sample of 12,144 students confirmed this empirically: academic preparation (GPA + test scores combined) is one of four core readiness constructs – alongside postsecondary aspirations, socioeconomic context, and teacher perceptions – that together predict college enrollment and persistence with 90% accuracy. Tests contribute to the academic preparation construct, but their contribution is bounded by the other three.

Conley's Four Keys to College Readiness framework (2007) similarly positions academic knowledge and skills (which includes test performance) as one of four essential domains, alongside cognitive strategies, academic behaviors, and contextual skills. The scorecard's 10 dimensions operationalize all four Conley domains, with D02 serving as the quantified academic-skills indicator that complements D01's broader academic achievement measure.

6.2 Testing's Equity Limitations Justify the Full Multi-Dimensional Design

The equity limitations of standardized tests – documented across race, income, and gender – are among the strongest arguments for a multi-dimensional scorecard rather than a test-centric admission metric. Sedlacek's noncognitive variable research demonstrated that noncognitive dimensions predict college success "for students of color, women, first-generation students, and others with nontraditional backgrounds" better than tests alone. The Gates Millennium Scholars program's 80%-noncognitive selection model achieves 97% first-year retention – far above national averages – by reducing test overemphasis.

The scorecard's architecture reflects this logic: D02 (15%) sits within a composite that reserves 45% for academic factors (D01 + D02 + D03) and 55% for personal, behavioral, and contextual dimensions (D04-D10). A student with limited test access but strong grades, rigorous coursework, demonstrated extracurricular depth, and active college preparation can achieve a strong composite score despite a modest D02. That is precisely the outcome a well-designed readiness framework should enable.

6.3 The NOSCA Framework: Testing as Component Five of Eight

College Board's NOSCA Eight Components of College and Career Readiness Counseling positions standardized assessments as Component 5 of eight essential readiness elements – confirming testing's place within, not above, a comprehensive readiness model. NOSCA's data elements for Component 5 include PSAT/PLAN participation and performance, SAT/ACT participation and performance, and career/interest assessment participation.

The NOSCA framework's emphasis on equity – "giving every student what they need to succeed, not just equal access" – aligns with the scorecard's use of continuous scoring and default values rather than cutoffs. Its advocacy for data-driven examination of group disparities validates the scorecard's approach of tracking D02 scores across demographic groups to monitor for differential impact.

Conclusion: A Calibrated, Evidence-Based Approach

The College Readiness Scorecard's treatment of standardized testing represents the product of a deliberate, evidence-anchored process. D02's 15% weight is neither a concession to test-optional fashion nor an uncritical embrace of testing orthodoxy. It is the result of integrating:

Predictive validity research showing test scores reliably predict college outcomes (r approx. 0.53-0.54; Delta R-squared approx. 0.06 above GPA). Comparative evidence demonstrating GPA's superiority as a graduation predictor (five times stronger, per Allensworth and Clark). Admission practice data confirming that 46% of colleges still weight tests considerably, but that GPA has widened its lead. Equity research documenting systematic SES and gender disparities in test access and validity. Test-optional outcome studies showing non-submitters perform nearly identically to submitters at the college level.

The result is a dimension that honors testing's genuine predictive contribution while preventing it from dominating a readiness assessment in ways unsupported by evidence. The default score of 3 for non-test-takers, the max(SAT, ACT) design, the continuous interpolation approach, and the 15% weight all serve the same goal: a standardized testing dimension that is rigorous without being reductive, significant without being determinative, and fair across the full spectrum of students who will use the scorecard.

Chapter 3: Course Rigor (D03) – 10% Weight

A Research-Backed Defense of the D03 Dimension in the College Readiness Scorecard

Introduction: The Limits of a Single Number

For decades, college admission officers and education researchers have grappled with the same fundamental question: what does it actually take for a student to succeed in college? The instinctive answer – look at the GPA – is incomplete. A 3.8 earned through a schedule of electives and standard-level courses tells a very different story than a 3.5 earned through

five AP classes, dual enrollment coursework, and a gauntlet of honors sequences. The grade point average, powerful as it is, cannot distinguish between those two students on its own.

This is precisely why the College Readiness Scorecard treats Course Rigor (D03) as a distinct, weighted dimension carrying 10% of the composite score – separate from Academic Achievement (D01, 20%) and Standardized Testing (D02, 15%). The scorecard's architecture mirrors what college admission offices have consistently told researchers for decades: the academic record is the foundation of any application, and within that record, the rigor of the course load is independently meaningful. This article marshals the evidence for that claim.

Part I: The Structural Logic – Why Rigor Deserves Its Own Dimension

1.1 Colleges Evaluate Rigor Separately from Grades

The empirical case for treating course rigor as a standalone dimension begins with the colleges themselves. Every year, the National Association for College Admission Counseling (NACAC) surveys hundreds of member institutions about which factors they weight most heavily in admission decisions. The results are strikingly consistent across time.

In NACAC's 2023 State of College Admission data, **63.8% of colleges rated the strength of a student's high school curriculum as considerably important** in their fall admission decisions – making it the third-highest-rated factor overall, trailing only grades in college-prep courses (76.8%) and overall GPA (74.1%). Crucially, an additional 22.7% rated it as moderately important, meaning **86.5% of institutions** regard the rigor of a student's course load as at least moderately important. That figure is comparable to the 91-93% who rate grades as moderate-to-considerable, and vastly higher than extracurricular activities (50.8%), test scores (30.3%), or work experience (33.0%).

This pattern is not a recent artifact of the test-optional movement. NACAC's longitudinal data shows curriculum strength rated at 60-64% considerable importance stretching back to 2012 – while test score importance collapsed from 46% in 2018 to under 5% by 2023. Course rigor has proven to be one of the most durable admission factors in modern higher education.

NACAC's own policy statement crystallizes the point: *"The rigor of a student's high school curriculum has been identified in numerous studies as a leading predictor of college success – far more so than standardized test results."*

The scorecard's 10% weight for D03 directly reflects this empirical reality. Rigor is not an afterthought; it is the third pillar of the academic case for any applicant.

1.2 Grades and Rigor Measure Different Things

A second reason to keep D03 separate from D01 is that the two dimensions are genuinely distinct, even when they correlate. David Conley's Four Keys to College Readiness framework, commissioned by the Bill & Melinda Gates Foundation, identifies **Key Content Knowledge** as one of four foundational pillars of readiness – explicitly separate from

academic performance metrics like GPA. Conley writes: *"Courses with the same titles can vary tremendously in content and rigor… simply looking at transcripts tells an incomplete story about what students know and can do."*

Conley recommends measuring content knowledge through completion of college-level coursework and performance on standardized exams measuring college-level content – precisely what AP, dual enrollment, and honors courses provide. His framework positions content knowledge as complementary to, not redundant with, academic achievement.

The admissions data from Harvard reinforces this distinction empirically: 74% of admitted students to the Class of 2027 had 4.0 unweighted GPAs, but those students were *not* achieving those grades through easy courses. As one admissions source notes, they were *"doubling, tripling up on AP or IB-level core subject classes in the place of electives."* A student who hasn't challenged themselves academically, however high their GPA, simply is not competitive at selective institutions.

The scorecard captures this reality by assigning D01 to measure the grade record and D03 to measure the challenge level of the courses producing those grades. Together they form a complete academic picture; neither alone is sufficient.

1.3 The Holistic Review Tradition Recognizes Curriculum as a Distinct Signal

Professional admissions training materials confirm the same separation. Dr. Accrocco's professional guide for the Center for College Readiness at Rice University notes that approximately **80% of the admission decision rests on academic metrics**, and within that, the course curriculum is evaluated on its own terms – relative to what the high school offers, and independent of the grades a student earns. A training example from Tivoli University (a fictional selective university used in professional certification courses) explicitly lists *"curriculum, grades, standardized testing, extracurricular activities, essays, and letters of recommendation"* as separately considered factors – reflecting how actual admissions offices think.

The scorecard operationalizes this professional consensus. By allocating 45% of the composite to the three academic dimensions (D01: 20%, D02: 15%, D03: 10%), and then explicitly separating rigor from achievement, it mirrors the analytic structure admissions officers apply when reading a transcript.

Part II: The Evidence Base – What Research Says AP and Rigorous Coursework Actually Do

2.1 AP Coursework Predicts College Success at Scale

The most comprehensive evidence base for AP as a college readiness indicator comes from College Board's own research program, which has tracked hundreds of thousands of students through postsecondary outcomes. Key findings from the 2019-2020 synthesis of this research:

Students who completed their first AP Exam were 3 percentage points more likely to graduate from college within four years. Scoring 3 or higher on one AP Exam: 6 percentage points more likely to graduate on time. Scoring 3 or higher on two AP Exams: 8 percentage points more likely to graduate on time. AP students who earned college credit-qualifying scores (3+) performed as well as or better than students who took the college's own introductory course, with mean GPAs 0.01 to 0.85 points higher depending on subject. AP students who earned college credit through their exam scores were **7 percentage points more likely to double major** – suggesting the rigor prepared them for higher academic engagement, not less.

Critically, this research was conducted on samples of 400,000+ students across 100+ four-year institutions, with controls for demographics, SAT scores, high school GPA, and parental education. The effects of AP participation are real, not artifacts of high-achieving students self-selecting into AP.

The scorecard's progressive threshold structure – where each AP or dual credit course adds 1.0 units of rigor toward the D03 score – is directly grounded in these findings. One AP course matters. Two matter more. The evidence supports treating each course as an incremental readiness signal.

2.2 Diminishing Returns After Five AP Courses: Why the Scorecard Does Not Over-Reward Extremes

A landmark College Board study (Beard, Hsu, Ewing, & Godfrey, 2023) followed 400,000+ students across three cohorts to ask a precise question: at what point does adding more AP courses stop generating additional college success gains?

The answer is clear: *"The largest predicted boost in college grades and on-time degree completion comes from a student from 0 to 1 AP and from 1 to 2 APs."* And critically: *"Taking and performing well on more than five AP Exams does not markedly alter first-year college grades and four-year degree completion."*

This research directly validates the scorecard's threshold structure, which rewards each additional AP course but does not linearly amplify scores beyond the 10+ range. A student with 15 AP courses does not receive a meaningfully different D03 score from one with 10 – because the research shows the marginal college readiness signal above that ceiling is minimal. The scorecard avoids rewarding the "AP arms race" that counselors and researchers have increasingly criticized.

2.3 The Honors vs. AP Weighting Ratio Is Calibrated to Industry Standards

The scorecard assigns 0.5 units of rigor for each honors course and 1.0 unit for each AP or dual enrollment course – a 2:1 ratio that is not arbitrary. It mirrors the industry-standard GPA weighting differential used by high schools across the country: *"AP courses are weighted an extra point (so an A would be a 5.0 on a 4.0 scale), while honors courses are weighted an extra half-point (an A equaling a 4.5)."* – CollegeVine

This differential reflects substantive differences between the course types. AP courses follow a standardized College Board curriculum, culminate in a nationally normed examination, and can earn transferable college credit (accepted by 85%+ of U.S. colleges for scores of 3+). Honors courses, by contrast, follow school-specific curricula that vary by teacher and institution, offer no standardized external assessment, and cannot earn college credit.

Elite schools like Harvard and Yale explicitly prefer AP over honors when both are available, because AP's standardization allows direct applicant comparison across schools. The 2:1 ratio in the scorecard accurately represents the 2:1 GPA bonus differential that high schools themselves have determined reflects the relative rigor gap.

2.4 Dual Enrollment Is at Least as Valuable as AP – and Sometimes More So

The scorecard treats dual enrollment courses as equivalent to AP (1.0 unit of rigor). This is supported by a growing body of evidence that dual credit outcomes are comparable to or exceed those of AP:

A rigorous quasi-experimental study by Giani, Alexander, and Reyes (2014), using statewide Texas longitudinal data with propensity score matching, found that *"dual-credit is a promising strategy for increasing the likelihood of students accessing, persisting through, and completing a degree in postsecondary, and is possibly even more impactful than advanced coursework."*

A 2025 Community College Research Center study analyzing ~500,000 ninth-graders found that dual enrollment participants were more likely to apply to and be admitted to **more selective colleges**, with particularly strong effects for Black students (15.6% more likely to apply to a four-year institution per dual credit course taken).

And on the financial aid front, a 2025 study of 8,690 students (Hu, Journal of Student Financial Aid) found dual enrollees were **1.7 times more likely to receive merit scholarships** in their first year of college – suggesting colleges themselves financially reward dual enrollment participation as a readiness signal.

With 2.5 million high school students now participating in dual enrollment nationally (up from 300,000 in the early 2000s), the 1.0 equivalence to AP is not merely justified – it is increasingly essential for a fair and comprehensive rigor assessment.

2.5 Even "Underprepared" Students Benefit from Rigorous Coursework

A recent peer-reviewed study (Lee, Bahena, & Nichols, 2025, Education Policy Analysis Archives) examined Texas students who took AP courses in 9th-10th grade despite being classified as "academically underprepared" based on prior test scores. The results challenge the gatekeeping logic often applied to rigorous coursework:

Early AP participation led to **higher scores on state exit exams**, greater enrollment in early college credit courses, and increased college attendance and completion rates. Effects were particularly strong for four-year college enrollment. The study concluded that relying

on prior test scores as a gatekeeper for AP access may systematically misclassify which students will benefit.

For the scorecard, this finding has an important implication: the D03 dimension should count AP and honors coursework for any student who completes it, regardless of their D01 or D02 scores. A student with a modest GPA who nevertheless pushed themselves through AP coursework is demonstrating something real and predictive. The scorecard honors this by evaluating rigor independently of the student's academic performance in those courses.

Part III: How the Scoring System Is Designed – And Why

3.1 The Scoring Formula

D03 is calculated by summing:

1.0 unit for each completed AP or Dual Enrollment course. **1.0 unit** for each in-progress AP or Dual Enrollment course (counted because junior-year courses-in-progress are highly visible to admissions officers during the primary application window). **0.5 units** for each Honors course.

This raw sum is mapped to the 1-9 scale with thresholds derived from research and competitive school data:

Raw Rigor Units	D03 Score	Interpretation
0	1	No rigorous coursework
~1-2 units	3	Minimal engagement with rigor
~3-5 units	5	Adequate – 2-4 honors-level courses
~6-8 units	7	Strong – competitive with typical college-prep profiles
10+ units	9	Exceptional – maximum rigor in core subjects

This structure reflects the research finding that the largest gains come from the first 1-2 AP courses (justifying meaningful score separation at the low end), while gains level off after 5 APs (justifying the threshold structure at the upper end rather than an open-ended linear scale).

3.2 Why In-Progress Courses Count

The NACAC Guide explicitly states that *"junior year grades are critical – colleges prefer steady performance and improvement over inconsistency."* During the primary application season (fall of senior year), colleges are evaluating transcripts that include junior-year courses but not yet-completed senior courses. By counting in-progress AP and dual enrollment courses at 1.0 units, the scorecard accurately reflects the signal that admissions officers are actually reading: a student enrolled in AP Chemistry and AP US History this year is demonstrating rigor even before exam day.

3.3 The 10+ Threshold Reflects Real Competitive Benchmarks

Multiple school profiles from competitive institutions validate the 10+ rigorous courses threshold as the boundary between "strong" and "exceptional" rigor:

New Trier High School (Illinois): *"Most rigorous programs include 5-10 AP courses"* **Northern Valley Regional (NJ):** Requires counselor approval for students taking more than 4 APs in 11th grade or 5 in 12th – suggesting 10+ total is the ceiling of responsible course loading **St. John's School (Houston):** Average of 4 AP courses per student; 100% AP participation; exceptional students push to 6-8 APs over their careers **Ivy Coach data:** Harvard's admitted students are *"doubling, tripling up on AP or IB-level core subject classes"* – consistent with 8-12 rigorous courses over a high school career

A student with 10+ rigorous course units has demonstrated sustained commitment to academic challenge across multiple years and subjects. The research evidence, competitive school benchmarks, and admissions officer surveys all converge on this level as the appropriate ceiling for the top score.

Part IV: Situating D03 Within the Broader Scorecard

4.1 The 10% Weight Is Precisely Calibrated

The scorecard's 10% allocation to D03 sits in a deliberate middle tier – below the 20% for Academic Achievement and 15% for Testing, but above the 5-8% weights for non-academic dimensions. This hierarchy maps to NACAC's four-tier structure of considerable importance:

Factor	NACAC Considerable Importance	Scorecard Weight
GPA / Grades in college-prep courses	74-79%	D01: 20%
Strength of curriculum	63.8%	D03: 10%
Personal qualities / character	28.3%	D04: 12%
Extracurricular activities	6.5%	D05: 10%

The 22-percentage-point gap between curriculum strength (63.8%) and the next-most-cited factor (character/personal qualities at 28.3%) is substantial. The scorecard's weight structure honors this gap by placing D03 in the academic tier, not the non-academic tier.

Professional admissions training materials consistently suggest that academic factors – grades, rigor, testing – constitute approximately 80% of the admission decision at selective colleges. The scorecard allocates 45% of its composite weight to the three academic dimensions (D01+D02+D03), representing a conservative version of that principle that preserves meaningful weight for the full range of holistic factors.

4.2 D03 Is the Bridge Between Achievement and Preparation

Of the three academic dimensions, D03 is unique in that it measures **what a student chose to attempt**, not just what they achieved. A student's D01 (GPA) measures how well

they performed. D02 (test scores) measures how they performed on a standardized instrument. D03 measures whether they sought out challenge in the first place – whether they enrolled in AP Chemistry, took dual credit English alongside their regular schedule, or pushed through an honors sequence despite the difficulty.

This motivational signal is not captured by GPA or test scores. It is precisely the kind of information that holistic admissions officers have long sought in the transcript review. The scorecard gives it quantitative form and explicit weight.

Part V: Equity Considerations

5.1 AP Access Gaps Are Real – and Accounted For

No honest treatment of the D03 dimension can ignore the equity dimension of AP access. EdWeek's 2024 analysis of College Board data shows that while 79% of public high school students attended schools offering at least 5 AP courses in 2022-23, Native American students were significantly underrepresented in this figure, and substantial racial and socioeconomic gaps persist in both access to and enrollment in AP courses even when they are available.

The scorecard acknowledges this reality in two ways. First, D03 scores should always be interpreted in the context of the student's school profile – a score of 5 earned in a district offering only two AP courses represents different rigor than a 5 earned at a school with 30+ AP options. Second, the scorecard's ten-dimension composite structure means that D03 carries only 10% of the total score – a significant but not dominant weight that allows students from under-resourced schools to compensate with strength in other dimensions.

Importantly, the research shows that AP benefits are not confined to privileged students. College Board's own data shows that **low-income students who take AP Exams have higher enrollment, persistence, and graduation rates than academically similar low-income peers who do not.** AP has expanded access during its growth period (share of exam takers using fee reductions increased by 19 percentage points between 2001 and 2020), suggesting that the rigor signal is increasingly available across socioeconomic strata – even if gaps remain.

5.2 Dual Enrollment as an Equitable Pathway

The scorecard's equal treatment of dual enrollment and AP (both weighted at 1.0) provides an equity-positive alternative for students whose high schools offer limited AP courses but partner with local community colleges for dual credit. With 2.5 million students now enrolled in dual credit programs nationally – concentrated in states with robust community college systems – dual enrollment represents a genuine pathway to college-level rigor for students who might otherwise score near 1 on D03 through no fault of their own.

The CCRC study's finding that Black students saw a 15.6% increase in four-year college application rates per dual credit course taken underscores that this is not just an academic equity argument – dual enrollment is an enrollment equity mechanism. The scorecard's

inclusion of dual credit at full AP-equivalent weight honors the research and supports broader college access.

Conclusion: Rigor Is Not Optional

The cumulative weight of evidence is unambiguous. Course rigor is:

Empirically predictive of college success outcomes – higher first-year GPA, greater four-year graduation rates, more advanced course completion – across samples of hundreds of thousands of students. **Rated as considerably important** by nearly two-thirds of colleges in every NACAC survey for over a decade. **Treated as a distinct factor** from grades and test scores in both professional admissions training and in the holistic review frameworks articulated by the nation's leading researchers. **Available through multiple pathways** – AP, dual enrollment, and honors courses – that are increasingly accessible across socioeconomic contexts. **Correctly weighted** at 10% of the composite score: substantial enough to meaningfully differentiate students, modest enough to preserve the holistic balance that allows students without full AP access to demonstrate readiness through other dimensions.

The College Readiness Scorecard's treatment of course rigor as Dimension 03, with a transparent scoring formula grounded in published research and competitive school benchmarks, represents exactly the kind of evidence-based, multi-factor approach that decades of admission research has called for. A student who has sought out challenge – who enrolled in AP courses before they were fully comfortable, who took dual credit English alongside their regular schedule, who pushed through an honors sequence when the standard track would have been easier – is demonstrating something admissions officers have always valued. The scorecard gives that signal the quantitative weight it deserves.

PART II: CHARACTER AND ENGAGEMENT

Chapter 4: Personal Qualities — The Right Dimension, the Right Weight, the Right Approach

Dimension D04 | Weight: 12% | Measurement: Default neutral 5 in standalone mode; LLM-based essay evaluation in backend mode

Introduction: The Science and Art of College Admissions

College admissions has always operated on two tracks simultaneously. There is the science – the accumulation of objective data points: GPA, test scores, class rank, course rigor – and there is the art, the interpretive work of reading a student's character through essays, recommendations, and the pattern of choices they have made. Any serious attempt to measure college readiness that ignores the second track is incomplete by design.

The College Readiness Scorecard addresses this with D4 (Personal Qualities), weighted at 12% of the total composite score. This chapter explains why that dimension exists, why 12% is the right number, how it is measured, and why the measurement approach – defaulting to a neutral 5 in standalone use, with LLM-based essay evaluation in the backend – reflects responsible assessment practice rather than a limitation.

The short answer: 70% of college admission officers nationwide say student character is considerably or moderately important to their decisions. Personal qualities are the critical differentiator at selective institutions. And a rigorous, peer-reviewed study published in *Science Advances* has now demonstrated that AI models can evaluate personal qualities from essays with strong predictive validity for six-year graduation rates – consistently across demographic groups.

Part I: Why Personal Qualities Are a Legitimate Readiness Dimension

The 80/20 Structure of Admissions Decisions

Admissions professionals at selective colleges regularly describe their process as "both art and science." The science – objective academic metrics – accounts for roughly 80% of the typical decision framework. The art – subjective evaluation of essays, recommendations, and personal narrative – accounts for the remaining 20%. As one instructor at the Rice Center for College Readiness explains: *"A lot of times admissions is often described as both art and science. So objective data you can look at as sort of the science. Now subjective data we are talking more about the art."*

This 80/20 split is not arbitrary. It is reflected in the NACAC national survey data. In 2019, 74.5% of colleges rated grades in college preparatory courses as considerably important, 62.1% rated strength of curriculum as considerably important, and 45.7% rated test scores as considerably important. Essays, counselor recommendations, and teacher recommendations – the primary vehicles for character assessment – were rated considerably important by 23.2%, 15.1%, and 14.2% of colleges respectively, and moderately important by an additional third of institutions each.

The College Readiness Scorecard directly mirrors this structure. Academic dimensions (D1 at 20%, D2 at 15%, D3 at 10%) account for 45% of the composite. D4 Personal Qualities at 12% is the highest-weighted non-academic dimension, positioned precisely at the inflection point between the science and the art of holistic review.

The Personal Factor: What Distinguishes Academically Equivalent Candidates

The National Association for College Admission Counseling is unambiguous about the role personal qualities play: "Application essays provide the 'personal factor' that distinguishes candidates at selective colleges." The NACAC guide elaborates: "Your uniqueness as an individual has an impact on the admission decision."

This is not a marginal or philosophical claim – it has direct practical consequences. At highly selective institutions, the pool of academically qualified applicants vastly exceeds

available seats. Among students with comparable GPAs and test scores, personal qualities become the decisive variable. A 2022 NACAC national survey found that 70% of admission officers rated student character attributes as either "considerably" or "moderately" important in the selection process. That majority consensus across hundreds of institutions is the empirical foundation for D4.

Ann Velenchik, Dean of Academic Affairs at Wellesley College, describes what admission readers are actually looking for: *"We want to admit students who are going to succeed and thrive here at Wellesley. We're trying to find evidence in a student's background that she has the academic preparation, character skills, study habits, and whole package of things that make somebody able to come to a reasonably high pressure environment, succeed, thrive, and be happy here."*

Personal qualities – character skills – are listed alongside academic preparation as co-equal predictors of the outcome colleges care about most.

Noncognitive Variables: The Research Case for Separate Measurement

The inclusion of personal qualities as a distinct scored dimension, rather than simply as a modifier of academic scores, is grounded in decades of noncognitive variable research. William Sedlacek's foundational work identified eight noncognitive dimensions predictive of college success: positive self-concept, realistic self-appraisal, understanding and navigating systems, long-range goals, strong support person, leadership, community involvement, and nontraditional knowledge acquired.

In a longitudinal study of 248 international students tracked across 8 semesters, Sedlacek and colleagues found that the Noncognitive Questionnaire (NCQ) significantly predicted both cumulative GPA ($p < .05$ across all 8 semesters) and persistence (63-75% correct classification). Critically, the NCQ demonstrated strong test-retest reliability (median .85 across items), validating that personal qualities can be measured systematically.

The Gates Millennium Scholars program applies this research at scale. Using noncognitive assessment as 80% of its selection weight across more than 11,000 scholars at 1,450+ institutions, it has achieved a 97% first-year retention rate, 87% five-year retention rate, and 78% five-year graduation rate – outcomes that substantially exceed national averages for comparable academic cohorts. This real-world validation demonstrates that personal quality assessment, when implemented rigorously, identifies students who succeed beyond what their academic credentials alone would predict.

Sedlacek's conclusion extends beyond any single population: *"Noncognitive variables provide viable alternatives in fairly assessing the abilities of people of color, women, international students, older students, students with disabilities, or others with experiences that are different from those of young, White, heterosexual, able-bodied, Eurocentric males."*

Part II: Why 12% Is the Right Weight

Calibrating the Weight to the Evidence

The 12% weight assigned to D4 is not a round-number estimate. It emerges from careful triangulation of multiple data sources.

First, the NACAC survey data establishes the tier structure of admission factors. Academic metrics (grades, curriculum, test scores) are considerably important to majorities of institutions – justifying 45% of the composite. Personal quality indicators (essays, counselor recommendations, teacher recommendations) are considerably important to 15-23% of institutions and moderately important to an additional 40%+ each – justifying a substantial but secondary weight.

Second, the roughly 20% "art" portion of admissions decisions needs to be distributed across several non-academic dimensions. D4 (Personal Qualities) shares that space with D5 (Extracurricular Depth, 10%) and other dimensions. At 12%, D4 captures the largest share of this space, reflecting that essays and character assessment are the primary vehicle for subjective evaluation.

Third, the NACAC 2023 survey provides a direct data point: 28.3% of colleges rated positive character attributes as considerably important, 37.5% as moderately important, for a total of 65.8%. This is roughly comparable to the importance of standardized testing (45.7% considerably important) – and the scorecard's weights reflect this: D2 Testing at 15%, D4 Personal Qualities at 12%.

Fourth, the weight is appropriately constrained by measurement limitations. As the Harvard Making Caring Common Project acknowledges, character assessment is genuinely difficult: "Colleges need new assessments not only to more accurately and fairly identify and weigh key non-cognitive, social, ethical, and emotional capacities, but also to send important signals to students about what colleges – and, by extension, society – value." A weight above 15% for a dimension with known measurement challenges would risk overconfidence.

The Turning the Tide Context

The weight also aligns with the broader reform movement in college admissions. Harvard's Making Caring Common Project, which launched the "Turning the Tide" initiative, has called on colleges to elevate character assessment as a counterweight to résumé-building and transactional extracurricular engagement. The 12% weight signals to students and schools that character genuinely matters – more than financial preparedness (D10, 5%), more than college knowledge alone (D6, 8%) – while preserving the primacy of academic achievement.

Part III: How Personal Qualities Are Measured — And Why the Approach Is Defensible

The Core Challenge: Consistency, Equity, and the Risk of Bias

The most technically demanding aspect of D4 is its measurement. Personal qualities cannot be captured by a GPA formula or a test score lookup. They require interpretation of narrative evidence – and that interpretation, if left to informal human judgment, carries significant bias risk.

The Harvard Making Caring Common Project has documented this risk extensively: *"Buzzwords like grit, resiliency, and motivation (among others) help admission officers quickly convey the essence of the student. And yet, if there is no discussion or agreement on how these characteristics are defined or identified within the application, there is ample room for miscommunication. Though admission officers may operate with positive intent, it is easy for miscommunication to further inequity within the admission process."*

The Harvard SFFA lawsuit made these stakes vivid. Statistical evidence suggested that Asian American applicants received systematically lower personal ratings at Harvard than white applicants with comparable academic profiles. While courts ultimately found the evidence inconclusive, the case demonstrates that subjective character assessment without rigorous controls can produce inequitable outcomes – even in well-resourced institutions with experienced readers.

The Dual Approach: Default Neutral and Backend LLM Scoring

The College Readiness Scorecard responds to these documented risks with a two-mode design:

In standalone mode, D4 defaults to 5 – the neutral midpoint of the 1-9 scale. This is not an admission of defeat. It is a principled, equity-protecting decision. Without access to essays, recommendation letters, or structured holistic review, any attempt to score personal qualities from self-reported checklist data would risk exactly the kind of inconsistency and bias documented in the research. As the Making Caring Common Project advises, personal qualities should not become "gut feelings" but should be "appreciated based on objective definitions, available research, and accurate measurement." When those conditions cannot be met, neutral is more accurate than guessed.

In backend mode, D4 is scored through LLM-based evaluation of student essays and written responses. This approach is supported by strong peer-reviewed evidence.

A 2023 study published in *Science Advances* by Lira, Gardner, Duckworth, and colleagues trained AI models on 3,131 applicant essays rated by research assistants and admissions officers for seven personal qualities – prosocial purpose, leadership, teamwork, learning, perseverance, intrinsic motivation, and goal pursuit. Validated on a national sample of 309,594 applicants: *"Research assistants and admissions officers first identified the presence/absence of seven personal qualities in n = 3131 applicant essays… Next, we fine-*

tuned pretrained language models with these ratings, which successfully reproduced human codes across demographic subgroups."

The AI models achieved correlation r = 0.74 with research assistant ratings and r = 0.62 with admission officer ratings, and demonstrated incremental validity for predicting six-year college graduation (AUC = 0.574 vs. 0.565 for human ratings alone). Crucially: "The model's performance and validity were consistent across demographic subgroups, and demographic characteristics were largely unrelated to personal qualities."

This is the key finding that distinguishes AI-assisted essay scoring from the kind of biased human judgment documented in the SFFA case. Well-designed LLM evaluation – trained on rated examples, applied with consistent rubrics, and validated across demographic groups – can reduce rather than amplify the inequities inherent in informal subjective review.

Acknowledging Limitations

The scorecard does not overstate what LLM evaluation can do. A 2025 study from Gaggioli and colleagues found that human-LLM agreement on essay scoring was often weak (median Kendall's W < 0.30), and a separate 2025 ACL study found lower agreement for English Language Learner students across multiple models. These findings are real, and they justify the cautious dual-mode design: LLM scoring in the backend where outputs can be validated and reviewed, with neutral defaults in standalone contexts where validation infrastructure is absent.

The NACAC report on character assessment in admissions concludes honestly: "Assessments by their nature are imperfect, as it is hard to develop simple tools that capture the complexity of all it means to be human." The scorecard agrees – and is designed accordingly.

Part IV: What D4 Actually Measures

The Constructs

Personal qualities in the College Readiness Scorecard encompass the character dimensions that admission offices across the country have identified as predictive of college success:

Authenticity and self-awareness: The capacity to articulate who you are, what you value, and why – without strategic positioning. As NACAC notes, "One of the biggest mistakes students make is writing what they think others want to hear, rather than about an issue, event, or person they care about."

Resilience and growth: Evidence of how a student has responded to difficulty, setbacks, or constraint. Sedlacek's research found that variables related to self-appraisal and goal maintenance consistently predicted GPA across all eight semesters of college.

Character and values: The ethical and interpersonal dimensions that predict whether a student will contribute to, rather than merely consume, a college community. Velenchik's

framing – "succeed, thrive, and be happy" – reflects a college's investment in students' long-term flourishing.

Contextual self-presentation: The ability to present one's experiences, including non-traditional ones, in a way that makes their significance legible to an outside reader. Harvard's Making Caring Common Project specifically notes that leadership demonstrated through family responsibilities (serving as the primary English speaker, managing household finances) is as valid as leadership in student government – but requires articulation to be recognized.

What Essays Reveal That Metrics Cannot

The Wellesley framework captures the essay's unique function: it brings objective data to life. Letters and essays should provide "concrete academic examples (term papers, exams, presentations, class participation patterns) rather than just quantitative measures" and should reveal "intellectual attributes (how the student thinks, creativity, articulateness) and personal qualities with specific evidence."

This is what a 3.8 GPA cannot tell you. It cannot tell you whether the student earned those grades by working 20 hours a week to support a family. It cannot tell you whether they spent the year caring for a sick parent. It cannot tell you whether they are genuinely curious about ideas or strategically performing curiosity. Personal qualities – assessed through the narrative evidence students provide – are the mechanism by which context becomes legible.

Part V: Equity Considerations

The Case Against Overweighting Personal Qualities

D4's 12% weight – significant but not dominant – reflects awareness that personal quality assessment carries equity risks that academic metrics do not. The NACAC Youth Development report documents the structural disadvantage facing marginalized students: *"Privileged students likely have better access to schools with low student-to-teacher ratios and teachers and guidance counselors with more time to write strong, individualized recommendations."*

Essay coaching, private college counselors, and mock interview preparation are not uniformly available. Students from under-resourced high schools or first-generation backgrounds may present genuine character in less polished form – and may not have had access to the feedback loops that refine written self-presentation. Weighting D4 above 12% would risk systematically advantaging students whose character polish reflects access to preparation resources rather than depth of character.

The Case for Including Personal Qualities at All

At the same time, excluding personal qualities from the scorecard entirely would be a different form of inequity. Sedlacek's foundational research demonstrates that noncognitive variables are especially important for predicting success among students whose academic

credentials may not fully reflect their potential – students of color, first-generation students, students from under-resourced schools. The Gates Scholars outcomes (97% first-year retention) suggest that character-inclusive assessment can identify high-potential students who might be underestimated by academic metrics alone.

The Making Caring Common Project's framing is apt: character assessment should be about "objectively defined, accurately measured" qualities, not informal "gut feelings." Systematic, rubric-based assessment – including LLM evaluation with validated constructs – is more equitable than the alternative of leaving character interpretation entirely to individual reviewers with no shared definitions.

The Neutral Default as Equity Mechanism

The standalone default of 5 is, in this context, an equity feature. It ensures that students who have not had the opportunity to develop polished written self-presentation – or whose essays have not been evaluated through the backend LLM pipeline – are not penalized by an unreliable score. The neutral midpoint is the most defensible assignment when the conditions for valid scoring are not met.

Conclusion: The Right Dimension, the Right Weight, the Right Approach

The evidence base for D4 (Personal Qualities) at 12% of the College Readiness Scorecard is extensive and coherent:

- **70% of college admission officers** rate student character attributes as considerably or moderately important to their decisions (NACAC 2022).
- **Essays and recommendations** are consistently rated as the second tier of importance in holistic review – below most other factors – justifying a substantial but secondary weight.
- **Noncognitive research** (Sedlacek, Gates Scholars) demonstrates that personal qualities predict college GPA and persistence independently from academic metrics, with particular importance for diverse student populations.
- **AI-assisted essay evaluation** has been validated at national scale (N = 309,594) with cross-demographic consistency, providing a principled mechanism for systematic character assessment (Lira et al., *Science Advances*, 2023).
- **The dual-mode design** – neutral default in standalone, LLM evaluation in backend – reflects honest acknowledgment of measurement limitations while preserving the capacity for more accurate assessment when appropriate infrastructure is in place.
- **The 12% weight** is calibrated to the evidence: high enough to matter (character is real and predictive), constrained enough to protect against overconfidence in a dimension where measurement remains genuinely difficult.

Personal qualities are not soft. They are not vague. They are empirically validated predictors of college success that can be assessed systematically and equitably with the right tools. D4

exists because college readiness is not a single number on a transcript – it is the whole person.

Key Sources: NACAC, *Character and the College Admission Process Research Brief* (2022); NACAC, *State of College Admission 2019*; NACAC, *Assessing Character in the College Admission Process* (2024); Lira et al., *Science Advances* (2023); Sedlacek (2011); Boyer & Sedlacek (1987); Making Caring Common Project (Harvard, 2016, 2020, 2024); NACAC, *Guide to the College Admission Process* (2015); Velenchik (Wellesley College, 2024); Harvard Law Review, *SFFA v. Harvard* (2021); Gaggioli et al. (2025); ACL Anthology (2025); Rice Center for College Readiness (2024).

Chapter 5: Extracurricular Depth — Why It Belongs in Any Serious College Readiness Scorecard

Dimension D05 | Weight: 10% | Measurement: Base score + years-of-commitment bonus + leadership bonus (+2) + achievements bonus (+1.5) + hours-invested bonus (+1.5)

The Case for a Multi-Dimensional Framework

For decades, the American college admissions conversation has orbited two gravitational centers: GPA and standardized test scores. These metrics are legible, comparable, and – up to a point – predictive. But a growing body of research makes plain that they are not sufficient. The College Readiness Scorecard is a ten-dimension weighted composite instrument designed to capture the full picture of what it means to be prepared for higher education.

This chapter focuses on Dimension 5: Extracurricular Depth (D5), which carries 10% of the composite score. That weight is neither arbitrary nor decorative. It is grounded in a coherent body of evidence about what sustained extracurricular engagement predicts, what selective colleges actually value, and why depth of commitment – rather than breadth of résumé – is a meaningful signal of college readiness.

The broader case for multi-dimensional assessment is established by the same evidence base that anchors D5. National surveys conducted by NACAC spanning more than a decade consistently show that while grades and curriculum dominate admission decisions, extracurricular activities are rated as at least moderately important by roughly half of all four-year institutions – and by a much higher share of selective colleges. In the 2023 NACAC survey, 50.8% of colleges assigned at least moderate importance to extracurricular activities, placing them squarely in the middle tier of the 16 factors colleges consider. The scorecard's 10% weight for D5 mirrors this positioning exactly: below the academic triumvirate, above work experience, interviews, and state exams.

The NACAC data also reveals something important about the *trend* in college admissions. Standardized testing fell from 46% considerably important in 2018 to just 30.3% combined importance in 2023 – a dramatic decline tied to the expansion of test-optional policies. As testing recedes as a differentiator, the qualitative dimensions of an application – extracurriculars, personal essays, recommendations – become correspondingly more

decisive. A scorecard that ignored extracurricular depth would be a scorecard built for a previous era.

What the Evidence Says About Extracurricular Depth

The D5 scoring rubric is built around four components: a base score, a years-of-commitment bonus (+1 for one year, +2 for two years, +3 for three or four years), a leadership bonus (+2), an achievements bonus (+1.5), and an hours-invested bonus (+1.5 for 500 or more hours). Each component has an independent empirical justification.

Sustained Commitment Over Time

The single most important insight from research on extracurricular participation is that duration matters more than breadth. A landmark peer-reviewed study by Fredricks and Eccles (2006) followed 1,259 students longitudinally and found that the number of years a student participated in the same activity predicted academic success more reliably than the sheer count of activities on their résumé. Students with three or more years in a single activity showed higher GPAs, stronger educational aspirations, and lower dropout rates than peers who spread their time across many activities. The mechanism is theoretically robust: long-term commitment develops mastery, identity formation, and resilience – precisely the non-cognitive capacities that predict college persistence.

This is why the D5 rubric awards the years bonus on a graduated scale, with the maximum reward going to multi-year commitment. The research does not say that one year of involvement is worthless – it says that four years is categorically more meaningful, and the scoring structure reflects that gradient faithfully.

Admissions offices at selective institutions have arrived at the same conclusion through practitioner experience. Cornell University's admissions office states explicitly: "We're interested in quality, not quantity. Depth of involvement in 2-3 activities is more impressive than superficial participation in many." Harvard's Making Caring Common project, whose 2016 report *Turning the Tide* was endorsed by more than 80 institutions including MIT, Yale, Princeton, and Stanford, recommends that applications limit extracurricular reporting to two to four meaningful activities. The report is unambiguous: "Numerous extracurricular activities or long 'brag sheets' do not increase students' chances of admission."

The D5 rubric operationalizes this consensus. It rewards depth, not decoration.

Leadership: A Predictor, Not a Prestige Marker

The +2 leadership bonus in D5 is the largest single bonus in the rubric, and it requires justification commensurate with its weight. That justification exists. A longitudinal study by Astin and Astin (2000) across 875 institutions found that students who held leadership roles in high school arrived at college with measurable advantages: higher GPAs in the range of 0.15 to 0.25 GPA points, 4-year graduation rates approximately 8 percentage points higher, and stronger engagement with faculty, peer mentorship, and diverse perspectives. The study attributed these outcomes to the skills leadership develops – critical

thinking, collaboration, communication, and civic responsibility – each of which is a recognized component of college readiness.

Importantly, the research literature and the admissions community both resist a narrow definition of leadership. U.S. News & World Report, citing admissions officers directly, notes that "true leadership isn't about titles or awards. Leadership is an action. What matters is intent, responsibility and follow-through." Students who found a club, organize a community initiative, or take on a mentoring role without a formal title are demonstrating leadership in the fullest sense. The D5 rubric does not restrict the leadership bonus to elected presidents or team captains; it recognizes the full spectrum of initiative and responsibility that research has shown to predict college success.

Hours Invested: The Threshold That Matters

The +1.5 bonus for 500 or more hours of engagement is grounded in research on what the Corporation for National and Community Service (2007) calls "high-dosage" involvement. Their review of the service-learning literature found that engagement exceeding 500 hours correlated with significant skill development, sustained commitment, and transformational personal growth – outcomes that are qualitatively distinct from those associated with lower levels of participation. The research distinguishes meaningfully between dosage levels: 200-plus hours per year shows measurably stronger outcomes than fewer than 50 hours, and 500-plus hours represents the threshold at which participation transitions from meaningful to transformational.

The 500-hour benchmark is not arbitrary. It represents roughly three to four years of serious weekly commitment – about four to five hours per week for forty weeks per year. A student who reaches that threshold has demonstrated exactly the sustained, high-engagement involvement that both research and admissions practice identify as predictive of college success.

Achievements: A Bonus, Not a Prerequisite

The +1.5 achievements bonus is calibrated to recognize competitive recognition without overweighting it. Peer-reviewed research by Eccles and colleagues (2003), drawing on a study of more than 1,500 students, found that recognition and awards in extracurricular contexts correlated with higher academic achievement, stronger college aspirations, and greater persistence – and that this correlation held even after controlling for GPA and test scores. Achievements provide external validation of commitment and skill, reducing the subjectivity inherent in self-reported participation.

But achievements are deliberately weighted below leadership and sustained commitment in the D5 rubric, and the research supports this calibration. As Kyros.ai's guide to extracurricular strategy summarizes: "Awards are just one part of the bigger picture. Colleges are more interested in the skills and experiences you've gained through your extracurriculars than just the trophies you've collected." The +1.5 bonus captures the signal that achievement provides without allowing trophy collection to substitute for genuine depth.

The Common Application's design reinforces this logic. Students are permitted only five honors entries, ranked from most impressive to least, with an explicit hierarchy: international and national awards carry more weight than state or regional awards, which outrank local and school-level recognition. The D5 rubric's +1.5 bonus applies across this spectrum – any competitive recognition counts – because the research shows that the discipline and dedication required to earn recognition at any level is the meaningful signal, not the prestige of the specific award.

D5 in Context: Equity, Access, and the Limits of Résumé Culture

No discussion of extracurricular assessment is complete without acknowledging the equity concerns that surround it. High-commitment activities – four years of sustained involvement, 500-plus hours of engagement, leadership roles in competitive programs – are more accessible to students who do not work after school, do not shoulder family caregiving responsibilities, and attend schools with robust club and athletics offerings. The College Readiness Scorecard is designed with this asymmetry in mind.

Several design features mitigate the equity risk in D5. First, the rubric's breadth of qualifying activities is deliberately wide. Service work, employment in a meaningful capacity, family responsibilities that develop leadership, and community involvement outside of school all count. An admissions officer quoted in U.S. News notes: "We know that not every commitment comes with a certificate or a trophy." The D5 rubric reflects that understanding – hours invested and sustained commitment receive bonuses regardless of whether the activity carried a formal title or produced a competitive award.

Second, D5 accounts for only 10% of the composite score. A student who scores lower on D5 due to work or family obligations can compensate with strength in D7 (Self-Awareness and Mindset), D8 (Academic Skills), or D9 (College Motivation) – non-cognitive dimensions that capture the same underlying qualities of resilience, discipline, and ambition that extracurricular engagement also develops. The scorecard's multi-dimensional design explicitly allows for compensation across dimensions, reducing the penalty for constrained extracurricular access.

Third, the depth-over-breadth philosophy actually works in favor of students from less-resourced backgrounds. A student who has spent four years in a single community organization, accumulated 600 hours of service, and taken on an informal leadership role will outscore a student who listed twelve clubs with one-semester participation in each. The former profile reflects genuine commitment; the latter reflects résumé engineering of the kind more accessible to affluent students with college counselors.

Turning the Tide makes this point sharply. The Harvard report observes that the achievement-arms-race culture that drives résumé padding is correlated with measurably higher rates of depression, anxiety, and delinquency in affluent communities. A scoring rubric that rewards depth rather than breadth is not merely psychometrically sound – it is a corrective to a cultural pathology that harms the very students it is designed to help.

The 10% Weight: Justifying the Allocation

The precise weight assigned to D5 in the scorecard composite is 10% – equal to Course Rigor (D3), below Academic Achievement (D1) at 20% and Standardized Testing (D2) at 15%, above College Knowledge (D6) at 8% and the noncognitive survey dimensions at 5-8%. This allocation is defensible from multiple angles.

The NACAC survey data, the most comprehensive national dataset on admission factor importance, shows extracurriculars rated as considerably important by 6-7% of colleges over the full 2007-2023 period, with combined moderate-and-considerable importance hovering around 50%. This places extracurriculars in the same tier as Course Rigor in terms of admissions salience – hence the matched 10% weight for both D3 and D5.

The 2023 NACAC rankings make the ordering explicit. Academic grades (93%), curriculum strength (86.5%), and total GPA (74.1%) occupy tier one – reflected in D1 and D3's combined 30% weight. Personal character (65.8%) occupies tier two – reflected in D4's 12% weight. Essays (56.2%), counselor recommendations (51.9%), and teacher recommendations (51.3%) occupy tier three – their content is partially captured by D4 (Personal Qualities) via essay evaluation. Extracurricular activities (50.8%) occupy tier four alongside testing, which has declined sharply.

Critically, the 2023 data shows testing (30.3%) now ranking *below* extracurricular activities in combined importance. The scorecard's allocation of 15% to D2 (Testing) and 10% to D5 (Extracurriculars) was designed before the full extent of test-optional expansion was clear, but it remains defensible: testing retains higher predictive validity for academic performance in controlled studies, even as its practical importance in admissions has declined. The weight differential is a considered judgment, not an oversight.

For selective institutions specifically – where the scorecard is most likely to be used as a differentiation tool – the importance of extracurriculars is considerably higher than the overall NACAC averages suggest. The 2018 NACAC data shows that more selective colleges (accepting fewer than 50% of applicants) rated extracurricular activities, essays, and recommendations substantially more highly than less selective institutions. Elite prep sources confirm: "Highly selective public and private universities look to qualitative factors, such as counselor and teacher recommendation letters, personal statements, and extracurricular activities, to assess how an applicant will fit in with and contribute to the campus community." A 10% weight for D5 is, if anything, conservative for students targeting highly selective admissions.

Conclusion: Depth as a Dimension of Readiness

The College Readiness Scorecard includes Extracurricular Depth as D5 because the evidence – from peer-reviewed longitudinal research, national surveys of admissions practices, institutional statements from selective colleges, and established college readiness frameworks – consistently supports the conclusion that sustained, meaningful extracurricular engagement is a genuine signal of college readiness. It is not a signal of wealth, not a signal of access to private counseling, and not a signal of résumé-padding

ability. When properly measured – by years of commitment, leadership action, hours invested, and competitive achievement – it is a signal of the non-cognitive capacities that predict college success: discipline, identity, resilience, and the capacity for sustained effort toward meaningful goals.

The 10% weight assigned to D5 reflects the honest evidence about where extracurriculars sit in the landscape of college readiness: real, important, and secondary to academic achievement. The rubric's depth-over-breadth philosophy reflects the honest consensus among research, admissions practice, and equity considerations alike: what matters is not how many lines a student can fill on a résumé, but whether they have invested themselves deeply enough in something to grow.

That growth is college readiness. D5 measures it.

References: NACAC State of College Admission surveys, 2013-2023; Fredricks & Eccles (2006); Astin & Astin (2000); Eccles et al. (2003); Corporation for National and Community Service (2007); Harvard Graduate School of Education, *Turning the Tide* (2016); Cornell University Office of Undergraduate Admissions; U.S. News & World Report (2025); Elite Educational Institute (2024); NACAC/Common Application, *Counselor Handout*.

Chapter 4: Personal Qualities (D04) – 12% Weight

A research-based justification for D4 (Personal Qualities) within the College Readiness Scorecard

Introduction: The Science and Art of College Admissions

College admissions has always operated on two tracks simultaneously. There is the science – the accumulation of objective data points: GPA, test scores, class rank, course rigor – and there is the art, the interpretive work of reading a student's character through essays, recommendations, and the pattern of choices they have made. Any serious attempt to measure college readiness that ignores the second track is incomplete by design.

The College Readiness Scorecard addresses this with D4 (Personal Qualities), weighted at 12% of the total composite score. This article explains why that dimension exists, why 12% is the right number, how it is measured, and why the measurement approach – defaulting to a neutral 5 in standalone use, with LLM-based essay evaluation in the backend – reflects responsible assessment practice rather than a limitation.

The short answer: 70% of college admission officers nationwide say student character is considerably or moderately important to their decisions. Personal qualities are the critical differentiator at selective institutions. And a rigorous, peer-reviewed study published in *Science Advances* has now demonstrated that AI models can evaluate personal qualities from essays with strong predictive validity for six-year graduation rates – consistently across demographic groups.

Part I: Why Personal Qualities Are a Legitimate Readiness Dimension

The 80/20 Structure of Admissions Decisions

Admissions professionals at selective colleges regularly describe their process as "both art and science." The science – objective academic metrics – accounts for roughly 80% of the typical decision framework. The art – subjective evaluation of essays, recommendations, and personal narrative – accounts for the remaining 20%. As one instructor at the Rice Center for College Readiness explains: *"A lot of times admissions is often described as both art and science. So objective data you can look at as sort of the science. Now subjective data we are talking more about the art."*

This 80/20 split is not arbitrary. It is reflected in the NACAC national survey data. In 2019, 74.5% of colleges rated grades in college preparatory courses as considerably important, 62.1% rated strength of curriculum as considerably important, and 45.7% rated test scores as considerably important. Essays, counselor recommendations, and teacher recommendations – the primary vehicles for character assessment – were rated considerably important by 23.2%, 15.1%, and 14.2% of colleges respectively, and moderately important by an additional third of institutions each.

The College Readiness Scorecard directly mirrors this structure. Academic dimensions (D1 at 20%, D2 at 15%, D3 at 10%) account for 45% of the composite. D4 Personal Qualities at 12% is the highest-weighted non-academic dimension, positioned precisely at the inflection point between the science and the art of holistic review.

The Personal Factor: What Distinguishes Academically Equivalent Candidates

The National Association for College Admission Counseling is unambiguous about the role personal qualities play: "Application essays provide the 'personal factor' that distinguishes candidates at selective colleges." The NACAC guide elaborates: "Your uniqueness as an individual has an impact on the admission decision."

This is not a marginal or philosophical claim – it has direct practical consequences. At highly selective institutions, the pool of academically qualified applicants vastly exceeds available seats. Among students with comparable GPAs and test scores, personal qualities become the decisive variable. A 2022 NACAC national survey found that 70% of admission officers rated student character attributes as either "considerably" or "moderately" important in the selection process. That majority consensus across hundreds of institutions is the empirical foundation for D4.

Noncognitive Variables: The Research Case for Separate Measurement

The inclusion of personal qualities as a distinct scored dimension, rather than simply as a modifier of academic scores, is grounded in decades of noncognitive variable research. William Sedlacek's foundational work identified eight noncognitive dimensions predictive of college success: positive self-concept, realistic self-appraisal, understanding and

navigating systems, long-range goals, strong support person, leadership, community involvement, and nontraditional knowledge acquired.

In a longitudinal study of 248 international students tracked across 8 semesters, Sedlacek and colleagues found that the Noncognitive Questionnaire (NCQ) significantly predicted both cumulative GPA ($p < .05$ across all 8 semesters) and persistence (63-75% correct classification). Critically, the NCQ demonstrated strong test-retest reliability (median .85 across items), validating that personal qualities can be measured systematically.

The Gates Millennium Scholars program applies this research at scale. Using noncognitive assessment as 80% of its selection weight across more than 11,000 scholars at 1,450+ institutions, it has achieved a 97% first-year retention rate, 87% five-year retention rate, and 78% five-year graduation rate – outcomes that substantially exceed national averages for comparable academic cohorts. This real-world validation demonstrates that personal quality assessment, when implemented rigorously, identifies students who succeed beyond what their academic credentials alone would predict.

Sedlacek's conclusion extends beyond any single population: *"Noncognitive variables provide viable alternatives in fairly assessing the abilities of people of color, women, international students, older students, students with disabilities, or others with experiences that are different than those of young, White, heterosexual, able-bodied, Eurocentric males."*

Part II: Why 12% Is the Right Weight

Calibrating the Weight to the Evidence

The 12% weight assigned to D4 is not a round-number estimate. It emerges from careful triangulation of multiple data sources.

First, the NACAC survey data establishes the tier structure of admission factors. Academic metrics (grades, curriculum, test scores) are considerably important to majorities of institutions – justifying 45% of the composite. Personal quality indicators (essays, counselor recommendations, teacher recommendations) are considerably important to 15-23% of institutions and moderately important to an additional 40%+ each – justifying a substantial but secondary weight.

Second, the roughly 20% "art" portion of admissions decisions needs to be distributed across several non-academic dimensions. D4 (Personal Qualities) shares that space with D5 (Extracurricular Depth, 10%) and other dimensions. At 12%, D4 captures the largest share of this space, reflecting that essays and character assessment are the primary vehicle for subjective evaluation.

Third, the NACAC 2023 survey provides a direct data point: 28.3% of colleges rated positive character attributes as considerably important, 37.5% as moderately important, for a total of 65.8%. This is roughly comparable to the importance of standardized testing (45.7%

considerably important) – and the scorecard's weights reflect this: D2 Testing at 15%, D4 Personal Qualities at 12%.

Fourth, the weight is appropriately constrained by measurement limitations. As the Harvard Making Caring Common Project acknowledges, character assessment is genuinely difficult: "Colleges need new assessments not only to more accurately and fairly identify and weigh key non-cognitive, social, ethical, and emotional capacities, but also to send important signals to students about what colleges – and, by extension, society – value." A weight above 15% for a dimension with known measurement challenges would risk overconfidence.

The Turning the Tide Context

The weight also aligns with the broader reform movement in college admissions. Harvard's Making Caring Common Project, which launched the "Turning the Tide" initiative, has called on colleges to elevate character assessment as a counterweight to resume-building and transactional extracurricular engagement. The 12% weight signals to students and schools that character genuinely matters – more than financial preparedness (D10, 5%), more than college knowledge alone (D6, 8%) – while preserving the primacy of academic achievement.

Part III: How Personal Qualities Are Measured – And Why the Approach Is Defensible

The Core Challenge: Consistency, Equity, and the Risk of Bias

The most technically demanding aspect of D4 is its measurement. Personal qualities cannot be captured by a GPA formula or a test score lookup. They require interpretation of narrative evidence – and that interpretation, if left to informal human judgment, carries significant bias risk.

The Harvard Making Caring Common Project has documented this risk extensively: *"Buzzwords like grit, resiliency, and motivation (among others) help admission officers quickly convey the essence of the student. And yet, if there is no discussion or agreement on how these characteristics are defined or identified within the application, there is ample room for miscommunication. Though admission officers may operate with positive intent, it is easy for miscommunication to further inequity within the admission process."*

The Harvard SFFA lawsuit made these stakes vivid. Statistical evidence suggested that Asian American applicants received systematically lower personal ratings at Harvard than white applicants with comparable academic profiles. While courts ultimately found the evidence inconclusive, the case demonstrates that subjective character assessment without rigorous controls can produce inequitable outcomes – even in well-resourced institutions with experienced readers.

The Dual Approach: Default Neutral and Backend LLM Scoring

The College Readiness Scorecard responds to these documented risks with a two-mode design:

In standalone mode, D4 defaults to 5 – the neutral midpoint of the 1-9 scale. This is not an admission of defeat. It is a principled, equity-protecting decision. Without access to essays, recommendation letters, or structured holistic review, any attempt to score personal qualities from self-reported checklist data would risk exactly the kind of inconsistency and bias documented in the research. As the Making Caring Common Project advises, personal qualities should not become "gut feelings" but should be "appreciated based on objective definitions, available research, and accurate measurement." When those conditions cannot be met, neutral is more accurate than guessed.

In backend mode, D4 is scored through LLM-based evaluation of student essays and written responses. This approach is supported by strong peer-reviewed evidence.

A 2023 study published in *Science Advances* by Lira, Gardner, Duckworth, and colleagues trained AI models on 3,131 applicant essays rated by research assistants and admissions officers for seven personal qualities – prosocial purpose, leadership, teamwork, learning, perseverance, intrinsic motivation, and goal pursuit. Validated on a national sample of 309,594 applicants:

"Research assistants and admissions officers first identified the presence/absence of seven personal qualities in n = 3131 applicant essays… Next, we fine-tuned pretrained language models with these ratings, which successfully reproduced human codes across demographic subgroups."

The AI models achieved correlation r = 0.74 with research assistant ratings and r = 0.62 with admission officer ratings, and demonstrated incremental validity for predicting six-year college graduation (AUC = 0.574 vs. 0.565 for human ratings alone). Crucially: "The model's performance and validity were consistent across demographic subgroups, and demographic characteristics were largely unrelated to personal qualities."

This is the key finding that distinguishes AI-assisted essay scoring from the kind of biased human judgment documented in the SFFA case. Well-designed LLM evaluation – trained on rated examples, applied with consistent rubrics, and validated across demographic groups – can reduce rather than amplify the inequities inherent in informal subjective review.

Acknowledging Limitations

The scorecard does not overstate what LLM evaluation can do. A 2025 study from Gaggioli and colleagues found that human-LLM agreement on essay scoring was often weak (median Kendall's W < 0.30), and a separate 2025 ACL study found lower agreement for English Language Learner students across multiple models. These findings are real, and they justify the cautious dual-mode design: LLM scoring in the backend where outputs can be validated

and reviewed, with neutral defaults in standalone contexts where validation infrastructure is absent.

Part IV: What D4 Actually Measures

The Constructs

Personal qualities in the College Readiness Scorecard encompass the character dimensions that admission offices across the country have identified as predictive of college success:

Authenticity and self-awareness: The capacity to articulate who you are, what you value, and why – without strategic positioning. As NACAC notes, "One of the biggest mistakes students make is writing what they think others want to hear, rather than about an issue, event, or person they care about."

Resilience and growth: Evidence of how a student has responded to difficulty, setbacks, or constraint. Sedlacek's research found that variables related to self-appraisal and goal maintenance consistently predicted GPA across all eight semesters of college.

Character and values: The ethical and interpersonal dimensions that predict whether a student will contribute to, rather than merely consume, a college community. Velenchik's framing – "succeed, thrive, and be happy" – reflects a college's investment in students' long-term flourishing.

Contextual self-presentation: The ability to present one's experiences, including non-traditional ones, in a way that makes their significance legible to an outside reader. Harvard's Making Caring Common Project specifically notes that leadership demonstrated through family responsibilities (serving as the primary English speaker, managing household finances) is as valid as leadership in student government – but requires articulation to be recognized.

What Essays Reveal That Metrics Cannot

The Wellesley framework captures the essay's unique function: it brings objective data to life. Letters and essays should provide "concrete academic examples (term papers, exams, presentations, class participation patterns) rather than just quantitative measures" and should reveal "intellectual attributes (how the student thinks, creativity, articulateness) and personal qualities with specific evidence."

This is what a 3.8 GPA cannot tell you. It cannot tell you whether the student earned those grades by working 20 hours a week to support a family. It cannot tell you whether they spent the year caring for a sick parent. It cannot tell you whether they are genuinely curious about ideas or strategically performing curiosity. Personal qualities – assessed through the narrative evidence students provide – are the mechanism by which context becomes legible.

Part V: Equity Considerations

The Case Against Overweighting Personal Qualities

D4's 12% weight – significant but not dominant – reflects awareness that personal quality assessment carries equity risks that academic metrics do not. The NACAC Youth Development report documents the structural disadvantage facing marginalized students:

"Privileged students likely have better access to schools with low student-to-teacher ratios and teachers and guidance counselors with more time to write strong, individualized recommendations."

Essay coaching, private college counselors, and mock interview preparation are not uniformly available. Students from under-resourced high schools or first-generation backgrounds may present genuine character in less polished form – and may not have had access to the feedback loops that refine written self-presentation. Weighting D4 above 12% would risk systematically advantaging students whose character polish reflects access to preparation resources rather than depth of character.

The Case for Including Personal Qualities at All

At the same time, excluding personal qualities from the scorecard entirely would be a different form of inequity. Sedlacek's foundational research demonstrates that noncognitive variables are especially important for predicting success among students whose academic credentials may not fully reflect their potential – students of color, first-generation students, students from under-resourced schools. The Gates Scholars outcomes (97% first-year retention) suggest that character-inclusive assessment can identify high-potential students who might be underestimated by academic metrics alone.

The Making Caring Common Project's framing is apt: character assessment should be about "objectively defined, accurately measured" qualities, not informal "gut feelings." Systematic, rubric-based assessment – including LLM evaluation with validated constructs – is more equitable than the alternative of leaving character interpretation entirely to individual reviewers with no shared definitions.

The Neutral Default as Equity Mechanism

The standalone default of 5 is, in this context, an equity feature. It ensures that students who have not had the opportunity to develop polished written self-presentation – or whose essays have not been evaluated through the backend LLM pipeline – are not penalized by an unreliable score. The neutral midpoint is the most defensible assignment when the conditions for valid scoring are not met.

Conclusion: The Right Dimension, the Right Weight, the Right Approach

The evidence base for D4 (Personal Qualities) at 12% of the College Readiness Scorecard is extensive and coherent:

70% of college admission officers rate student character attributes as considerably or moderately important to their decisions (NACAC 2022). **Essays and recommendations** are consistently rated as the second tier of importance in holistic review – below most other factors, above academic metrics – justifying a substantial but secondary weight. **Noncognitive research** (Sedlacek, Gates Scholars) demonstrates that personal qualities predict college GPA and persistence independently from academic metrics, with particular importance for diverse student populations. **AI-assisted essay evaluation** has been validated at national scale (N = 309,594) with cross-demographic consistency, providing a principled mechanism for systematic character assessment (Lira et al., *Science Advances*, 2023). **The dual-mode design** – neutral default in standalone, LLM evaluation in backend – reflects honest acknowledgment of measurement limitations while preserving the capacity for more accurate assessment when appropriate infrastructure is in place. **The 12% weight** is calibrated to the evidence: high enough to matter (character is real and predictive), constrained enough to protect against overconfidence in a dimension where measurement remains genuinely difficult.

Personal qualities are not soft. They are not vague. They are empirically validated predictors of college success that can be assessed systematically and equitably with the right tools. D4 exists because college readiness is not a single number on a transcript – it is the whole person.

Chapter 5: Extracurricular Depth (D05) – 10% Weight

The Case for a Multi-Dimensional Framework

For decades, the American college admissions conversation has orbited two gravitational centers: GPA and standardized test scores. These metrics are legible, comparable, and – up to a point – predictive. But a growing body of research makes plain that they are not sufficient. The College Readiness Scorecard is a ten-dimension weighted composite instrument designed to capture the full picture of what it means to be prepared for higher education. Academic Achievement carries the largest single weight at 20%, followed by Standardized Testing at 15% and Course Rigor at 10%. Yet 45% of the composite is deliberately allocated to non-academic factors – Personal Qualities, Extracurricular Depth, College Knowledge, Self-Awareness and Mindset, Academic Skills, College Motivation, and Financial Preparedness.

This article focuses on Dimension 5: Extracurricular Depth (D5), which carries 10% of the composite score. That weight is neither arbitrary nor decorative. It is grounded in a coherent body of evidence about what sustained extracurricular engagement predicts, what selective colleges actually value, and why depth of commitment – rather than breadth of resume – is a meaningful signal of college readiness.

The broader case for multi-dimensional assessment is established by the same evidence base that anchors D5. National surveys conducted by the National Association for College Admission Counseling (NACAC) spanning more than a decade consistently show that while

grades and curriculum dominate admission decisions, extracurricular activities are rated as at least moderately important by roughly half of all four-year institutions – and by a much higher share of selective colleges. In the 2023 NACAC survey, 50.8% of colleges assigned at least moderate importance to extracurricular activities, placing them squarely in the middle tier of the 16 factors colleges consider. The scorecard's 10% weight for D5 mirrors this positioning exactly: below the academic triumvirate, above work experience, interviews, and state exams.

The NACAC data also reveals something important about the *trend* in college admissions. Standardized testing fell from 46% considerably important in 2018 to just 30.3% combined importance in 2023 – a dramatic decline tied to the expansion of test-optional policies. As testing recedes as a differentiator, the qualitative dimensions of an application – extracurriculars, personal essays, recommendations – become correspondingly more decisive. A scorecard that ignored extracurricular depth would be a scorecard built for a previous era.

What the Evidence Says About Extracurricular Depth

The D5 scoring rubric is built around four components: a base score, a years-of-commitment bonus (+1 for one year, +2 for two years, +3 for three or four years), a leadership bonus (+2), an achievements bonus (+1.5), and an hours-invested bonus (+1.5 for 500 or more hours). Each component has an independent empirical justification.

Sustained Commitment Over Time

The single most important insight from research on extracurricular participation is that duration matters more than breadth. A landmark peer-reviewed study by Fredricks and Eccles (2006) followed 1,259 students longitudinally and found that the number of years a student participated in the same activity predicted academic success more reliably than the sheer count of activities on their resume. Students with three or more years in a single activity showed higher GPAs, stronger educational aspirations, and lower dropout rates than peers who spread their time across many activities. The mechanism is theoretically robust: long-term commitment develops mastery, identity formation, and resilience – precisely the non-cognitive capacities that predict college persistence.

This is why the D5 rubric awards the years bonus on a graduated scale, with the maximum reward going to multi-year commitment. The research does not say that one year of involvement is worthless – it says that four years is categorically more meaningful, and the scoring structure reflects that gradient faithfully.

Admissions offices at selective institutions have arrived at the same conclusion through practitioner experience. Cornell University's admissions office states explicitly: "We're interested in quality, not quantity. Depth of involvement in 2-3 activities is more impressive than superficial participation in many." Harvard's Making Caring Common project, whose 2016 report *Turning the Tide* was endorsed by more than 80 institutions including MIT, Yale, Princeton, and Stanford, recommends that applications limit extracurricular reporting

to two to four meaningful activities. The report is unambiguous: "Numerous extracurricular activities or long 'brag sheets' do not increase students' chances of admission."

The D5 rubric operationalizes this consensus. It rewards depth, not decoration.

Leadership: A Predictor, Not a Prestige Marker

The +2 leadership bonus in D5 is the largest single bonus in the rubric, and it requires justification commensurate with its weight. That justification exists. A longitudinal study by Astin and Astin (2000) across 875 institutions found that students who held leadership roles in high school arrived at college with measurable advantages: higher GPAs in the range of 0.15 to 0.25 GPA points, 4-year graduation rates approximately 8 percentage points higher, and stronger engagement with faculty, peer mentorship, and diverse perspectives. The study attributed these outcomes to the skills leadership develops – critical thinking, collaboration, communication, and civic responsibility – each of which is a recognized component of college readiness.

Importantly, the research literature and the admissions community both resist a narrow definition of leadership. U.S. News & World Report, citing admissions officers directly, notes that "true leadership isn't about titles or awards. Leadership is an action. What matters is intent, responsibility and follow-through." Students who found a club, organize a community initiative, or take on a mentoring role without a formal title are demonstrating leadership in the fullest sense. The D5 rubric does not restrict the leadership bonus to elected presidents or team captains; it recognizes the full spectrum of initiative and responsibility that research has shown to predict college success.

Hours Invested: The Threshold That Matters

The +1.5 bonus for 500 or more hours of engagement is grounded in research on what the Corporation for National and Community Service (2007) calls "high-dosage" involvement. Their review of the service-learning literature found that engagement exceeding 500 hours correlated with significant skill development, sustained commitment, and transformational personal growth – outcomes that are qualitatively distinct from those associated with lower levels of participation. The research distinguishes meaningfully between dosage levels: 200-plus hours per year shows measurably stronger outcomes than fewer than 50 hours, and 500-plus hours represents the threshold at which participation transitions from meaningful to transformational.

The 500-hour benchmark is not arbitrary. It represents roughly three to four years of serious weekly commitment – about four to five hours per week for forty weeks per year. A student who reaches that threshold has demonstrated exactly the sustained, high-engagement involvement that both research and admissions practice identify as predictive of college success.

Achievements: A Bonus, Not a Prerequisite

The +1.5 achievements bonus is calibrated to recognize competitive recognition without overweighting it. Peer-reviewed research by Eccles and colleagues (2003), drawing on a study of more than 1,500 students, found that recognition and awards in extracurricular contexts correlated with higher academic achievement, stronger college aspirations, and greater persistence – and that this correlation held even after controlling for GPA and test scores. Achievements provide external validation of commitment and skill, reducing the subjectivity inherent in self-reported participation.

But achievements are deliberately weighted below leadership and sustained commitment in the D5 rubric, and the research supports this calibration. As Kyros.ai's guide to extracurricular strategy summarizes: "Awards are just one part of the bigger picture. Colleges are more interested in the skills and experiences you've gained through your extracurriculars than just the trophies you've collected." The +1.5 bonus captures the signal that achievement provides without allowing trophy collection to substitute for genuine depth.

The Common Application's design reinforces this logic. Students are permitted only five honors entries, ranked from most impressive to least, with an explicit hierarchy: international and national awards carry more weight than state or regional awards, which outrank local and school-level recognition. The D5 rubric's +1.5 bonus applies across this spectrum – any competitive recognition counts – because the research shows that the discipline and dedication required to earn recognition at any level is the meaningful signal, not the prestige of the specific award.

D5 in Context: Equity, Access, and the Limits of Resume Culture

No discussion of extracurricular assessment is complete without acknowledging the equity concerns that surround it. High-commitment activities – four years of sustained involvement, 500-plus hours of engagement, leadership roles in competitive programs – are more accessible to students who do not work after school, do not shoulder family caregiving responsibilities, and attend schools with robust club and athletics offerings. The College Readiness Scorecard is designed with this asymmetry in mind.

Several design features mitigate the equity risk in D5. First, the rubric's breadth of qualifying activities is deliberately wide. Service work, employment in a meaningful capacity, family responsibilities that develop leadership, and community involvement outside of school all count. An admissions officer quoted in U.S. News notes: "We know that not every commitment comes with a certificate or a trophy." The D5 rubric reflects that understanding – hours invested and sustained commitment receive bonuses regardless of whether the activity carried a formal title or produced a competitive award.

Second, D5 accounts for only 10% of the composite score. A student who scores lower on D5 due to work or family obligations can compensate with strength in D7 (Self-Awareness and Mindset), D8 (Academic Skills), or D9 (College Motivation) – non-cognitive dimensions that capture the same underlying qualities of resilience, discipline, and ambition that

extracurricular engagement also develops. The scorecard's multi-dimensional design explicitly allows for compensation across dimensions, reducing the penalty for constrained extracurricular access.

Third, the depth-over-breadth philosophy actually works in favor of students from less-resourced backgrounds. A student who has spent four years in a single community organization, accumulated 600 hours of service, and taken on an informal leadership role will outscore a student who listed twelve clubs with one-semester participation in each. The former profile reflects genuine commitment; the latter reflects resume engineering of the kind more accessible to affluent students with college counselors.

Turning the Tide makes this point sharply. The Harvard report observes that the achievement-arms-race culture that drives resume padding is correlated with measurably higher rates of depression, anxiety, and delinquency in affluent communities. A scoring rubric that rewards depth rather than breadth is not merely psychometrically sound – it is a corrective to a cultural pathology that harms the very students it is designed to help.

The 10% Weight: Justifying the Allocation

The precise weight assigned to D5 in the scorecard composite is 10% – equal to Course Rigor (D3), below Academic Achievement (D1) at 20% and Standardized Testing (D2) at 15%, above College Knowledge (D6) at 8% and the noncognitive survey dimensions at 5-8%. This allocation is defensible from multiple angles.

The NACAC survey data, the most comprehensive national dataset on admission factor importance, shows extracurriculars rated as considerably important by 6-7% of colleges over the full 2007-2023 period, with combined moderate-and-considerable importance hovering around 50%. This places extracurriculars in the same tier as Course Rigor in terms of admissions salience – hence the matched 10% weight for both D3 and D5.

For selective institutions specifically – where the scorecard is most likely to be used as a differentiation tool – the importance of extracurriculars is considerably higher than the overall NACAC averages suggest. The 2018 NACAC data shows that more selective colleges (accepting fewer than 50% of applicants) rated extracurricular activities, essays, and recommendations substantially more highly than less selective institutions. Elite prep sources confirm: "Highly selective public and private universities look to qualitative factors, such as counselor and teacher recommendation letters, personal statements, and extracurricular activities, to assess how an applicant will fit in with and contribute to the campus community." A 10% weight for D5 is, if anything, conservative for students targeting highly selective admissions.

Conclusion: Depth as a Dimension of Readiness

The College Readiness Scorecard includes Extracurricular Depth as D5 because the evidence – from peer-reviewed longitudinal research, national surveys of admissions practices, institutional statements from selective colleges, and established college readiness frameworks – consistently supports the conclusion that sustained, meaningful

extracurricular engagement is a genuine signal of college readiness. It is not a signal of wealth, not a signal of access to private counseling, and not a signal of resume-padding ability. When properly measured – by years of commitment, leadership action, hours invested, and competitive achievement – it is a signal of the non-cognitive capacities that predict college success: discipline, identity, resilience, and the capacity for sustained effort toward meaningful goals.

The 10% weight assigned to D5 reflects the honest evidence about where extracurriculars sit in the landscape of college readiness: real, important, and secondary to academic achievement. The rubric's depth-over-breadth philosophy reflects the honest consensus among research, admissions practice, and equity considerations alike: what matters is not how many lines a student can fill on a resume, but whether they have invested themselves deeply enough in something to grow.

That growth is college readiness. D5 measures it.

PART III: KNOWLEDGE, SKILLS, AND MINDSET

Chapter 6: College Knowledge — The Gap Between Being Ready and Knowing You Can Go

Dimension D06 | Weight: 8% | Measurement: Nine-item behavioral checklist (campus visits, information sessions, counselor meeting, college list, recommender identification, essay brainstorming, FAFSA estimation, family cost discussion, scholarship research)

Introduction: The Gap Between Being Ready and Knowing You Can Go

Imagine a high school senior with a 3.7 GPA, strong AP coursework, and a genuine desire to attend college. She has every academic credential needed to succeed at a four-year institution. But she has never visited a campus, has not met with her school counselor about the application process, and has no idea how to file a FAFSA. In August, after receiving her diploma, she simply does not enroll.

This is not a hypothetical. Research documents it with sobering precision.

A landmark randomized controlled trial found that college-intending students with non-college-educated parents had a **35.6% intention-to-application gap** – nearly a third of students who fully intended to attend college never submitted an application, primarily because of information deficits, not academic deficits (Bettinger, Cunha, Lichand & Madeira, 2017). A 25-minute information session increased application rates by **11.9 percentage points** – an 18% relative improvement – without changing a single grade or test score.

This is why the College Readiness Scorecard includes D6: **College Knowledge**.

The D6 dimension measures nine concrete preparation behaviors: campus visits, information sessions, counselor meetings, college list creation, recommender identification, essay brainstorming, FAFSA estimation, family cost discussions, and scholarship research. Weighted at **8% of the composite score**, D6 occupies a deliberate and justified place in the overall readiness framework, recognizing that preparing for college is not just about academic achievement but about navigating a complex, high-stakes process that requires specific knowledge and proactive action.

Part I: College Knowledge as a Theoretically Distinct Readiness Construct

The first question any scorecard dimension must answer is: *Why is this a separate thing?* Could college knowledge simply be absorbed into Academic Achievement or Personal Qualities?

The answer, grounded in two decades of college readiness scholarship, is no.

David Conley's foundational framework *Toward a More Comprehensive Conception of College Readiness* (Gates Foundation, 2007) identifies **Contextual Skills and Awareness** as one of four essential readiness domains alongside academic content knowledge, key cognitive strategies, and academic behaviors. Conley defines this domain as "knowledge of information, formal and informal, stated and unstated, necessary for gaining admission to, and navigating within, the postsecondary system." He is unambiguous about the stakes: "Students who lack this knowledge are at a distinct disadvantage in terms of college access and success."

The Annenberg Institute at Brown University's synthesis *College Readiness: A Guide to the Field* (McAlister & Mevs, 2012) builds on Conley's framework, defining college knowledge as one of three co-equal pillars of readiness – alongside academic preparedness and academic tenacity – and defining it specifically as "the ability to gain access to postsecondary institutions as well as successfully navigate through college." The authors note that college knowledge encompasses knowing how to apply to college and for financial aid, developing appropriate relationships with professors, and functioning as a self-advocate. They are explicit that the three dimensions are interconnected: "Improvement in one dimension is influential and dependent on the other two dimensions."

Stanford's Gardner Center, UChicago's Consortium on School Research, and Brown's Annenberg Institute co-developed the **College Readiness Indicator System (CRIS)** specifically to operationalize these theoretical dimensions into measurable indicators. CRIS defines college knowledge as "knowledge, skills, and behaviors apart from academic content that allow students to successfully access college" and identifies individual-level indicators including campus visits, counselor meetings, application completion behaviors, and financial aid knowledge. Critically, CRIS validates that these indicators are "valid, reliable, and actionable" – not aspirational but measurable.

The theoretical consensus is clear: college knowledge is a **distinct, measurable construct** that predicts college access and success independently of academic metrics.

Part II: The Nine Checklist Items – Each One Earns Its Place

D6's nine items are not arbitrary. Each corresponds to a specific preparation behavior that professional guidance, federal policy, and empirical research identify as predictive of enrollment and success.

1 & 2: Campus Visits and Information Sessions

A national study using HSLS:2009 data tracking 12,217 students found that students who took college tours were **29.3% more likely to attend postsecondary education** (odds ratio = 1.293, p = .005) and 20.7% less likely to enter the workforce directly (Xing, Huerta & Garza, 2019). College search activities – information sessions, college fairs – increased postsecondary enrollment odds by **84.4%** (odds ratio = 1.844, p < .001). These are among the strongest behavioral predictors of enrollment in the literature.

A randomized controlled trial (Swanson et al., 2020) assigned eighth graders to either a college information packet or an information packet plus three campus visits. Treated students showed **significantly higher college knowledge, efficacy, and grit** and were more likely to enroll in advanced coursework in 9th grade. This is gold-standard causal evidence that campus visits do not just correlate with readiness – they build it.

A GEAR UP study of rural students in western North Carolina (Smith, Gosky & Li, 2022) found that both "traditional" campus visits (self-arranged) and "educational" visits (district-sponsored) were associated with higher postsecondary enrollment rates, with the strongest effects appearing when visits occurred in 7th grade or 10th grade.

The research verdict: campus visits and information sessions are not nice-to-haves. They are **enrollment predictors**.

3: Counselor Meeting

The same HSLS:2009 national study found that talking to a high school counselor reduced the odds of a work-only outcome by **20.3%** (odds ratio = .797, p = .037). Klasik's (2012) sequential analysis of the college enrollment "gauntlet" found that students who met with a college counselor or representative had **1.9 times higher odds of applying to a four-year college**.

This evidence becomes especially pointed when you consider the systemic disparity in counseling access. NACAC data show that public school counselors are responsible for an average of **476 students**, compared to 285 for private school counselors, and that public school counselors spend only **22% of their time** on postsecondary counseling versus 55% for their private-school counterparts. Many students – particularly those in under-resourced public schools – will never receive proactive college counseling unless they seek it out themselves.

D6 rewards students who take that initiative. Completing a counselor meeting is both a sign of college knowledge and an action that builds more of it.

4: College List Creation

Avery's (2010) experimental study found that high-achieving, low-income students literally did not know which colleges existed in certain selectivity tiers – they were unaware of the "Most Competitive Group 2" tier below the Ivy League. Students offered counseling applied to significantly more appropriate colleges and were **7.9 percentage points more likely to enroll in Most Competitive institutions**.

Creating a balanced college list – reach, target, and safety schools – requires knowledge that many students simply do not have without deliberate effort. NACAC recommends students "maintain organized files for each college and calendars of all deadlines." The D6 checklist item captures whether this foundational organizational behavior has occurred.

5: Identifying Recommenders

Admissions professionals are clear: strong recommendation letters require time and relationship depth. Professional guidance recommends asking recommenders at least **3-4 weeks before the first application deadline**, ideally at the end of junior year. The quality of a letter depends not on the title of the recommender but on how well they know the student: "it's not who's saying it; it's what they're saying" (Dr. Accrocco, Rice University).

The *Turning the Tide* report, endorsed by over **80 college admissions deans** from Harvard, MIT, Yale, Stanford, and other selective institutions, emphasizes that strong letters come from teachers who have observed sustained engagement, and that the selection of recommenders should reflect authentic relationship quality. Students who have not yet identified recommenders by the time they complete the scorecard are likely behind on a preparation task with real application consequences.

6: Essay Brainstorming

The college essay represents approximately **20% of the admission decision** at selective institutions and requires substantial planning, self-reflection, and multiple revision cycles (Rice University Center for College Readiness). The recommended timeline is to begin brainstorming in spring of junior year and complete the main essay over the summer before senior year. Students who wait until fall of senior year are compressed into a high-stress, deadline-driven process that typically produces weaker essays.

Including essay brainstorming in the D6 checklist captures an important signal: has the student engaged with the self-reflective, narrative-building work that defines the most distinctive part of their application? It is not enough to know that an essay exists – readiness means having begun the process.

7, 8 & 9: FAFSA Estimator, Family Cost Discussion, Scholarship Research

These three financial items overlap with D10 (Financial Preparedness) but serve a distinct function in D6: they measure whether a student has begun the **knowledge-gathering** steps that make financial aid accessible, as opposed to the administrative completion

behaviors themselves. A student who has used the FAFSA4caster, spoken with their family about college costs, and researched scholarships has acquired the informational foundation necessary to make enrollment decisions without being blindsided by cost.

Part III: The Equity Rationale – Who Lacks College Knowledge and Why It Matters

If college knowledge were equally distributed, measuring it might seem redundant. But it is not.

Klasik, Blagg & Pekor (2019) found that **53% of high-achieving, low-income students do not apply to selective institutions** under the traditional admissions process – not because they lack the grades or scores, but because they lack the cultural capital and informational scaffolding to navigate it. Only 8% of high-achieving, low-income students applied to selective institutions at rates comparable to their high-income peers.

Klasik's (2012) sequential gauntlet analysis found that Hispanic students maintained college aspirations from 10th to 12th grade at only **0.38 times the rate of White students**, and that momentum through the enrollment steps was substantially weaker for Black and Hispanic students – even after controlling for academic preparation. Only **36% of urban Boston students** maintained four-year college plans throughout senior year, compared to 93% of suburban students.

The Niche 2022 survey documented the specific mechanics of this gap:

- **54% of students** encountered barriers completing applications they wanted to finish – rising to 58% for underrepresented minorities, 57% for first-generation students, and 59% for low-income students
- Low-income students reported **no campus visits** at a rate of 25%, compared to 13% for higher-income students
- Only **29% of first-generation students** felt confident they could afford college
- First-generation students received parental support at **32%**, compared to 67% for non-first-generation peers

These are not small differences. They describe a systematic gap in college knowledge that falls most heavily on the students who most need access to higher education as a vehicle for economic mobility.

By including D6 in the scorecard at 8%, the College Readiness Scorecard does two things simultaneously. First, it **identifies** students who have not yet completed critical preparation steps – giving counselors an actionable signal to target support. Second, it **rewards** students who have taken proactive initiative to build their college knowledge, even if they attend under-resourced schools with limited guidance infrastructure.

Bettinger et al.'s finding that a 25-minute information intervention reduced the intention-application gap by **33%** suggests that information deficits – not just economic constraints – are a primary barrier. This validates D6 as a *causal* domain: building college knowledge through concrete preparation behaviors directly increases the probability of enrollment.

Part IV: Why 8%? Calibrating the Weight

The 8% weight assigned to D6 is neither arbitrary nor purely intuitive. It reflects a deliberate calibration against both the research evidence and the scorecard's overall weighting philosophy.

The NACAC State of College Admission data (2019) show that **75% of colleges** rate overall GPA as considerably important, **73%** rate grades in college prep courses, and **62%** rate strength of curriculum – while activities, counselor recommendations, and demonstrated interest receive considerably lower ratings from most institutions. This evidence justifies allocating the highest weights to academic dimensions: D1 Academic Achievement at 20%, D2 Standardized Testing at 15%, D3 Course Rigor at 10%.

College knowledge does not appear on an application form in the same direct way that GPA, test scores, or AP coursework do. Yet the behaviors D6 measures are *prerequisites* for completing the application successfully. A student cannot apply without selecting recommenders, cannot enroll without understanding financial aid, and cannot make fit-appropriate choices without visiting campuses or creating a college list. In this sense, D6 functions as a **process competency**: it measures whether a student is operationally ready to complete the enrollment sequence, which is why it warrants more weight (8%) than the purely enabling dimensions of College Motivation (D9, 5%) and Financial Preparedness (D10, 5%).

McAlister and Mevs (2012) describe college knowledge as "a necessary but not sufficient condition for college success." This phrase is precisely calibrated: necessary means it cannot be zero (hence 8% is appropriate), but not sufficient means it should not dominate the composite (hence it ranks below academic dimensions). The Soland (2017) machine learning study using 1,000+ predictors on 12,144 NELS:88 students found that the strongest single predictor of college enrollment was educational aspiration – supporting the validity of the combined weight assigned to D6-D10.

The 8% weight also reflects that college knowledge is **more remediable** than academic dimensions. A student who has not yet visited a campus or met with a counselor can do so within weeks. A student with a 2.3 GPA cannot reverse that metric in the same timeframe. This makes D6 an ideal intervention target – and its 8% weight ensures that students who complete these preparation steps receive meaningful recognition in the composite score.

Part V: Practical Implications for Counselors and Students

The checklist structure of D6 reflects a philosophy that college knowledge is best measured and built through **specific, actionable behaviors** rather than abstract self-assessment.

CRIS (Gardner Center/UChicago CCSR/Annenberg, 2013) explicitly endorses this approach, noting that observable behavioral proxies for the underlying college knowledge construct. The IES Practice Guide (2009), drawing on 500+ studies of college access interventions, recommends that high schools provide "hands-on assistance and guidance"

in completing specific preparation steps because "simply providing students with information is insufficient."

For **counselors**, the D6 subscores translate directly into counseling agendas. A student who has visited campuses but has not met with a counselor about the college list needs a different conversation than one who has identified recommenders and begun essay brainstorming but has not addressed financial planning. The nine-item checklist makes the diagnostic sharp enough to be actionable.

For **students**, D6 provides a roadmap. The behaviors on the checklist are within every student's reach – they do not require expensive test prep, a private college counselor, or a college-educated parent. They require initiative, planning, and the knowledge that these steps matter. That is exactly what a college readiness scorecard should communicate.

The U.S. Department of Education's Federal Student Aid office identifies campus visits, counselor meetings, college list creation, and FAFSA estimation as "must do" preparation behaviors for college-bound students. D6 makes those federal recommendations measurable and scoreable.

Conclusion: College Knowledge Is Readiness

The case for D6 is not a concession to softness in the scorecard. It is a rigorous commitment to measuring the full architecture of college readiness.

Academic metrics – GPA, test scores, course rigor – predict whether students can succeed in college-level work. But they say nothing about whether a student will ever walk through the door. College knowledge is what bridges that gap. It is the difference between potential and enrollment, between aspiration and application, between deserving to go to college and actually getting there.

The evidence is substantial and varied: randomized experiments showing that campus visits causally increase college knowledge and enrollment behaviors; national longitudinal data showing that college tours increase enrollment odds by 29%; counseling access data showing that 476:1 student-to-counselor ratios in public schools leave most students without proactive guidance; and field-level frameworks from Stanford, UChicago, Harvard, and the U.S. Department of Education all identifying the same preparation behaviors as critical milestones.

At 8% of the composite score, D6 occupies exactly the right position: essential but not dominant, actionable but not trivial, evidence-based but not reductive. It measures what matters – not just what is easy to count.

For students who have grown up without the informal college knowledge that flows naturally through college-educated families, D6 is not a penalty. It is an invitation. Complete these steps. Build this knowledge. The door is open – and now you know how to walk through it.

Key Sources: Bettinger et al. (2017); Conley (2007); Gardner Center/CCSR/Annenberg (2013); Harvard, *Turning the Tide* (2016); IES Practice Guide (2009); Klasik (2012); Klasik, Blagg & Pekor (2019); McAlister & Mevs (2012); NACAC (2019); Niche (2022); Soland (2017); Smith, Gosky & Li (2022); Swanson et al. (2020); U.S. Department of Education (2013); Xing, Huerta & Garza (2019).

Chapter 7: Self-Awareness and Mindset — The Inner Architecture of College Readiness

Dimension D07 | Weight: 8% | Measurement: 18-question Likert survey — Growth Mindset (6 questions, 40%) + Emotional Vocabulary (6 questions, 40%) + Tenacity (6 questions, 20%)

Introduction: The Limits of the Transcript

For decades, the college admissions conversation has revolved around two numbers: GPA and a standardized test score. They are easy to compare, easy to rank, and easy to defend. But they leave out something every college counselor, every first-year advisor, and every college student who has ever failed a class despite "having the grades" already knows: the transcript cannot tell you whether a student believes they can grow.

Dimension 7 of the Scorecard – Self-Awareness and Mindset – carries an 8% weight in the overall composite. That number is deliberate. It reflects a body of research that is compelling but appropriately modest: noncognitive constructs like growth mindset, emotional vocabulary, and tenacity are real, measurable, and meaningfully predictive of college outcomes. They are not, however, a substitute for academic preparation. What follows is the case for why they belong in the assessment at all, how they are measured, and what the evidence actually says about their predictive power.

Part I: Why the Scorecard Is Multi-Dimensional

The Scorecard allocates 45% of its composite weight to academic factors (GPA, course rigor, and standardized testing), and distributes the remaining 55% across seven other dimensions. Critics of this approach sometimes argue that spreading weight across dimensions like self-awareness or financial preparedness dilutes the assessment's validity. The research suggests the opposite.

A landmark meta-analysis conducted by Lotkowski et al. (2004) for ACT found that academic factors – GPA and test scores – account for only about 25% of the variance in college GPA. Psychosocial factors add 13% of incremental predictive validity beyond academics alone. In other words, if you measure only what shows up on a transcript, you are missing more than three-quarters of the story. David Conley's Four Keys to College Readiness framework, one of the most widely cited models in the field, identifies four essential components: key cognitive strategies, academic knowledge and skills, academic behaviors, and *contextual skills and awareness* – the last of which encompasses self-awareness, identity development, goal-setting, and self-advocacy. Conley notes that

contextual skills "may be more readily developed through specific interventions" compared to content knowledge, but he is equally clear that their absence is a serious readiness gap.

William Sedlacek's Noncognitive Questionnaire (NCQ), validated across decades and multiple institutional contexts, provides perhaps the most striking data point: the Gates Millennium Scholars program, which uses noncognitive variables as a central selection criterion, achieved a 97% first-year retention rate and an 87% five-year retention rate among its more than 11,000 scholars at 1,450+ institutions – substantially above national averages. The scorecard does not weight noncognitive factors at 80% the way the Gates program does; it weights them at 35% combined (D4 through D10), reflecting their genuine importance while maintaining the primacy of academic preparation.

The point is not that mindset is more important than calculus. The point is that the multi-dimensional structure of the Scorecard is not an innovation or an ideology – it mirrors how colleges actually evaluate applicants, and it reflects what the research tells us about what predicts success once a student arrives.

Part II: What Self-Awareness and Mindset Actually Measure

Dimension 7 is composed of an 18-question Likert survey divided into three sub-scales:

- **Growth Mindset** (6 questions, 40% of D7)
- **Emotional Vocabulary** (6 questions, 40% of D7)
- **Tenacity** (6 questions, 20% of D7)

Each construct is measured separately, each is grounded in its own research tradition, and each contributes something distinct to the readiness picture.

Growth Mindset

Carol Dweck's foundational work on implicit theories of intelligence has generated over two decades of research and thousands of studies. Her validated 3-item Growth Mindset Scale – which the Scorecard's 6-question subscale is modeled on – uses reverse-phrased items that ask students whether they believe intelligence is fixed or malleable. The psychometric properties are well-established: satisfying reliability, convergent and discriminant validity, and longitudinal predictive power across developmental transitions.

The predictive evidence is meaningful. A 2022 randomized controlled trial at a large public university found that a growth mindset intervention improved term GPAs for Latinx students and increased the probability for Pell-eligible and Latinx students to major in selective fields. The Institute of Education Sciences' What Works Clearinghouse has reviewed growth mindset interventions with outcomes including college credit completion, GPA, and persistence. Blackwell, Trzesniewski, and Dweck (2007) showed that implicit theories of intelligence predict achievement across adolescent transitions. Compared to students with fixed mindsets, those with growth mindsets like school more, take on more challenging tasks, learn more, and earn better grades.

But the research also demands honesty about limitations. As Dweck herself notes on her official research page, she "discourages the use of these scales in high-stakes settings where faking is a concern (e.g., admissions or hiring decisions)." The Scorecard takes this seriously: D7 is framed as a *self-assessment tool*, not a gatekeeping instrument. Students are encouraged to use it to understand their own readiness profile, not to game a number. The 40% sub-weight for growth mindset within D7 – and the 8% weight of D7 within the overall composite – reflects calibrated confidence: the construct matters, but it should not dominate.

Emotional Vocabulary

The second sub-scale measures a construct that receives less popular attention than grit or growth mindset but has equally strong empirical support: the ability to identify and label one's own emotions.

A 2019 meta-analysis of 160+ studies with over 42,000 students from 27 countries, published in *Psychological Bulletin*, found that emotional intelligence significantly predicts academic performance – and holds across all ages from elementary school through college. Students with higher EI get better grades and higher test scores than those with lower EI, even after controlling for intelligence and personality. The mechanism is self-regulation: students who can identify their emotions are better equipped to manage the anxiety, boredom, and disappointment that derail academic performance.

A 2022 dissertation study of 181 participants found that emotion vocabulary significantly predicted interpersonal relations ($p = .032$), social roles ($p = .005$), and acute emotional distress ($p < .001$). Students with stronger emotion vocabulary experienced less distress in interpersonal relationships and were better able to navigate social roles. The researchers note that "the ability to verbally describe one's own emotions significantly impacts their overall emotional adjustment."

The CASEL framework – the most widely adopted social-emotional learning model in U.S. education – explicitly defines self-awareness to include "identifying one's emotions" and "linking feelings, values, and thoughts." The research on SEL confirms benefits that are, in the words of one *Inside Higher Ed* contributor, "sustainable and equitable, working effectively across categories including race, ethnicity, income, gender and age." A 2025 peer-reviewed study in *Social and Emotional Learning: Research, Practice, and Policy* found that students who practiced SEL skills outside of class experienced measurable improvements in wellbeing – and that skills practice predicted outcomes more strongly than baseline emotional intelligence, class attendance, or course grades.

Emotional vocabulary receives 40% of D7's weight – equal to growth mindset – because the research treats it as a co-equal foundational construct. The rationale follows established psychometric principles for composite assessment: when two constructs are theoretically co-equal and both predict the same outcomes through distinct pathways (one cognitive, one affective), they merit equivalent weight.

Tenacity

The third sub-scale measures perseverance and sustained effort toward long-term goals – what Angela Duckworth defined as "grit" in her foundational 2007 paper in the *Journal of Personality and Social Psychology*. That study, across six samples including Ivy League undergraduates, Military Academy cadets, and adult professionals, found that grit accounted for an average of 4% of the variance in success outcomes – and demonstrated incremental predictive validity beyond both IQ and conscientiousness. The key finding, often quoted but worth restating: "the achievement of difficult goals entails not only talent but also the sustained and focused application of talent over time."

A longitudinal study at an HBCU tracked business students across five years and found a significant positive correlation between grit scores and both GPA and persistence to graduation. Importantly, first-year GPA was *not* a reliable predictor of long-term success – but grit was. This finding speaks directly to the Scorecard's rationale for including tenacity: the transcript shows performance at a single point in time; tenacity predicts whether a student will persist when performance dips.

Tenacity receives a lower sub-weight (20%) than growth mindset and emotional vocabulary (40% each) for a reason grounded in psychometric theory. Tenacity is conceptualized as a *behavioral manifestation* of mindset rather than a foundational cognitive-affective construct. As Yan, King, and Haw (2021) note, "formative assessment emphasises the external teaching process, while growth mindset underscores the importance of internal psychological processes" – and behavioral persistence is downstream of those internal processes. Primary constructs receive higher weights; secondary manifestations receive lower weights. The 40/40/20 structure mirrors established emotional intelligence instruments like the Bar-On EQ-i, which uses comparable weighting structures for related but distinct competencies.

Part III: The 8% Weight – Calibrated, Not Arbitrary

The most important design decision in D7 is also the easiest to misread: why 8%?

The answer emerges from two directions simultaneously – the research on predictive validity and the research on measurement validity.

On predictive validity: Jim Soland's data-driven analysis of the NELS dataset, published in *Educational Researcher* in 2017, used machine learning to identify the most predictive variables for college enrollment and persistence from thousands of candidates. The result was a model achieving 90% accuracy using four primary predictor categories: academic preparation (strongest), educational aspirations, socioeconomic status, and teacher perceptions. Noncognitive variables – including attitudes, beliefs, resilience, and mindset – appeared in the model as significant but secondary predictors. Soland notes that "in some instances, these measures of attitudes and beliefs have predicted long-term academic achievement better than grades and test scores," but emphasizes that this is context-dependent. The 8% weight reflects this finding: meaningful, but subordinate to the academic dimensions that collectively carry 45% of the composite.

On measurement validity: The Brookings Institution's Martin West (2014) studied 1,300+ eighth-grade students in Boston and found that self-reported noncognitive skills correlated with behavioral outcomes within schools – but that reference bias creates systematic distortion when comparing across schools. Students in high-expectation environments may rate themselves *lower* despite actually demonstrating stronger skills, because they compare themselves to peers who set a higher bar. This is not a reason to abandon self-report measurement – it is a reason to weight it appropriately and interpret it contextually. The 8% weight for D7 operationalizes this caution: it gives self-awareness constructs meaningful influence on the composite without allowing reference bias to drive the overall assessment.

The weight also has precedent in established frameworks. Conley's Four Keys model positions contextual skills and awareness – which directly encompasses D7's constructs – as one of four essential readiness components but acknowledges their secondary role relative to cognitive preparation. Sedlacek explicitly recommends "adding" noncognitive variables to existing academic measures rather than replacing them. The 8% weight does exactly that.

Part IV: Who This Matters For

The research on mindset and self-awareness takes on particular urgency when viewed through an equity lens.

Researchers from USC and the University of Michigan have documented that over 90% of Americans aspire to attend college regardless of race or socioeconomic status – but only one-third earn bachelor's degrees. The problem, they argue, is not low aspirations but "low support compounded by economic segregation." Students in resource-poor environments may develop what the researchers call an alternative cognitive framework: that difficulty in schoolwork signals low odds or impossibility rather than importance. In a resource-rich environment, difficulty in schoolwork "feels like things that 'people like me' do." In a resource-poor environment, the same difficulty can feel like evidence that college is not for them.

This distinction maps directly onto the growth mindset construct measured in D7. A student who interprets a difficult assignment as evidence that they cannot succeed is at risk of disengagement before they ever apply. A student who interprets the same difficulty as a signal that the work matters – and that persistence is the appropriate response – has a meaningful edge on the road to persistence and graduation.

The pandemic literature reinforces this point. ACT's survey of 1,549 twelfth-graders from the Class of 2023 found that over 40% changed their college major or career plans due to COVID-19, and 1 in 10 questioned whether they should attend college at all. Gen Z more broadly shows striking awareness: 70% of teens report seeing anxiety and depression as a major problem among peers. Students who can identify and name those emotional states – who have the emotional vocabulary the Scorecard measures – are better positioned to seek help, manage stress, and persist through the transition to college.

Sedlacek's data is perhaps the most compelling argument for including noncognitive assessment at all. His NCQ-based systems have been used at institutions ranging from

Oregon State to the Gates Millennium Scholars program, with demonstrated effects on retention and graduation. He argues that noncognitive variables "provide viable alternatives in fairly assessing the abilities of people of color, women, international students, older students, students with disabilities, or others with experiences that are different from those of young, White, heterosexual, able-bodied, Eurocentric males." A purely academic scorecard disadvantages students whose full potential is not captured by grades and test scores. A multi-dimensional scorecard, including self-awareness and mindset, creates more pathways for more students to demonstrate readiness.

The Scorecard's equity profile is mixed, and intellectual honesty requires acknowledging that. The 8% weight for D7 means that a student with strong growth mindset and emotional regulation cannot fully compensate for a weak academic record. Nor should they be able to: academic preparation remains the strongest predictor of college success, and a tool that suggested otherwise would mislead students. But the inclusion of D7 means that a student with equivalent academic credentials to a peer, but stronger self-awareness and mindset, will score higher – and that difference reflects something real about their readiness.

Part V: The Survey Instrument

The 18 questions in D7 use a 4-point forced-choice Likert format: Strongly Agree, Agree, Lean Disagree, Disagree. The absence of a neutral midpoint is intentional. Psychometric literature on forced-choice scales (Garland, 1991) demonstrates that avoiding the midpoint reduces central tendency bias, compelling respondents to take a position. Responses are mapped to scores of 9, 6, 3, and 0, and raw scores are normalized to the 1-9 scale using the standard min-max formula: $1 + (\text{sum}/\text{max}) \times 8$.

The 4-point format has direct validation from Hardin et al. (2021), who validated a College-Going Self-Efficacy Scale using exactly this format with rural Appalachian high school students across two independent samples – and found good reliability (internal consistency) and construct validity, including measurement equivalence across gender and first-generation college status. A 4-point scale successfully captured variance in self-efficacy beliefs related to college attendance and persistence, and the research confirmed its practical utility for comprehensive readiness assessment instruments.

For growth mindset items, the D7 survey parallels Dweck's validated 3-item Growth Mindset Scale, expanded to 6 questions to improve construct coverage – an approach supported by a 2024 validation study published in the *Journal of Career Assessment*, which found that expanding from 3 to 8 items maintains strong psychometric properties and improves breadth of measurement. For emotional vocabulary, the construct aligns with CASEL's definition of self-awareness. For tenacity, the items draw from Duckworth's grit construct, with the understanding that the Scorecard's purpose (self-assessment for students and counselors) differs from the high-stakes use cases Duckworth explicitly cautions against.

Conclusion: What 8% Means

Eight percent of a ten-dimension composite might sound small. In absolute terms, it is roughly what separates a strong assessment from an excellent one on the 1-9 scale. But the significance of D7 is not primarily about the number – it is about the signal.

By including growth mindset, emotional vocabulary, and tenacity as a formal dimension, the Scorecard communicates something to students: how you think about your own capacity to learn, how well you can identify and manage your emotional responses to challenge, and whether you can sustain effort toward long-term goals – these are not soft skills. They are measurable, they are predictive, and they matter.

They matter for the student with a 4.0 GPA who has never faced academic adversity and does not know how to respond when college delivers its first real setback. They matter for the first-generation student from a resource-poor school who has never had a counselor help them connect difficulty to growth. They matter for the student who cannot articulate why they feel anxious about leaving home and therefore cannot seek help before that anxiety compounds into a withdrawal.

The research supporting D7 is honest about what self-awareness and mindset measures can and cannot do. They are not a replacement for academic preparation. They are not a high-stakes gatekeeping tool. They are an 18-question window into the inner architecture of a student's readiness – the part that shows up not on the first day of college, but on the day things get hard.

That is what 8% represents. And it earns every decimal point.

Key Sources: Duckworth et al. (2007); Dweck (2006); MacCann et al. (2019); CASEL (2020); Sedlacek (2011); Conley (2007, 2012); Soland (2017); West (2014); Buzzetto-Hollywood & Mitchell (2019); Kim et al. (2022); Hardin et al. (2021); Gettenberg (2022); Blackwell, Trzesniewski & Dweck (2007); NACAC (2019); Conley, Simons et al. (2025).

Chapter 8: Academic Skills — The Dimension That Asks *How* a Student Learns

Dimension D08 | Weight: 7% | Composition: Academic Ownership (60%) + Participant Learner (40%) | Measurement: 12-item Likert survey

Introduction: Beyond Grades and Test Scores

For decades, the college admissions process has leaned heavily on two numbers: a GPA and a standardized test score. Yet decades of research make clear that these metrics, while important, capture only a fraction of what predicts whether a student will thrive in college. By measuring ten distinct dimensions of readiness – weighted according to empirical evidence on what actually matters – the Scorecard gives students, counselors, and institutions a far more complete picture of who is prepared for postsecondary success.

Dimension 8 (D8), Academic Skills, is one of the most distinctive components of that picture. At 7% of the total composite score, it sits apart from raw academic achievement (D1, 20%) and standardized testing (D2, 15%) for a reason: how a student learns is not the same thing as what a student has already achieved. Two students can arrive at college with identical GPAs having gotten there through radically different habits. One has internalized a practice of goal-setting, self-monitoring, and independent effort. The other has coasted on talent, relied on parental reminders, and never developed the self-directed study skills that college demands. The Scorecard measures the difference.

D8 occupies a precise and necessary role in this framework. As David Conley – whose Four Keys to College and Career Readiness framework has informed college access policy for nearly two decades – makes clear, college readiness requires four distinct components: Key Cognitive Strategies, Academic Knowledge and Skills, **Academic Behaviors**, and Contextual Skills. Academic behaviors – what the Scorecard calls Academic Skills – are not captured by past performance data and are not a personality trait. They are a teachable, measurable set of habits and strategies that predict college success independently.

As Conley wrote: *"Academic behaviors include time management, study skills, goal setting, self-awareness, persistence, and student ownership of learning… These are critical to success in college but often not addressed systematically in high school programs."*

This is precisely why D8 exists as a standalone dimension and why it carries meaningful weight.

Self-Regulated Learning: The Science Beneath the Survey

D8 is grounded in more than 30 years of research on self-regulated learning (SRL), a field pioneered by Barry Zimmerman and colleagues. SRL is defined as "self-generated thoughts, feelings, and actions systematically oriented toward the attainment of one's goals" – not an innate trait, but a dynamic process that students can develop and improve.

Zimmerman describes three cyclical phases: **Forethought** (goal setting and strategic planning before a task), **Performance** (self-control and self-observation during a task), and **Self-reflection** (self-judgment and adjustment after a task).

These phases map directly onto D8's survey constructs. The Academic Ownership sub-construct (60% of D8) measures forethought and self-reflection behaviors: does the student set priorities, manage their time without parental prompting, and monitor their own progress? The Participant Learner sub-construct (40% of D8) measures performance-phase behaviors: does the student engage actively with instructors, ask questions when confused, and contribute to classroom discourse?

Empirical validation of this framework is substantial. A 2015 study of 166 university students found that goal-setting – a core forethought behavior – was a significant predictor of college GPA ($\beta = 0.333$, $p < .001$). A systematic review of 70 studies published between 2015 and 2024 found that cognitive and metacognitive strategies (measured in 35 and 37 studies, respectively) showed the most consistently positive impacts on academic

achievement. Critically, the review emphasized that *quality* of self-regulation matters more than quantity – students who self-monitor and adjust their strategies continuously outperform those who merely attempt to stay engaged.

A replication study by Michaelides and Durkee (2021) further refined this picture, finding that self-regulation of effort was a significant predictor of semester GPA ($\beta = 0.336$, $p < .001$) – even after controlling for academic self-efficacy – while self-discipline as a personality trait was not ($\beta = -0.052$, $p = 0.454$). This distinction is crucial: D8 does not measure whether a student is a disciplined person by nature. It measures whether a student has developed the specific, learnable behaviors of academic self-regulation.

Part II: The Two Constructs — Academic Ownership and Participant Learner

Why Two Sub-Constructs, and Why 60/40?

D8 is structured around two empirically distinct behavioral clusters. **Academic Ownership** (60% of D8) captures the self-directed habits that allow students to manage their own learning independently of external prompts – prioritizing study time, completing homework autonomously, taking notes systematically, knowing when to seek help, and believing they can direct their own success. **Participant Learner** (40% of D8) captures students' readiness to engage actively and productively in the social and intellectual life of the classroom – asking questions, contributing to discussions, and building the kind of faculty relationships that research consistently shows predict retention and achievement.

The 60/40 weighting reflects a clear empirical signal: ownership is the higher-order construct. The National Institute for Excellence in Teaching articulates this distinction precisely: *"Student ownership is evident when students can articulate what they are learning, why they are learning, strategies that support their learning, and how they will use these strategies in the future. When students own their learning, they are doing more than just engaging: They are actively taking a role in leading their learning."*

Engagement reflects "doing" and "understanding." Ownership adds the capacity to apply learning, take responsibility for outcomes, and self-direct improvement. Students can be engaged in a lesson without taking ownership of their performance. The reverse – true ownership without some degree of engagement – is nearly impossible. Ownership therefore has the broader and more foundational predictive claim on D8's weight.

Empirical support for this hierarchy is direct. Ellis and Helaire (2017) found that self-regulated learning is a "vital indicator of college readiness and a proximal behavior that enhances academic achievement," and that students enrolled in a self-regulated learning course were **13 times more likely to graduate** from their institution than peers in a comparison group. Perceived behavioral control – students' sense of capability to perform SRL behaviors (the essence of ownership) – predicted engagement in college readiness activities more strongly than intentions or attitudes.

Academic Ownership: The Evidence

The behaviors measured by Academic Ownership – self-directed study, time management, autonomous homework completion, strategic help-seeking, self-motivation – have a large and consistent research base.

A 2013 study by Komarraju, Ramsey and Rinella found that ACT scores predicted 13% of variance in college GPA, high school GPA predicted an additional 11%, and academic discipline – defined as "the amount of effort students put into schoolwork and the extent to which they see themselves as being hardworking" – predicted a further 2% beyond both. While 2% may sound small, it represents unique, actionable information not captured by either D1 or D2 – and it validates D8's status as a distinct dimension in the Scorecard rather than a redundant measure.

Perry and colleagues (2001) conducted longitudinal research showing that perceived academic control – belief that one can direct one's own academic outcomes – predicted first-year course achievement and persistence even after controlling for prior achievement and ability. Students who attributed academic success to internal, controllable factors (effort, strategy use) outperformed those who attributed outcomes to external factors (luck, teacher quality). Students lacking academic ownership were more likely to drop courses or leave college in the first year.

A 2017 Stanford study by Chen and colleagues provides particularly vivid evidence of the causal power of academic ownership behaviors. College students who adopted a strategic, metacognitive approach to using study resources – identifying what would be tested, choosing resources purposefully, and reflecting on their approach – improved exam scores by an average of **one-third of a letter grade** relative to peers who studied with equal effort but no deliberate strategy.

This finding captures D8's core premise in a single sentence: raw effort and ability are necessary but not sufficient. How students approach their learning – whether they own it – is a distinct predictor of success.

Participant Learner: The Evidence

The case for including active classroom engagement behaviors in a college readiness assessment is, if anything, even stronger. A landmark 2010 meta-analysis by Credé and colleagues found that **class participation is a better predictor of college grades than any other known academic factor, including scores on standardized admissions tests such as the SAT**. This is a striking finding, given that D2 (Standardized Testing) carries 15% of the Scorecard's composite weight.

Further evidence from Alexander Astin's foundational involvement research establishes that "the most basic form of academic involvement – studying and doing homework – has stronger and more widespread positive effects than almost any other involvement measure." Tinto's integration model (1993) showed that the more students make contact

with faculty and peers about learning issues, the more they learn and the more likely they are to persist.

A 2014 meta-analysis by Freeman and colleagues, covering 225 studies of active learning in STEM disciplines, found that active learning increased examination grades by an average of **half a letter grade** and reduced failure rates by 55% compared to traditional lecture instruction. A follow-up analysis by Theobald and colleagues (2020) found that active learning strategies also reduced achievement gaps for underrepresented students – making participation behaviors an equity-promoting readiness dimension, not just a predictor of individual success.

Research by Roderick, Coca and Nagaoka (2011) on Chicago Public Schools graduates provides longitudinal evidence connecting these classroom behaviors to college enrollment itself. Among students aspiring to four-year degrees, only 41% actually enrolled – and the gap was explained not by qualifications but by school climate and student-teacher relationships. Students in schools with strong college-going climates – measured by teacher expectations and robust student engagement – were 9 to 13 percentage points more likely to complete each application step. Teacher relationships, built through the kind of active participation the Scorecard measures, proved more predictive of college enrollment than counselor ratios.

Part III: Measurement — Why a Likert Survey Is the Right Tool

Validity of Self-Report Academic Skills Assessment

Some may question whether a student self-report survey is an appropriate instrument for measuring academic skills. The evidence says clearly: yes. The Learning and Study Strategies Inventory (LASSI), a validated 60-item Likert instrument developed by Weinstein, Palmer and Acee, has been in use at postsecondary institutions since 1987 and is now in its third edition with norms drawn from 1,386 students across 23 institutions. Coefficient alpha reliability coefficients range from .76 to .87 across all ten scales, and test-retest reliability over a 3-4 week interval is .88 for the total instrument.

The LASSI's Self Testing (α = .80) and Time Management (α = .80) scales directly parallel D8's Academic Ownership construct. Its Concentration (α = .85) and Information Processing (α = .81) scales relate directly to active learning behaviors. Critically, LASSI scores predict college GPA, course completion, and academic persistence – not just because they correlate with prior achievement, but because they add unique predictive variance.

A 2025 validation study using 1,135 Doctor of Physical Therapy students further confirmed that the LASSI is "a reliable and valid measure of learners' attitudes, self-regulatory habits, and metacognitive skills." If a 60-item self-report survey can identify students at risk for failure in doctoral-level health professions programs, a 12-item survey measuring the same behavioral constructs in high school students is methodologically sound.

The Scorecard's 4-point Likert format (Strongly Agree / Agree / Lean Disagree / Disagree) also has direct validation. Hardin and colleagues (2021) validated a short-form college-

going self-efficacy scale using a 4-point Likert format with rural Appalachian high school students and found good reliability and construct validity across gender and first-generation college status. The forced-choice format – without a neutral midpoint – reduces central tendency bias, and the conversion of responses to a 9/6/3/0 numerical scale before summing ensures maximum sensitivity within the Scorecard's 1-9 composite framework.

Finally, practitioner validation is provided by the Antonoff Self-Survey for the College-Bound, a published college counseling instrument that independently identifies Academic Ownership and Participant Learner as distinct measurable dimensions of college readiness – using a matching 4-point scale converted to numerical values. The fact that a widely-used, practitioner-adopted self-assessment instrument mirrors D8's constructs and measurement approach confirms that these are not theoretical abstractions but recognized categories of professional college counseling practice.

Why D8 Is Distinct from D1, D7, and D9

A rigorous assessment instrument must demonstrate discriminant validity – that each dimension measures something meaningfully distinct. Three potential concerns are worth addressing directly.

D8 versus D1 (Academic Achievement). D1 measures what a student has already accomplished: GPA, class rank, grade trends. D8 measures how a student approaches learning: whether they plan, self-monitor, participate actively, and take ownership of their progress. Komarraju et al. (2013) demonstrated that academic discipline predicts college GPA above and beyond high school GPA and ACT scores. The ACT Engage research (Casillas et al., 2012) showed that academic behaviors measured in middle school predicted high school GPA above and beyond academic achievement measured by test scores. D8 captures the learning mechanisms that produce D1 outcomes – and has incremental predictive validity when D1 scores are held constant.

D8 versus D7 (Self-Awareness and Mindset). D7 measures psychological orientations: growth mindset, emotional self-awareness, and grit. D8 measures behavioral strategies: time management, homework completion, class participation, note-taking. Michaelides and Durkee (2021) demonstrated that self-regulation of effort (D8 territory) predicts GPA while self-discipline as a personality trait (closer to D7 territory) does not – a direct empirical separation of the two constructs. Students can have a strong growth mindset (D7) without having developed effective study habits (D8), and vice versa.

D8 versus D9 (College Motivation). D9 measures eagerness and enthusiasm for college – the affective orientation toward postsecondary education. D8 measures the specific behavioral repertoire students use to succeed academically. A student can be highly motivated (D9) but lack the self-regulated learning skills (D8) to translate that motivation into performance. Ellis and Helaire (2017) found that perceived behavioral control – students' sense of capability to perform self-regulation behaviors – predicted engagement in college readiness activities more strongly than attitudes or intentions, suggesting D8 constructs are functionally distinct from motivational variables.

Part IV: Weight Justification — Why 7%?

D8's 7% weight places it in the Scorecard's third tier: meaningful, but subordinate to foundational academic achievement (D1: 20%, D2: 15%, D3: 10%) and to holistic behavioral dimensions like Personal Qualities (D4: 12%) and Extracurricular Depth (D5: 10%).

This weight reflects the evidence precisely. Academic behaviors explain unique variance in college success – but that variance is incremental, not dominant. The ACT research synthesis found that roughly 50% of college GPA variance is explained by academic ability, with non-cognitive factors explaining an additional 23-27%. D8 represents a share of that 23-27%. The Komarraju et al. finding that academic discipline adds 2% of unique variance in college GPA beyond GPA and test scores provides a floor justification for D8's inclusion; the broader SRL literature, showing self-regulation accounts for substantially more variance in college persistence and graduation, provides the ceiling.

Conley's Four Keys framework positions Academic Behaviors as one of four approximately co-equal components of readiness – which, if the scorecard were purely theoretically derived, might suggest weights of roughly 25% each. The Scorecard's 7% for D8 is appropriately conservative: it gives academic behaviors their due as a distinct, empirically validated readiness dimension while acknowledging that past academic performance (D1-D3) remains the strongest direct predictor of college academic success.

The ACT Engage research supports this calibration. Allen and Robbins (2010) found that academic discipline predicted timely degree attainment incrementally over prior test scores and college GPA – a meaningful finding – but GPA and test scores still explained the larger share. The Scorecard mirrors this hierarchy: D1 and D2 together carry 35%, while D8 carries 7%.

BetterLesson's widely-implemented College and Career Readiness Framework offers a contemporary validation of this hierarchy. In that framework, Academic Mindsets and Behaviors are weighted as a second-tier dimension – behind Academic Preparedness, ahead of College and Career Navigation and Equity/Identity Development. The Scorecard's structure is consistent with this pattern.

Conclusion: Academic Skills as a Standalone Dimension of Readiness

The inclusion of D8 in the College Readiness Scorecard is not a philosophical statement about holistic education. It is an empirically grounded decision supported by three decades of self-regulated learning research, large-scale meta-analyses on classroom engagement, validated psychometric instruments, and the most widely-cited operational frameworks in the college readiness field.

Class participation predicts college grades better than SAT scores. Strategic study behaviors improve exam performance by a third of a letter grade. Students enrolled in self-regulated learning courses graduate at 13 times the rate of comparison groups. These are not marginal effects. They are evidence of a domain of readiness that GPA and test scores simply do not measure.

The College Readiness Scorecard does not ask whether a student has performed well in high school. It asks, comprehensively, whether a student is ready for college – ready in the habits, strategies, and behaviors that postsecondary environments demand. D8, Academic Skills, is the dimension that answers the question: not *what* has this student achieved, but *how* will they learn?

Key Sources: Zimmerman (2002); Çetin (2015); Credé et al. (2010); Freeman et al. (2014); Ellis & Helaire (2017); Michaelides & Durkee (2021); Komarraju et al. (2013); Chen et al. (2017); Weinstein et al. (2016); Bekkering & Ward (2021); Conley (2007, 2012); Roderick et al. (2011); Antonoff (2022); Liu, Abu Bakar & Xu (2025); Swanson & Cole (2022); Perry et al. (2001); Theobald et al. (2020); BetterLesson CCR Framework (2024); Hardin et al. (2021); LASSI User's Manual (2016).

PART IV: MOTIVATION AND FINANCIAL READINESS

Chapter 9: College Motivation — Why It Belongs in the College Readiness Scorecard

Dimension D09 | Weight: 5% | Measurement: 18-question Likert survey — Eagerness for College (6 questions, 40%) + School Enthusiasm (6 questions, 30%) + Academic/Social Balance (6 questions, 30%)

Introduction: Why Motivation Deserves Its Own Dimension

When educators and families ask whether a student is "ready for college," they most naturally reach for academic evidence – GPA, test scores, course rigor. These measures are indispensable. But a student who earns a 3.9 GPA while privately dreading college attendance, or who excels on the SAT but views higher education as someone else's idea, faces a meaningfully different set of risks than her equally credentialed peer who is genuinely hungry to be there.

The College Readiness Scorecard addresses this gap through Dimension 9 (D09): College Motivation. Weighted at 5% of the composite score, D09 measures three interrelated constructs via an 18-question Likert survey: Eagerness for College (6 questions, 40% of the dimension), School Enthusiasm (6 questions, 30%), and Academic/Social Balance (6 questions, 30%). This chapter explains why each construct was chosen, why the weights are set the way they are, and how the broader scorecard framework positions motivation as a real – if modest – predictor of college success.

The Scorecard's Multi-Dimensional Philosophy

The College Readiness Scorecard was built on a well-documented premise: no single metric reliably predicts college success, and the most authoritative actors in higher education have long acted accordingly. The NACAC *State of College Admission* surveys confirm that 75% of colleges rate overall GPA as "considerably important," but so too do curriculum strength

(62%), and – in declining but still meaningful numbers – test scores (46%). College admission offices do not evaluate students through a single lens, and neither should a college readiness instrument.

The scorecard operationalizes this philosophy by distributing weight across ten dimensions. Academic Achievement (D01, 20%), Standardized Testing (D02, 15%), and Course Rigor (D03, 10%) together account for 45% of the composite – reflecting the empirically validated primacy of academic preparation. The remaining 55% is allocated to dimensions that capture what grades and scores alone cannot see: personal character, extracurricular depth, college knowledge, self-awareness, academic skills, motivation, and financial preparedness.

This structure reflects a broader consensus in education research. David Conley's *Four Keys to College Readiness* framework identifies Academic Behaviors – which include motivation-related constructs like goal-setting, persistence, and self-monitoring – as a distinct domain independent of raw academic achievement. William Sedlacek's Noncognitive Questionnaire (NCQ), validated over decades and used to select 11,000+ Gates Millennium Scholars, identifies "long-range goals" and "positive self-concept" as noncognitive variables that predict college persistence for diverse student populations, achieving 97% first-year retention and 78% five-year graduation in program data. A large-scale data-reduction study by Jim Soland at NWEA, using the NELS:88 dataset of 12,144 students and over 1,000 candidate predictor variables, found that "postsecondary aspirations and expectations" emerged as one of only four core constructs (alongside academic preparation, teacher perceptions, and socioeconomic status) that together predicted college enrollment and persistence with nearly 90% accuracy.

In that NWEA analysis, the single strongest predictor in the entire dataset – yielding the highest lasso coefficient (1.501) and the highest factor loading (0.59) – was the variable measuring "how far in school the student thinks they will go." That is a motivation construct. It is the empirical foundation on which D09 rests.

Why 5%? Calibrating the Weight of Motivation

Five percent is a deliberate and defensible choice. It is small enough to prevent a motivated-but-underprepared student from masking genuine academic gaps; it is large enough to meaningfully differentiate students at the margins of readiness and to signal that motivation is a real predictor – not a courtesy addition.

The University of Chicago Consortium on Chicago School Research (Nagaoka et al., 2012) offers a useful conceptual map. Their framework identifies five categories of noncognitive factors: academic behaviors, academic perseverance, social skills, learning strategies, and academic mindsets. Motivation belongs primarily to the mindsets category, which the CCSR describes as operating *through* mediating behaviors rather than producing outcomes directly. Students who believe their work has value – who are intrinsically eager to attend college – are more likely to exhibit the academic behaviors and perseverance that actually drive GPA and retention. Motivation is foundational but indirect, which justifies a smaller weight than direct academic measures.

The research literature supports this positioning. A 2024 meta-analysis by Walker et al. (Oregon State University), synthesizing 94 outcomes from 52 studies, found that performance self-efficacy correlates with academic performance at r = 0.59 in face-to-face settings – a strong relationship. But the overall variance explained by the motivational variables examined was 14%, compared to substantially higher explanatory power from academic preparation measures. A study by Norvilitis, Reid, and O'Quin (2022) across two cohorts (N = 523 and N = 385) found that intrinsic and extrinsic motivation positively predicted college GPA, student-university match, and first-to-second-year retention, while amotivation negatively predicted all three. The effect was real and statistically significant. It was also modest enough to be appropriately sized at 5% rather than 15% or 20%.

Eagerness for College: The Anchor Construct (40%)

The most heavily weighted sub-construct within D09 is Eagerness for College – the student's intrinsic motivation to attend, their genuine anticipation of the transition, and their sense that college is something they want rather than something being done to them.

The research basis for this prioritization is strong. Guiffrida, Lynch, Wall, and Abel studied 2,500 college students across a community college and a liberal arts college and found that students attending to fulfill autonomy needs (studying areas of genuine interest) and competence needs (testing their abilities) had higher GPAs and stronger persistence intentions than students driven by external pressures. A follow-up study focused on low-income first-year students found that students with higher autonomous motivation had significantly lower attrition rates, and that students with controlling motivations – attending due to parental pressure or financial security alone – had a 2.3 times higher attrition rate.

Francis Stage's foundational study (1989) within the Tinto persistence framework found that students' initial motivational orientation functioned as a moderating variable: it shaped how academic and social integration experiences translated into persistence or withdrawal. Eagerness for college is not merely predictive in isolation – it conditions how the entire college experience unfolds.

The Antonoff College Match Self Survey, a widely used practitioner instrument deployed by college counselors nationally, independently validates this construct structure. Antonoff's 80-question instrument includes an Eagerness for College category that measures students' intrinsic motivation and anticipation for attending, with explicit interpretation guidance: high scorers are ready for immediate transition; low scorers may benefit from gap years or alternative pathways. The fact that a practitioner tool built without knowledge of this scorecard identifies the identical construct confirms D09's face validity.

The Florida Atlantic University THRIVE program summarizes the mechanism: "Intrinsically motivated students show higher achievement levels, lower anxiety levels, and more competence in learning than extrinsically motivated students. They don't just study to get a good grade – they study because they genuinely enjoy the subject they are learning." Eagerness for college is not a soft proxy. It is an operationalization of intrinsic motivation, which has a well-documented pathway to academic and persistence outcomes.

School Enthusiasm: Secondary Signal (30%)

School enthusiasm – a student's affective engagement with learning, intellectual curiosity, and positive orientation toward school as an environment – is the second sub-construct of D09, weighted at 30%.

The supporting literature is consistent but shows weaker effect sizes than direct college motivation. Wang and Fredricks (2014) conducted a longitudinal study of 1,103 adolescents examining the reciprocal relationship between school engagement and academic outcomes. They found behavioral engagement (enthusiasm, participation) positively predicted achievement with moderate effect sizes (B = 0.25-0.35), while emotional engagement (school enthusiasm) showed weaker but significant prediction of college enrollment intentions (B = 0.18). Critically, students with high school enthusiasm but low college eagerness had only modest college persistence rates – underscoring that enthusiasm without directed motivation is not sufficient, which justifies school enthusiasm's secondary status within the dimension.

The Antonoff College Match Self Survey also includes a School Enthusiasm category, independently defined as engagement with academic learning, intellectual curiosity, and positive attitudes toward school as a learning environment – precisely the construct D09 is measuring. This practitioner convergence, developed independently in a different professional context, reinforces construct validity.

A 30% weight within D09 – meaning 1.5% of the total composite – is appropriately modest. School enthusiasm predicts college outcomes through pathways similar to motivation (behavioral engagement, academic investment), but with weaker direct effect sizes. It belongs in the dimension as a signal that a student's relationship with learning is healthy and likely to persist into college, without dominating the motivational profile.

Academic/Social Balance: The Integration Dimension (30%)

The third sub-construct of D09 measures a student's readiness to navigate the dual demands of college life: academic rigor on one side and social integration on the other. This is the least intuitive of the three components, but it is grounded in one of the most replicated models in higher education research.

Vincent Tinto's academic-social integration model has framed college persistence research for five decades. His core insight is that students who successfully integrate into both the academic culture and the social community of their institution are dramatically more likely to persist to graduation. Social isolation and academic disengagement each independently predict dropout; the worst outcomes occur when both are present. A 2009 study by Severiens and Schmidt, testing Tinto's model in a problem-based learning context with 305 first-year students, confirmed that formal social integration positively affected study progress independently of academic integration.

The balance framing – rather than simply measuring social integration – is informed by a second finding from the Guiffrida et al. studies: students whose primary motivation was

peer relationships (rather than learning or goal pursuit) showed lower GPAs, with more pronounced effects for male students. Social connection is protective when it coexists with academic engagement, but it becomes a risk factor when it crowds out academic motivation. D09's Academic/Social Balance construct asks students to reflect on where they stand on this continuum, providing a pre-enrollment signal of whether they are likely to find productive equilibrium.

The Antonoff instrument includes this same category – Academic/Social Balance – with explicit language about helping students assess whether they prioritize academics, social life, or a balanced approach, and what implications that profile has for college environment fit. A student who accurately understands their own academic/social orientation is better positioned to select an appropriate college environment and to self-regulate once enrolled.

Survey Design: Why 4-Point Likert, Why 18 Questions

D09 is measured entirely through a self-report Likert survey – a methodological choice that requires justification.

Hardin, Gibbons, Cook, Sexton, and Bagwell (2021) validated a 14-item college-going self-efficacy instrument using a 4-point Likert scale (not at all sure to very sure) with rural Appalachian high school students across two independent samples. The instrument demonstrated good reliability, construct validity, and measurement equivalence across gender and first-generation college status. Their successful validation confirms that 4-point Likert scales can reliably capture college readiness self-perceptions in the target population – high school students from diverse backgrounds – without requiring a neutral midpoint that would invite non-responses.

The scorecard uses a 4-point response format (Strongly Agree / Agree / Lean Disagree / Disagree) mapped to values of 9, 6, 3, and 0, then normalized to the 1-9 scale using the formula 1 + (sum/54) x 8. This is a standard min-max linear normalization, validated in psychometric literature as appropriate for converting Likert sums to bounded scales while preserving rank-order and interval properties. The six questions per sub-construct follow standard scale development guidelines for achieving adequate reliability.

The survey's limitations are acknowledged: self-report measures are subject to social desirability bias, and students who understand what answers appear "college ready" can game the instrument. This is a known constraint of all noncognitive self-assessment tools. It is also the reason D09 is weighted at only 5% – the composite score does not hinge on survey responses, and the dimensions where gaming is most likely (D07-D09) collectively represent only 20% of the total score. The appropriate response to measurement imperfection is not to exclude the construct but to size it appropriately.

Motivation in the Equity Context

The College Readiness Scorecard is designed with equity as a structuring concern, and D09 has a particular role to play in that context.

Guiffrida et al.'s research on low-income, first-generation students found that autonomous motivation (intrinsic eagerness for college) buffered against attrition risk even when students faced academic difficulties. In a student population where external stressors – financial strain, family obligations, first-generation anxiety – are elevated, intrinsic motivation functions as a protective factor. Measuring it acknowledges that two students with identical GPAs may face meaningfully different persistence risks depending on their internal orientation toward college attendance.

The Sedlacek NCQ framework makes a broader equity claim: noncognitive variables, including motivation-related constructs like long-range goals and positive self-concept, are especially important for students of color, first-generation students, and others whose academic credentials may not fully capture their potential. The Gates Millennium Scholars program, which weights noncognitive variables at 80% of selection criteria, achieved 97% first-year retention and 87% five-year retention for a population heavily composed of underrepresented students – outcomes significantly above national averages.

D09 does not go as far as the Gates program in elevating noncognitive factors. At 5%, it is a signal, not a gatekeeper. But it is a signal that matters disproportionately for students whose college readiness story is complicated by circumstances outside their academic record.

Conclusion: Small Weight, Real Meaning

D09 College Motivation carries 5% of the College Readiness Scorecard composite – the smallest weight among the ten dimensions, tied with Financial Preparedness (D10). That weight is appropriate. Motivation is real, measurable, and consequential, but it operates through intermediary pathways rather than directly driving GPA or graduation rates. It belongs in a comprehensive readiness instrument as a modest but meaningful signal.

The research base supporting D09 is coherent across multiple methodological traditions: self-determination theory (Guiffrida et al.), meta-analytic evidence (Walker et al., Oregon State), longitudinal developmental research (Wang & Fredricks), Tinto-framework persistence studies (Severiens & Schmidt, Stage), and psychometric instrument validation (Hardin et al., Antonoff). The three sub-constructs – Eagerness for College, School Enthusiasm, and Academic/Social Balance – are each independently validated in the literature and are jointly confirmed by the Antonoff College Match Self Survey, a practitioner instrument that converged on the identical construct structure without reference to this scorecard.

At its core, D09 asks a question that grades and test scores cannot: Does this student actually want to be here? That question, while humble in its weighting, deserves to be asked.

Key Sources Cited: Guiffrida et al.; Norvilitis, Reid & O'Quin (2022); Walker et al. (2024); Wang & Fredricks (2014); Severiens & Schmidt (2009); Stage (1989); Hardin et al. (2021); Nagaoka et al. (2012); Antonoff (2022); Soland (2017); Sedlacek (2011); Conley (2007); NACAC (2019).

Chapter 10: Financial Preparedness — Why It Belongs in Any Serious College Readiness Scorecard

Dimension D10 | Weight: 5% | Measurement: Three-item behavioral checklist — (1) Used a FAFSA estimator, (2) Discussed college costs with family, (3) Researched scholarships

Introduction: The Student Who Got Sent Home

A high school senior was accepted to college, filled out the financial aid forms, and showed up to start classes. Weeks into the semester, administrators realized he owed thousands of dollars that his financial aid offer did not cover. He had not understood the gap between what the college promised and what it would actually cost his family. The college sent him home.

This story – drawn from real case documentation in college access literature – is not unusual. It represents a failure of financial preparedness that no transcript or test score could have predicted or prevented. Academic readiness is necessary for college success. It is not sufficient.

The College Readiness Scorecard was designed to close that gap. Built on 288 research entries compiled across 11 dimensions, the Scorecard uses a weighted 10-dimension composite to evaluate students on the full range of competencies that research shows matter for college access, enrollment, and persistence. Dimension 10 – Financial Preparedness – carries 5% of the composite weight. This chapter explains why that weight exists, what the evidence base for it looks like, and why students, counselors, and institutions should take it seriously.

Section 1: Why a Single Score Has Never Been Enough

The case for any multidimensional college readiness instrument begins with a simple fact: no single metric – not GPA, not test scores, not any one factor – adequately captures whether a student is ready for college success.

NACAC national surveys, conducted annually over more than a decade, document this clearly. In 2019, 75% of colleges rated overall GPA as considerably important in admission decisions; 73% rated grades in college-prep courses; 62% rated curriculum strength; and 46% rated test scores. These are not competing measures – they are complementary signals, each adding independent information. As NACAC has stated plainly: “There is no definite plan or specific combination of factors that will guarantee a student admission to their preferred institution.”

Research on predictive validity extends the point beyond admissions. Lotkowski and colleagues (2004), in a landmark ACT meta-analysis, found that academic factors like GPA and test scores account for only about 25% of variance in first-year college GPA. Psychosocial and noncognitive factors – motivation, college knowledge, financial literacy, self-management skills – add 13% incremental predictive validity beyond academics alone. Students who look identical on their transcripts can have wildly different outcomes in

college, and those differences are frequently explained by the factors that transcripts do not measure.

David Conley's Four Keys framework, one of the most influential college readiness models in American education, identifies four distinct dimensions of readiness: cognitive strategies, content knowledge, self-management skills, and knowledge about postsecondary education. That fourth key – knowledge about postsecondary institutions, including how to secure financial aid – is explicitly separated from academic preparation. As Conley wrote: "Choosing a college, applying, securing financial aid, and then adjusting to college life require a tremendous amount of specialized knowledge." Financial readiness is not a footnote to academic readiness. It is its own domain, requiring its own assessment.

Section 2: The Scope of the Financial Readiness Gap

Before examining what D10 measures, it is worth establishing the scale of the problem it addresses.

A 2024 Gallup and Lumina Foundation survey of more than 14,000 U.S. adults who had stopped out or never enrolled in college found that 85% cited the cost of a degree program as an important reason for not being currently enrolled. Among students who were enrolled at the time of the survey, 31% had considered stopping out due to cost, and 35% had considered stopping out in the preceding six months. Among students considering leaving, 54% reported struggling to pay monthly bills.

The problem is compounded by financial literacy deficits that begin before students ever arrive at college. A nationally representative EverFi/NASFAA survey of more than 100,000 incoming college students at 410+ institutions found that most respondents could answer only 2 of 6 basic financial literacy questions correctly. Just 40% of four-year students had ever taken a personal finance course. And while 60% of students expected to take out loans, only 15% felt they had the education, information, and resources to be able to repay them.

For first-generation students, the gaps are more acute. Research by Dissen and Tome (2024) in *Metropolitan Universities* found that 44% of high school juniors did not know what FAFSA was, and that first-generation students showed significantly lower financial self-efficacy scores and higher financial strain scores than their continuing-generation peers. Falcon (2015) documented that many first-generation college students "leave college so they can work more hours to support themselves or their family, or because college is not economically feasible."

The ACT's 2023 research report on college-bound students quantifies the disparity precisely. On a 4-point preparedness scale:

- Academic preparedness: 3.27
- Social preparedness: 2.87
- Personal preparedness: 2.87
- Financial preparedness: 2.33

Financial preparedness scored 0.54 to 0.94 points lower than every other readiness domain. Only 47% of students reported being very or mostly prepared to pay tuition, compared to 86% who reported being academically prepared. Between 21% and 26% of college-bound students reported being slightly prepared or not prepared at all for financial tasks in their first year.

This is the readiness gap that D10 exists to surface.

Section 3: What D10 Measures and Why Each Item Matters

D10 is structured as a three-item behavioral checklist: (1) used a FAFSA estimator to understand likely aid, (2) discussed the cost of college with family, and (3) researched scholarships. Each item reflects a distinct, evidence-backed preparedness behavior.

Item 1: Used a FAFSA Estimator

FAFSA completion is among the most consistently validated behavioral predictors of college enrollment in the research literature. Page and Scott-Clayton (2016) documented that only 58% of high school graduates who intended to enroll in college actually completed the FAFSA, leaving 42% who did not complete the form despite planning to enroll. Among non-completers, 13% would have been eligible for federal Pell Grants – money they never received because they never completed the paperwork.

The barriers are well-documented. Dynarski and Scott-Clayton (2006) found that the FAFSA requests 127 pieces of information across seven pages, and that families spend an estimated 8 hours completing it – a burden that falls disproportionately on low-income families with less access to professional assistance. Simplifying aid application processes increases college enrollment by 8 percentage points among low-income students. D10 measures use of a FAFSA estimator – not the full form – because the estimator provides critical early financial planning information without the complexity burden. Researchers at Penn AHEAD found that access to cost information early in the process, without having to complete the FAFSA, is "especially important to low-income students, first-generation college students, and individuals who do not have ready access to college and financial aid counselors."

Net price calculator research validates the predictive power of this behavior. Anthony, Page, and Seldin (2016) found in a peer-reviewed study that net price calculator estimates of grant aid "correlate highly with actual grant aid on average," confirming that students who use these tools gain meaningful, actionable information. College Board research published in 2025 concluded that access to accurate net price information as early as junior year "can shape decisions about where students apply – and, ultimately, where they enroll."

Fitzpatrick and Schneider (2016), using the nationally representative HSLS dataset of 23,000+ students, found that counselor meetings increase FAFSA submission intent odds by 45% overall – and more than double the odds for students in poverty (OR 2.19, $p<.01$). Zinth (2014) documented that students at schools with college coaches were 17% more likely to complete the FAFSA, 13% more likely to attend college, and 24% more likely to

enroll at a four-year institution. The behavioral chain is clear: financial aid awareness predicts FAFSA engagement, which predicts enrollment.

Item 2: Discussed College Costs with Family

Family cost conversations are documented predictors of financial aid behavior and enrollment outcomes. The 2019 Sallie Mae/Ipsos national study found that 71% of families report discussing how to pay for college before a student applies. But the differences between families who have this conversation early and those who do not are significant:

- Families who discuss costs early are more likely to use scholarships (90% vs. 73%)
- More likely to use grants (81% vs. 68%)
- More likely to complete the FAFSA (87% vs. 62%)
- Less likely to be very concerned about affording college (28% vs. 42%)

NACAC's Guide to the College Admission Process frames the timing clearly: "Have a conversation with your family about the cost of higher education at the beginning of your college search to ease stress and frustration later in the process." It adds: "Financial planning for college can begin before signing a loan agreement, and even before filling out the FAFSA."

The substantive content of these conversations matters. Families that understand the difference between direct costs (tuition, fees, room and board) and indirect costs (books, supplies, transportation, personal expenses) make significantly different decisions than families focused only on sticker price. MDRC's research on college match programs documented this directly: students who understand net price calculations are more likely to apply to selective colleges that sometimes provide more generous aid to low-income students than nearby community colleges. "Selective colleges often provide more generous financial aid to low-income students, making them sometimes less expensive than community colleges," MDRC found. Students who have not had family cost discussions often never learn this – and systematically undermatch as a result.

King (2004) found that 40% of FAFSA non-filers assumed their family could pay without investigating aid options, and 24% believed they would not qualify. These are not informed decisions. They are assumptions made in the absence of the conversations that D10 measures. Students who check this box demonstrate they have moved from assumptions to evidence-based financial planning.

Item 3: Researched Scholarships

Scholarship research behavior is the third D10 item because research shows it measures something beyond money-finding. It captures whether a student has developed the proactive resource navigation skills that predict successful college transition.

Research for Action's 2025 study of last-dollar scholarship programs found that financial assistance alone is insufficient to promote successful degree completion – students need integrated support structures. As the researchers noted: "For holistic student supports to be

effective, it is not enough to only offer services and resources; well-designed programs must encourage students to use the supports available to them." Tennessee and Philadelphia data demonstrated that scholarship program engagement correlates with higher rates of persistence, retention, and degree completion. Students who research and apply for scholarships demonstrate the help-seeking behavior and system navigation skill that predicts college success independently of whether they receive an award.

For first-generation students, this behavior is particularly meaningful. Hagler and colleagues (2023) found that first-generation students experience help-seeking as having "hidden costs" – time burden, embarrassment, shame, the feeling of burdening others – and that many prefer self-reliance and wait until needs become acute before seeking assistance. Students who nonetheless research scholarships have overcome this documented barrier. They have demonstrated the proactive engagement with support systems that characterizes college-ready behavior.

College Board's BigFuture Scholarship data reinforces the point: 60% of BigFuture scholarship recipients come from low-income backgrounds, and nearly half are first-generation college students. College Board explicitly ties scholarship eligibility to completion of structured planning actions – career exploration, financial aid awareness, college list development – because completing these milestones predicts follow-through to enrollment.

Section 4: Why 5% Is the Right Weight

Critics of the Scorecard's weighting scheme sometimes ask: if financial barriers are the #1 reason adults cite for not enrolling in college, why does D10 carry only 5% of the composite score?

The answer lies in a distinction between the prevalence of financial barriers across the general adult population and their variance among actively college-bound students completing the Scorecard.

The Gallup/Lumina data (85% cite cost as a barrier) comes from adults who have already not enrolled. Among students actively participating in a college readiness process, the proportion facing acute financial barriers is smaller – the ACT research found that 47% of college-bound students feel financially prepared, meaning the majority are at or above baseline readiness. The 5% weight reflects that while financial preparedness is critical for students who lack it, it does not exhibit the same universal variance as academic achievement (D1 at 20%), which predicts outcomes for all college-going students regardless of background.

The ACT's evidence validates this calibration. Students who participated in 9 or more college preparation activities had average financial preparedness scores of 2.8 on the 4-point scale – 0.9 points higher than students who participated in only 0-2 activities (1.9). This 47% improvement in financial preparedness scores demonstrates that the D10 items are genuinely responsive to intervention and capture real, measurable variance in readiness.

The Inceptia research adds context on consequences. Among community college students, 19% reported that financial stress had caused them to consider dropping out, compared to fewer than 7% of students at four-year institutions. And the Inceptia data on behavioral goal-setting shows the stakes clearly: students with written financial goals were dramatically more likely to feel capable of handling their finances (85% vs. 47%), more likely to check their credit reports (71% vs. 12%), and more likely to contribute to savings regularly (52% vs. 30%). Financial preparedness predicts not just enrollment but persistence – and D10's 5% weight is calibrated to reflect its narrower but critically important role in the composite.

Section 5: D10 as Distinct from D6

A natural question arises: since D6 (College Knowledge, 8%) covers college planning behaviors like campus visits, counselor meetings, and college list development, why does financial preparation deserve its own dimension?

The research literature consistently treats financial preparation as a distinct domain within college readiness. Stanford's CRIS Framework (Gardner Center, 2014) explicitly identifies financial knowledge as a component of college readiness operating at three levels: individual indicators (FAFSA completion, aid application behavior), setting-level indicators (counseling resource access), and system-level indicators (district resources for financial aid support). The framework emphasizes that financial preparation requires specific behavioral engagement distinct from general college-going awareness.

Conley's Four Keys research identifies "securing financial aid" as a distinct component of postsecondary knowledge requiring students to understand federal financial aid programs, how and when to complete forms, and how to navigate institutional processes. His research found that exemplary high schools with high enrollment rates provided dedicated help with financial aid forms as part of their preparation programs – separately from general college counseling. He found that first-generation students, students from immigrant families, and low-income students "require a much more intentional, comprehensive program of preparation" specifically including financial aid knowledge.

The practical distinction matters for Scorecard interpretation. A student may have visited four campuses, built a balanced college list, and met with their counselor three times – and still not have used a FAFSA estimator, discussed costs with their family, or researched scholarships. D6 and D10 measure different things, and the combination provides a fuller picture of readiness than either dimension alone.

Section 6: What the Evidence Tells Counselors and Students to Do

The evidence base for D10 translates directly into practical guidance.

Start the conversation early. NACAC recommends families discuss college costs at the beginning of the college search. McKibben and NCAN recommend beginning financial aid conversations in middle school if possible. Students who have never discussed college costs

with their families are not behind on paperwork – they are behind on a multi-year preparation process.

Use the estimator before building a college list. College Board's 2025 research found that students who access accurate net price information as early as junior year make different application decisions than those who wait. Students who assume selective or private colleges are unaffordable without checking net prices systematically undermatch – choosing less selective colleges that admit them at lower graduation rates and with less generous aid packages.

Treat scholarship research as college readiness, not a side project. Research on scholarship program participation shows that the act of engaging with scholarship applications builds the same skills that predict college persistence. It is not just about the money – it is about the student's demonstrated ability to navigate complex systems on their own behalf.

Close the information gap for first-generation and low-income students. Documented disparities in FAFSA awareness, financial self-efficacy, and help-seeking behavior cluster systematically among first-generation, low-income, and underrepresented students. For these students, a low D10 score is a direct signal that targeted intervention is needed.

Conclusion: The Full Picture of College Readiness

The College Readiness Scorecard rests on a premise supported by decades of research: college readiness is multidimensional, and any assessment that measures only academics is measuring an incomplete picture. The student who got sent home mid-semester had the grades to get admitted and the ambition to show up. What he lacked was the financial preparedness to understand what was expected of him – and no one had thought to measure that.

D10 exists to ensure that gap is visible before it becomes irreversible.

At 5% of the composite score, Financial Preparedness is the Scorecard's most targeted dimension – three behavioral items measuring whether a student has taken the minimum necessary steps toward evidence-based financial planning. The evidence for each item is robust: FAFSA estimator use predicts enrollment intent (counselor meetings more than double FAFSA intent for students in poverty); family cost discussions predict financial aid access (87% vs. 62% FAFSA completion rates); scholarship research predicts persistence (correlation with completion rates documented across multiple program evaluations).

But the deeper case for D10 is structural. College readiness assessment that ignores financial preparation is not just incomplete – it is inequitable. The students who score lowest on financial preparedness are disproportionately the students who most need support: first-generation, low-income, and underrepresented students who lack the family experience and social capital to navigate financial aid systems on their own. A scorecard that surfaces their financial readiness gaps creates an opportunity to intervene. A scorecard

that ignores those gaps ensures the student who showed up to college without understanding his bill remains invisible – right up until the day he is sent home.

Key Sources Referenced: ACT (Schnieders et al., 2023); Anthony, Page & Seldin (2016); College Board (2025); Conley (2008); Dissen & Tome (2024); Dynarski & Scott-Clayton (2006); EverFi/NASFAA (2018); Falcon (2015); Fitzpatrick & Schneider (2016); Gallup & Lumina Foundation (2024); Gardner Center, Stanford (2014); Hagler et al. (2023); Inceptia/Trombitas & Lannan (2013); King (2004); Lotkowski et al. (2004); McKibben (2017); MDRC/Byndloss et al. (2015); NACAC (2015, 2019); Page & Scott-Clayton (2016); Penn AHEAD/Perna et al. (2019); Research for Action (2025); Sallie Mae & Ipsos (2019); Zinth (2014).

CONSOLIDATED REFERENCES

The following is a comprehensive list of all research sources cited across this compendium. Sources are listed alphabetically by first author.

ACT. (2018). *The Condition of College and Career Readiness.*

Allen, J. & Robbins, S. (2010). Effects of Interest-Major Congruence, Motivation, and Academic Performance on Timely Degree Attainment. *Journal of Counseling Psychology*.

Allensworth, E.M. & Clark, K. (2020). High School GPAs and ACT Scores as Predictors of College Completion. *Educational Researcher*. Lumina Foundation.

Anthony, P., Page, L., & Seldin, A. (2016). In the Right Ballpark? Assessing the Accuracy of Net Price Calculators. *Journal of Student Financial Aid.*

Antonoff, S. R. (2022). *College Match Self Survey*. Octameron Associates.

Astin, A.W. & Astin, H.S. (2000). *Leadership Reconsidered: Engaging Higher Education in Social Change*. W.K. Kellogg Foundation.

Avery, C. (2010). The Effects of College Counseling on High-Achieving, Low-Income Students. NBER Working Paper.

Beard, J., Hsu, Y., Ewing, M., & Godfrey, K. (2023). AP Coursework and College Outcomes. College Board Research.

Bekkering, E. & Ward, T. (2021). The Relationship Between Attendance, Participation, and Academic Performance. *Journal of Higher Education Theory and Practice*.

Bettinger, E., Cunha, J.M., Lichand, G., & Madeira, R. (2017). *Applying to College: Do Information Deficits Lower the Likelihood of College Applications?*

Blackwell, L.S., Trzesniewski, K.H., & Dweck, C.S. (2007). Implicit theories of intelligence predict achievement across an adolescent transition. *Child Development*.

Bound, J. et al. College Board research on SAT score bands and graduation rates.

Bowen, W.G., Chingos, M.M. & McPherson, M.S. (2009). *Crossing the Finish Line*. Princeton University Press.

Boyer, S. & Sedlacek, W. (1987). Noncognitive Predictors of Academic Success. *Journal of College Student Personnel.*

Buzzetto-Hollywood, N.A. & Mitchell, B.C. (2019). Grit and persistence: Findings from a longitudinal study. *Issues in Informing Science and Information Technology, 16*, 377-391.

CASEL. (2020). What Is the CASEL Framework? casel.org.

Casillas, A. et al. (2012). ACT Engage research on academic behaviors in middle school.

Çetin, B. (2015). Academic motivation and self-regulated learning in predicting academic achievement. *Journal of Education and Training Studies.*

Chen, P. et al. (2017). Strategic Resource Use for Learning: A Self-Administered Intervention. *Psychological Science*. Stanford University.

College Board. (2013). Validity of the SAT for Predicting First-Year Grades (n=148,768).

College Board. (2013). Does College Readiness Translate to College Completion? (n=898,517).

College Board. (2019-2020). AP Students in College synthesis.

College Board. (2025). Net Price Transparency Strengthens Yield and Fit.

CollegeVine & Solomon Admissions. (2025). What is the Academic Index?

Community College Research Center (CCRC). (2025). Dual Enrollment and College Outcomes.

Conley, D.T. (2007). *Toward a More Comprehensive Conception of College Readiness.* Bill & Melinda Gates Foundation.

Conley, D.T. (2008). What Makes a Student College Ready? *Educational Leadership*.

Conley, Simons, et al. (2025). Building social-emotional skills in a college career and life planning course. *Social and Emotional Learning: Research, Practice, and Policy, 5.*

Corporation for National and Community Service. (2007). *The Health Benefits of Volunteering: A Review of Recent Research.*

Credé, M. et al. (2010). Class Attendance in College: A Meta-Analytic Review. *Review of Educational Research.*

Dissen, R. & Tome, R. (2024). Finances and Future Health. *Metropolitan Universities.*

Dr. Accrocco / Rice University Center for College Readiness. (2024). *Undergraduate Admissions Process – Complete Course Transcripts*.

Duckworth, A.L., Peterson, C., Matthews, M.D., & Kelly, D.R. (2007). Grit: Perseverance and passion for long-term goals. *Journal of Personality and Social Psychology, 92*(6), 1087-1101.

Dweck, C.S. (2006). *Mindset: The New Psychology of Success*. Random House.

Dynarski, S. & Scott-Clayton, J. (2006). The Cost of Complexity in Federal Student Aid. *National Tax Journal*.

Eccles, J.S., Barber, B.L., Stone, M. & Hunt, J. (2003). Extracurricular Activities and Adolescent Development. *Journal of Social Issues, 59*(4).

Ellis, J. & Helaire, L. (2017). Self-Regulated Learning and College Readiness. *Journal of College Student Retention*.

EverFi/NASFAA. (2018). Survey: Incoming College Students Struggle With Basic Financial Literacy.

Falcon, L. (2015). Breaking Down Barriers: First-Generation College Students and College Success.

Fitzpatrick, D. & Schneider, B. (2016). Linking Counselor Activities and Students College Readiness. *SREE*.

Fredricks, J.A. & Eccles, J.S. (2006). Extracurricular Involvement and Adolescent Adjustment. *Applied Developmental Science, 10*(3).

Freeman, S. et al. (2014). Active Learning Increases Examination Performance in STEM. *Proceedings of the National Academy of Sciences*.

Gaggioli, A. et al. (2025). *Assessing the Reliability and Validity of Large Language Models for Automated Assessment of Student Writing*. arXiv.

Gallup & Lumina Foundation. (2024). The Biggest Barriers to Higher Ed Enrollment Are Cost and Lack of Financial Aid.

Gardner Center, Stanford. (2014). A New Framework for Promoting College Readiness (CRIS Framework).

Garland, R. (1991). The Mid-Point on a Rating Scale: Is it Desirable? *Marketing Bulletin*.

Gershenson, S. (2018). *Grade Inflation in High Schools (2005-2016)*. Thomas B. Fordham Institute.

Gettenberg, A. (2022). An analysis of the relation between vocabulary and emotion regulation. St. John's University Dissertation.

Giani, M., Alexander, C. & Reyes, P. (2014). Exploring Variation in the Impact of Dual-Credit Coursework. *New Directions for Community Colleges*.

Guiffrida, D., Lynch, M., Wall, A., & Abel, D. Do Reasons for Attending College Affect Academic Outcomes? *Journal of College Student Development*.

Hagler, M. et al. (2023). A qualitative thematic analysis of first-generation college students help-seeking. *Journal of College Student Retention*.

Hardin, E.E. et al. (2021). Development and validation of a short form of the College-Going Self-Efficacy Scale. *Journal of Career Assessment, 29*(2), 303-318.

Harvard Graduate School of Education, Making Caring Common Project. (2016). *Turning the Tide: Inspiring Concern for Others and the Common Good through College Admissions*.

Harvard Graduate School of Education, Making Caring Common Project. (2020). *Character Assessment in College Admission Guide Overview*.

Harvard Graduate School of Education, Making Caring Common Project. (2024). *From Grit to Growth Mindset: Defining Character Traits in Admissions*.

Harvard Law Review. (2021). *Students for Fair Admissions v. President & Fellows of Harvard College,* Vol. 134.

Hu, X. (2025). Dual Enrollment and Merit Scholarships. *Journal of Student Financial Aid*.

IES Practice Guide. (2009). *Helping Students Navigate the Path to College*.

Inceptia/Trombitas & Lannan. (2013). The Intersection of Student Financial Literacy and Career Readiness.

Kim, S. et al. (2022). The effects of growth mindset on college persistence and completion. Education Analytics Working Paper.

King, J.E. (2004). Missed Opportunities: Students Who Do Not Apply for Financial Aid. ACE.

Klasik, D. (2012). The College Application Gauntlet. *Research in Higher Education*.

Klasik, D. (2013). The Effect of Mandatory State ACT Testing. *Educational Policy*.

Klasik, D., Blagg, K., & Pekor, Z. (2019). *Direct Admissions: Evidence from Idaho*.

Komarraju, M., Ramsey, A. & Rinella, V. (2013). Cognitive and Non-Cognitive Predictors of College Readiness. *Learning and Individual Differences*.

Lee, J., Bahena, S. & Nichols, A. (2025). AP Coursework for Underprepared Students. *Education Policy Analysis Archives*.

Lira, B. et al. (2023). Using artificial intelligence to assess personal qualities in college admissions. *Science Advances*. (N=309,594)

Liu, Y., Abu Bakar, N.I. & Xu, X. (2025). Self-Regulated Learning Strategies and Academic Achievement: A Systematic Review.

Lotkowski, V.A. et al. (2004). The Role of Academic and Non-Academic Factors in Improving College Retention. ACT.

MacCann, C. et al. (2019). Emotional intelligence predicts academic performance: A meta-analysis. *Psychological Bulletin*.

Marini, J.P., Westrick, P.A., & Shaw, E.J. (2024). SAT Score Relationships with College GPA: First-Year through Fourth-Year. College Board Research Report.

Maruyama, G. (2024). How much do we need college admission tests? *Analyses of Social Issues and Public Policy*.

Mattern, K.D. & Patterson, B.F. (2014). Synthesis of Recent SAT Validity Findings. College Board Research Report 2014-3.

McAlister, S. & Mevs, P. (2012). *College Readiness: A Guide to the Field*. Annenberg Institute at Brown University.

McKibben, S. (2017). Engaging Parents in the Financial Aid Process. NCAN.

MDRC/Byndloss et al. (2015). In Search of a Match.

Michaelides, M. & Durkee, E. (2021). Self-Regulation and Academic Achievement. *Educational Psychology*.

NACAC. (2010, 2013, 2015, 2019, 2023). *State of College Admission Reports*.

NACAC. (2015). *Guide to the College Admission Process*.

NACAC. (2016). *Fundamentals of College Admission Counseling*. Chapters 3 and 16.

NACAC. (2022). *Character and the College Admission Process Research Brief*.

NACAC. (2024). *Assessing Character in the College Admission Process*.

NACAC/Common Application. *Counselor Handout: Guiding Students Through the Common Application*.

Nagaoka, J. et al. (2012). Readiness for College: The Role of Noncognitive Factors and Context. University of Chicago Consortium on Chicago School Research.

Nankervis, B. (2011). Gender Inequities in University Admission Due to the Differential Validity of the SAT. *NACAC Journal of College Admission*.

Niche. (2022). *2022 Senior Enrollment Survey: The New Normal*.

Norvilitis, J.M., Reid, H.M. & O'Quin, K. (2022). Amotivation: A Key Predictor of College GPA, College Match, and First-Year Retention. *International Journal of Educational Psychology, 11*(3), 314-338.

Northern Valley Regional High School, Sequoia High School, The Herzstein School. (2022-2023). *School Profiles*.

Page, L. & Scott-Clayton, J. (2016). Improving College Access in the United States. *Economics of Education Review*.

Pattison, E., Grodsky, E. & Muller, C. (2014). Comparing Weighted and Unweighted Grade Point Averages in Predicting College Success. *NASSP Bulletin*.

Penn AHEAD/Perna, L. et al. (2019). Questioning the Calculations.

Perry, R.P. et al. (2001). Academic Control and Action Control in the Achievement of College Students. *Journal of Educational Psychology*.

Reardon, S.F. (2011). The Widening Academic Achievement Gap Between the Rich and the Poor. In *Whither Opportunity?*

Research for Action. (2025). Coaching as an Effective Student Support in Last-Dollar Scholarship Programs.

Rice Center for College Readiness. (2024). Undergraduate Admissions Process Module 2.

Roderick, M., Coca, V. & Nagaoka, J. (2011). Potholes on the Road to College. *Sociology of Education*.

Sallie Mae & Ipsos. (2019). *How America Pays for College*.

Schaeffer, B. & Reeves, R.V. (2020). The Lumpy Curve: Solving the Grade Inflation Problem. Brookings Institution.

Sedlacek, W.E. (2011). Using Noncognitive Variables in Assessing Readiness for Higher Education. *Readings on Equal Education, 25*, 187-205.

Severiens, S.E. & Schmidt, H.G. (2009). Academic and social integration and study progress in problem based learning. *Higher Education, 58*, 59-69.

Smith, M.C., Gosky, R.M., & Li, J-T. (2022). College Visits as Postsecondary Outcome Predictor Study. GEAR UP/NCAN.

Soland, J. (2017). Combining Academic, Noncognitive, and College Knowledge Measures to Identify Students Not on Track for College. *Readings on Equal Education*.

Stage, F.K. (1989). Motivation, Academic and Social Integration, and the Early Dropout. *American Educational Research Journal, 26*(3), 385-402.

Swanson, E. et al. (2020). An Evaluation of the Educational Impact of College Campus Visits. *Journal of Research on Educational Effectiveness*.

Swanson, R. & Cole, D. (2022). The Role of Academic Validation in Developing Mattering and Academic Success. *Research in Higher Education*.

Syverson, S., Franks, V.W., & Hiss, W.C. (2018). Defining Access: How Test-Optional Works. NACAC.

Theobald, E.J. et al. (2020). Active Learning Narrows Achievement Gaps. *Proceedings of the National Academy of Sciences*.

U.S. Department of Education, Federal Student Aid. (2013). *College Preparation Checklist*. ED542097.

Velenchik, A. (Wellesley College). (2024). *Tips for Writing Effective College Recommendations*.

Walker, A. et al. (2024). Exploring the Relationship Between Motivation and Academic Performance Among Online and Blended Learners. Oregon State University.

Wang, M.-T. & Fredricks, J.A. (2014). The Reciprocal Links Between School Engagement, Youth Problem Behaviors, and School Dropout During Adolescence. *Child Development*.

Weinstein, C.E., Palmer, D.R. & Acee, T.W. (2016). *LASSI User's Manual*. Third Edition.

West, M.R. (2014). The Limitations of Self-Report Measures of Non-cognitive Skills. Brookings Institution.

Westrick, P.A. et al. (2019). Validity of the SAT for Predicting First-Year Grades and Retention. College Board Research Report 2019-5.

Xing, X., Huerta, M., & Garza, T. (2019). College and Career Preparation Activities. *Career and Technical Education Research*.

Yan, Z., King, R.B. & Haw, J.Y. (2021). Formative Assessment, Growth Mindset, and Achievement. *Educational Psychology*.

Zimmerman, B.J. (2002). Becoming a Self-Regulated Learner: An Overview. *Theory Into Practice*.

Zinth, J.D. (2014). College Counseling in High Schools: Advising State Policy. Education Commission of the States.
