



AP Score Changes: 2024

Summary

Recent advances in data collection and analysis, including newly available technologies, have enabled researchers to identify with more precision than ever before what students have learned in their Advanced Placement classes. Accordingly, a more robust approach to setting standards for AP Exams has been applied to a variety of AP subjects over the past three years. Specifically, from 2022–2024, researchers applied evidence-based standard setting (EBSS) processes¹ to determine appropriate performance standards for students in a range of AP courses. This methodology collects input from hundreds of experts and then assembles fine-grained student performance data for analysis by subject-matter experts.

The EBSS process has consistently confirmed that 60%–80% of AP examinees, depending on the subject, perform with competency across the topics and skills of the exam, and thus merit an AP score of 3 or higher. In AP subjects that have historically had a success rate of 60%–80%, the EBSS process has confirmed, but not raised, that success rate. Among 9 AP subjects with recent success rates lower than typical, the EBSS process justified increases to the same 60%–80% success rates that have consistently been warranted in the other AP subjects.

Results

26 AP subjects. Most AP subjects have consistently had an exam success rate of ~60%–80%. (The “success rate” is the percentage of exams earning scores of 3 or higher on the 5-point AP score scale.) The list of AP subjects with these typical AP success rates can be found in Table 1. When the EBSS process has been applied to a selection of subjects in this table, the success rate has not increased in these subjects. Rather, the EBSS process has consistently confirmed the appropriateness of these typical AP success rates. Accordingly, scores in these 26 subjects have been consistent over time.

9 AP subjects. In contrast, the 9 AP subjects in Table 2 have recently had success rates lower than the 60%–80% typical of AP subjects. (Exam performance data for all AP subjects for recent years is available on the College Board’s website at: <https://apcentral.collegeboard.org/about-ap/ap-data-research>.) When EBSS methodologies were used to evaluate student performance in these 9 subjects, student performance analyses supported raising these success rates, such that they are now in alignment with the ~60%–80% success rates typically achieved in other AP subjects. The 9 subjects with increased success rates resulting from the EBSS process are listed in Table 2.²

1. For AERA’s peer-reviewed overview of the evidence-based standard setting methodology and process, see Katie Larsen McClarty, Walter D. Way, Andrew C. Porter, Jennifer N. Beimers, and Julie A. Miles, “Evidence-Based Standard Setting: Establishing a Validity Framework for Cut Scores.” *Educational Researcher* 2013:42.

2. The 5 AP subjects not appearing in Table 1 or Table 2 are discussed later in this brief.

Table 1: AP Exams That Have Not Had Score Changes

AP Subject	2024 AP Exam Success Rate	AP Subject	2024 AP Exam Success Rate
AP Calculus AB	64%	AP Japanese Language	74%
AP Calculus BC	80%	AP Spanish Language	83%
AP Computer Science Principles	65%	AP Spanish Literature	66%
AP Computer Science A	66%	AP 2D Art and Design	83%
AP Statistics	62%	AP 3D Art and Design	72%
AP Physics 2	68%	AP Drawing	84%
AP Physics C: Mechanics	75%	AP Art History	63%
AP Physics C: Electricity/Mag	67%	AP Music Theory	61%
AP Comparative Govt/Politics	73%	AP Seminar	89%
AP Chinese Language	87%	AP Research	85%
AP French Language	71%	AP Psychology	61%
AP German Language	67%	AP African American Studies	73%
AP Italian Language	71%	AP Precalculus	76%

Table 2: AP Exam with Score Changes: 2022, 2023, and 2024³

AP Subject	Year of EBSS Standard Setting	2024 AP Exam Success Rate
AP World History	2022	64%
AP English Literature	2022	72%
AP Biology	2022	68%
AP Macroeconomics	2023	62%
AP Microeconomics	2023	65%
AP Chemistry	2023	75%
AP U.S. Government	2024	73%
AP U.S. History	2024	72%
AP European History	2024	72%

3. The score distributions (the percentage of exams earning 1s, 2s, 3s, 4s, and 5s) for the 2024 AP Exams are posted on the College Board's website when scoring is complete in June. These can be found at: <https://apstudents.collegeboard.org/about-ap-scores/score-distributions>. These are worth reviewing for AP Biology, since the EBSS standard setting for AP Biology resulted in changes to the percentages of 4s and 5s, rather than 3s, so the EBSS-related changes are not visible by looking at the success rate alone.

Methodology

Prior to 2022, the AP Program utilized standard-setting panels to confirm or change AP scores every 5–10 years. These panels of 10–18 educators followed established and well-documented protocols⁴ (e.g., the Bookmark methodology, the modified Angoff methodology, the Body of Work methodology) and drew upon their individual expertise to estimate the percentage of AP students who should earn each AP score. In years prior to the emergence of EBSS, these standard-setting panels were industry standard, and remain viable for many assessment programs today.

But in contrast to standard-setting panels which, by their very nature, include a small number of experts, evidence-based standard setting collects input from hundreds of educators and assembles multiple sources of information about student performance. Moreover, because the EBSS process is so heavily anchored in student performance data, it is especially well-suited for ensuring that AP standards and scores are not tugged higher by the well-documented increases in college grades over the past 30 years.⁵ Accordingly, application of the EBSS process does not aim to match AP scores to college grades.

Over the past decade, two key developments have made it possible for AP to use EBSS rather than smaller panels:

1. The evolution of digital data collection technologies has made large-scale data collection and analysis possible within very tight operational timeframes between students sitting for the exam in May, faculty scoring of the exams in June, and the reporting of scores to colleges, educators, and students in early July each year. Evidence-based standard setting utilizes emergent technology to collect, organize, and analyze inputs from hundreds of teachers and faculty, rather than just the perspectives of 10–18 panelists.
2. Beginning in fall 2019, the AP Program provided all AP teachers with a new digital library of AP course materials—titled AP Classroom—and an accompanying course and exam description binder. This material, for the first time in the AP Program’s history, established coherent units, topics, learning objectives, and skills for each AP course that could be tagged to each exam question, generating a comprehensive collection of metadata about the specific content knowledge and skills students demonstrated on the exam. As a result, more granular and targeted student performance data is now available within a very short operational window for analysts to utilize for identifying student performance at basic, moderate, and exceptional levels.

For example, the exam performance data for AP U.S. History enables distinctions between the following:

- Students who could only read and explain two historians’ different interpretations of the origins of the women’s rights movement in the early 19th century, versus students who were also able to summon historical evidence to support or refute those historians’ claims.
- Students who were only able to describe a British policy in colonial North America, versus students who could go further in explaining the ways different groups in North America responded to those policies, and students who could also draw causal connections to the American colonists’ victory over Great Britain in the Revolutionary War.

4. Cizek, G. J., (2012). *Setting Performance Standards: Foundations, Methods, and Innovations*, 2nd Edition, New York, NY.

5. National Center for Education Statistics, *National Postsecondary Student Aid Study: 2020 Undergraduate Students; 1990 Undergraduate Students*.

- Students who could only develop a basic thesis about the ways slavery shaped U.S. society between 1783 and 1840, and students who could analyze multiple primary sources and assemble them as evidence, and students who could marshal additional context and evidence.
- Students who could only develop a basic thesis about the causes of the growth of a national culture in the United States from 1800 to 1848, and students who could develop moderately successful essays with some appropriate evidence, and students who could develop well-supported essays with substantial evidence.

Subject-matter experts determined that students receiving 3s, 4s, and 5s must be students whom the data indicated were consistently demonstrating baseline proficiency across the array of course skills, such as summoning and using historical evidence on a range of topics and questions. In turn, students earning scores of 4 and 5 demonstrated the ability to marshal more extensive historical evidence than students earning a 3, while students receiving 2s could not go much beyond the crafting of basic thesis statements and the recognition of key historical developments when reading primary source documents. When these standards were applied to this year’s AP U.S. History students, 72% demonstrated this proficiency with evidence and historical knowledge essential for an AP 3.

Targeted analysis of student performance data is one of the key ways the EBSS process prevents AP grades from drifting over time: The standard for a score of 3 remains anchored in demonstrated student performance, rather than panelists’ fluctuating perceptions of what a college A, B, or C might mean at any one particular institution.

Maintaining AP Standards Over Time

The EBSS process gathered an unprecedented amount of data about the tens of thousands of students taking U.S. history survey courses in college, so that current college grading practices could be identified and compared to past college grading practices and AP Exam success rates.

Table 3 shows the course grades of more than 90,000 college students who took U.S. history survey courses in college from spring of 2021 through fall of 2023 compared to the 2024 AP U.S. History scores derived through the EBSS process:

Table 3: 2021-23 College Students’ Grades in U.S. History and 2024 AP U.S. History Scores

College Grades: U.S. History		Scores: AP U.S. History	
A+/A	27%	AP 5	12%
A-/B+/B	38%	AP 4	34%
B-/C+/C	19%	AP 3	26%
C-/D+/D	8%	AP 2	18%
D-/F	8%	AP 1	9%
C or higher	84%	3 or higher	72%

Converting the 2021–2023 college grades and the 2024 AP grades to the GPA scale and taking the average, the 2021–23 college grades are at 2.90, 12% higher⁶ than average college grades in 1990 of 2.59. By way of comparison, the increases in AP U.S. History scores for 2024 result in an AP average of 2.47 on the GPA scale, well below the college grades of 2.59 in 1990.

This pattern holds true across 8 of the 9 AP subjects for which the scores increased when the EBSS process was applied in 2022, 2023, or 2024. AP Chemistry, in contrast, is taken by such a self-selecting population of advanced chemistry students that, like AP Calculus BC, the data reviewed through the EBSS process indicated that a full 75% of AP Chemistry students demonstrated proficient content knowledge and skills.

In summary, AP sets standards that are significantly more stringent than the standards represented by colleges' own grade distributions: Colleges' grades in humanities courses are typically 85% Cs or better, and colleges' grades in STEM courses are typically 75% Cs or better, whereas for most AP subjects, the evidence shows that 60%–75% of AP Exams should receive scores of 3 or higher in order to maintain the historical standards associated with AP scores.

Table 4: 9 Subjects That Have Had Score Changes, Compared to College Grades in 2021-23

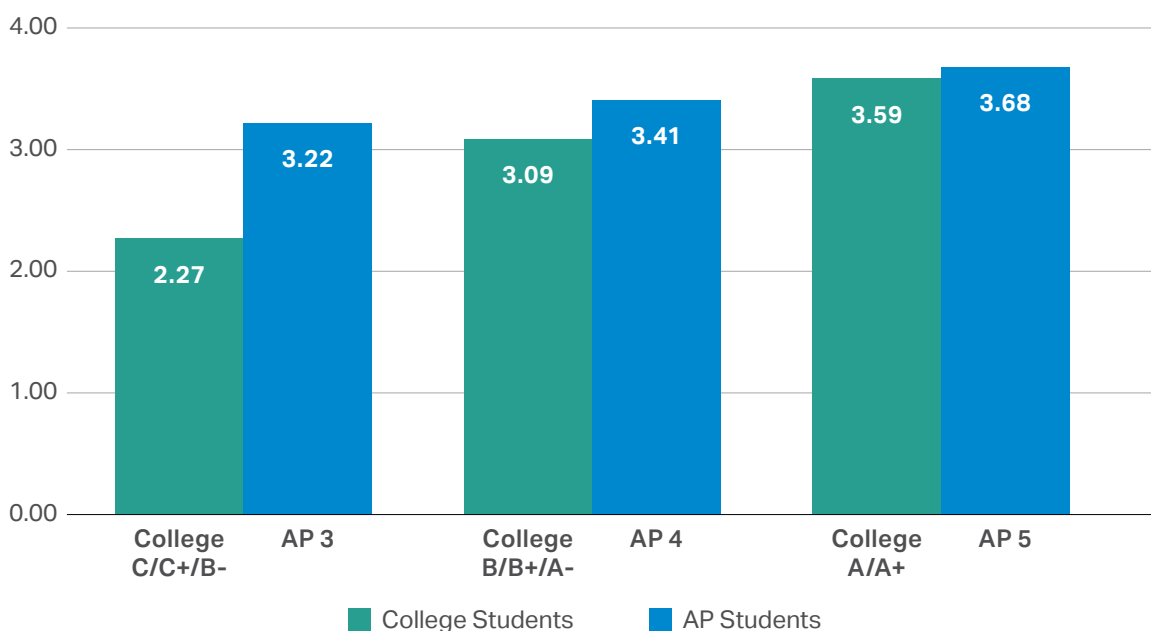
AP Subject	Year of EBSS Standard Setting	2024 AP Exam Success Rate	% of College Students Earning Grades of C or Higher When Taking the College Equivalent of This AP Course*
AP World History	2022	64%	85%
AP English Literature	2022	72%	92%
AP Biology	2022	68%	77%
AP Macroeconomics	2023	62%	91%
AP Microeconomics	2023	65%	87%
AP Chemistry	2023	75%	74%
AP U.S. Government	2024	73%	85%
AP U.S. History	2024	72%	84%
AP European History	2024	72%	84%

*AP scores of 3 or higher are designed to be equivalent to a college C or higher. The average college GPA rose by 12% over the past 30 years; in contrast, the AP score verification process aligns AP scores with college grading standards prior to 1990. (National Center for Education Statistics, National Postsecondary Student Aid Study: 2020 Undergraduate Students; 1990 Undergraduate Students)

6. National Center for Education Statistics, National Postsecondary Student Aid Study: 2020 Undergraduate Students; 1990 Undergraduate Students.

As part of the EBSS process, data were assembled on the most recent cohorts of AP U.S. History students for whom years of subsequent college data are available—the students who took AP U.S. History in 2017 and 2018—observing their performance in upper-division history courses in college, in comparison to GPAs in upper-division history courses of college students who took the U.S. history survey course on campus and received specific college grades. This analysis shows that AP U.S. History students who have been earning an AP score of 3 proceed to earn subsequent history GPAs of 3.22, far outpacing the group with whom they should be similar, college students who earned a C, C+, or B- in the college U.S. history survey coursework. Such findings confirm that AP U.S. History students are highly successful in college history programs and that increases in AP scores as a result of EBSS will not reduce the accuracy with which AP scores of 3 or higher place students ahead in college. (See Figure 1.)

Figure 1: GPAs in Subsequent College History Classes: Students Who Took Colleges’ American History Survey Compared to AP U.S. History Students



The Strong Academic Readiness of AP Students for Challenging Coursework

Under the EBSS process, the question was asked: Is there any reason to believe that AP students are not prepared to learn at a college level and qualify for college credit? To answer this question, researchers utilize objective performance data for both college students and AP students to make appropriate comparisons.

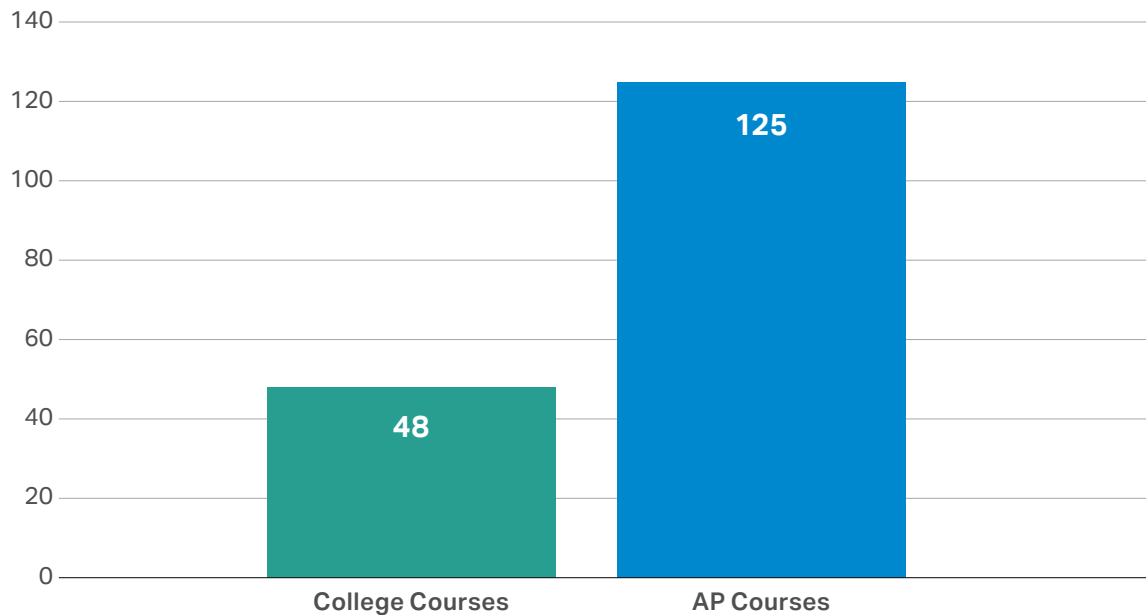
For example, scores on the PSAT-related assessments—well-established predictors of college preparedness and utilized for National Merit Scholarships—were compared among college U.S. history students who receive grades of A, B, C, D, and F and the most recent cohort of AP U.S. History students for whom such data are available. The results, in Table 5, show the remarkable academic preparation for college-level coursework among the AP students: AP U.S. History students who have historically received an AP score of 2 have higher average PSAT 10 scores than college students who have been receiving college grades of A+ and A.

Table 5: Average PSAT 10 Scores for Students in Introductory College Courses and AP Students – U.S. History

Course Grade (AP Score)	College-Level Introductory Course		2022 AP U.S. History Exam	
	Count	Mean Score	Count	Mean Score
A+ and A (AP 5)	30,648	1058	39,221	1311
A-, B+, and B (AP 4)	32,359	1015	54,093	1231
B-, C+, and C (AP 3)	16,758	978	71,554	1161
C-, D+, and D (AP 2)	5,796	959	70,217	1084
D- and F (AP 1)	6,644	959	76,686	956

The EBSS process also collects information from colleges and AP classrooms nationwide on the number of instructional hours students spend on the course content. EBSS analyses have consistently found that AP students receive more than double the number of classroom instructional hours that college students spend in the comparable college course. Figure 2 illustrates this comparison for college students taking the Introduction to American Government course on campus and high school students taking AP U.S. Government.

Figure 2: Modal Number of Instructional Hours: College American Government Classes and AP U.S. Government Classes



Other Notes:

“Easier” rubrics: In 2024, the AP History essay rubrics—which are the scoring guidelines used by the AP graders to determine how many points students should receive for their essays—were changed to ensure the scoring process better distinguished between each of the five levels of AP performance. In prior years, the rubrics had too many points that only measured the differences between extremely strong students (AP 5) and strong students (AP 4). But graders also need to identify the differences between low performance (AP 1) and basic performance (AP 2), and between basic performance and proficient performance (AP 3). If all points on the rubric are equally difficult to obtain, the scoring process will not generate as much data about students at the novice and intermediate levels of performance as it will about the most advanced students. In other words, the best rubric for exams like AP is a rubric that has some novice-level points, some moderately challenging points, and some very challenging points, so that students can effectively be placed within one of the five AP score categories. Accordingly, this year’s changes ensured that one point was dedicated to differentiating between AP 1s and AP 2s. The other points on the rubric continued to focus on identifying more proficient and advanced levels of performance—the points students need to accumulate on their essays to receive a 3, 4, or 5. A change was also made to an advanced point that was rarely used by graders and so provided no measurement value. Instead, this point was changed so that it could be used to help distinguish whether a student should receive a 3, 4, or 5.

Differences in success rates among similar exams: Because AP World History, AP European History, and AP U.S. History focus on the same historical thinking skills, with the same exam design, and the same rubrics for scoring the essays, some might expect that the success rates would be the same for these three exams. Instead, applying standards derived from the EBSS data simultaneously and equivalently to these 3 subjects in 2024 resulted in 64% of AP World History Exams earning score of 3+, in contrast to 72% of European History and 72% of U.S. History Exams. World History students earned fewer points on their essays than European History and U.S. History students, hence the lower success rate. This is unsurprising, as World History students are typically a grade younger than U.S. History students, and European History students are a more self-selected and smaller group than World History students.

Future applications of the EBSS methodology: Upcoming exam revisions in these five remaining AP subjects will enable the application of the EBSS methodology, which is now the standard methodology for AP standard settings: AP English Language, AP Environmental Science, AP Human Geography, AP Latin, and AP Physics 1.

External review and validation: The AP Program relies on external experts like the American Council on Education (ACE) to independently review and verify AP processes. In their most recent report, ACE stated that the validity evidence for AP scores with success rates in the typical 60-80% range was “exceptionally strong.”

Briefing the AP Community: As we do each time we conduct a score verification process, we have shared the findings with thousands of college faculty and teachers through presentations, briefings, and memos at the AP Readings and via the AP online Teacher Communities. We also reviewed the standard-setting results with the College Board’s governance committees and advisory groups that include college faculty, enrollment, and admissions leaders.