



How Is the CPA Exam Scored?

Effective January 1, 2011

Prepared by the American Institute of Certified Public Accountants

How is the CPA Exam Scored?

Anyone who has taken the Uniform CPA Examination, prepared for the Exam, or been involved in the CPA licensure process knows that the passing score is 75. But very few understand what that 75 means. We regularly hear from candidates, state board representatives, review course providers, educators, and others who wonder how the AICPA scores the Exam. To help answer the most frequently asked questions, we have prepared this non-technical overview of the scoring process. Please review and keep this document as a handy reference guide. We welcome additional questions and anticipate updating the document periodically based on the feedback we receive.

The Structure of the CPA Examination

Before we explain how the Exam is scored, let us first review how it is structured. The CPA Exam is comprised of four sections: Auditing and Attestation (AUD), Business Environment and Concepts (BEC), Financial Accounting and Reporting (FAR), and Regulation (REG). All four sections contain multiple-choice questions (MCQs). AUD, FAR, and REG sections have an additional portion for task-based simulation (TBS) questions; BEC has a portion for written communication questions, but no TBS questions.

The MCQs within each section are administered in three blocks, called “testlets.” Each testlet contains operational and pretest questions. Operational questions are scored, while pretest questions are not scored. Instead, a candidate’s response to a pretest question is used to evaluate the question’s statistical performance. It’s important to note that a majority of the questions are operational questions; however, pretest questions are mixed into the Exam and are not identified as pretest questions. From the candidate’s perspective, pretest questions are indistinguishable from operational questions. Pretest questions that meet certain statistical criteria are used as operational questions on future exams. This strategy for pretesting questions is common practice in high stakes testing.

Multiple-choice testlets vary in difficulty—there are two levels that are labeled “medium” and “difficult.” Within the testlets, items often vary substantially in their difficulty levels, but across testlets, those labeled “difficult” contain harder questions on average than testlets labeled “medium.” Every candidate receives a medium testlet first. Succeeding testlets can be either medium or difficult, depending on a candidate’s performance (for more information about multi-stage testing, see FAQ #3, below). The scoring procedures take the difficulty of all questions into account so that candidates are scored fairly regardless of the difficulty of the testlets they take.

In addition to multiple-choice questions, the AUD, FAR, and REG sections also contain task-based simulations. Task-based simulations are case studies that allow candidates to demonstrate their knowledge and skills by generating responses to questions rather than simply selecting the correct answer. Task-based simulations typically require candidates to use spreadsheets and/or research authoritative literature provided in the Exam.

Unlike the other three sections, the BEC section does not have any task-based simulations. Instead, it contains three written communication questions. Each question requires the candidate to write a short passage, letter, or memo.

Overview of the CPA Exam by Section

	AUD	FAR	REG	BEC
Time	4 Hours	4 Hours	3 Hours	3 Hours
Multiple-Choice Questions (MCQ)				
Testlets	3	3	3	3
Operational	75	75	60	60
Pretest	15	15	12	12
Total MCQ	90	90	72	72
Task-Based Simulations (TBS)				
Operational	6	6	5	n/a
Pretest	1	1	1	n/a
Total TBS	7	7	6	n/a
Written Communication Tasks				
Operational	n/a	n/a	n/a	2
Pretest	n/a	n/a	n/a	1
Total Written Communications	n/a	n/a	n/a	3

Before appearing on the CPA Examination, all operational and pretest questions have passed through several extensive and rigorous subject matter reviews to ensure that they are technically correct, have a single best or correct answer, are current, and measure entry level content as specified in the Content Specification Outlines (CSOs). The CSOs specify the percentage that each section of multiple-choice questions should be devoted to each content area. The current CSOs were put in effect in 2011 based on the results of a practice analysis and state board of accountancy responses to an exposure draft of the recommended CSOs. Operational questions have also been statistically evaluated to ensure they meet the psychometric requirements of the CPA Exam.

Score Scale and Passing Score

Section scores are reported on a scale that ranges from 0 to 99. A total *reported* score of 75 is required to pass each section. This is not a percent correct score and cannot be interpreted as a percentage. The total score in the AUD, FAR and REG sections is a weighted combination of scaled scores from the multiple-choice and task-based simulation questions. For the BEC section, the total score is a weighted combination of the scaled scores from the multiple-choice questions and written communication tasks. Scaled scores on the multiple-choice and task-based simulation portions of the Exam

are calculated using formulas that take into account factors such as whether the question was answered correctly and the relative difficulty of each question.

Percent of Score Contributed by Item (Question) Type				
	AUD	FAR	REG	BEC
Multiple-Choice Questions	60%	60%	60%	85%
Task-Based Simulations	40%	40%	40%	n/a
Written Communication Tasks	n/a	n/a	n/a	15%

Frequently Asked Questions and Answers

1. Can you compare scores of candidates who are administered Exams with variable difficulty and different questions?

Yes. Scores from the different administrations of the Exam are placed on a standard scale so they can be compared to each other. This process accounts for any differences in difficulty among the versions. All total scores are reported on the CPA Examination 0 to 99 scale (note that scores administered after the implementation of CBT-e (January 1, 2011) are *not* comparable with scores from prior administrations). The use of a standard reporting scale is common practice in the testing industry. You may be familiar, for example, with the 200 to 800 SAT scale or the 1 to 36 ACT scale.

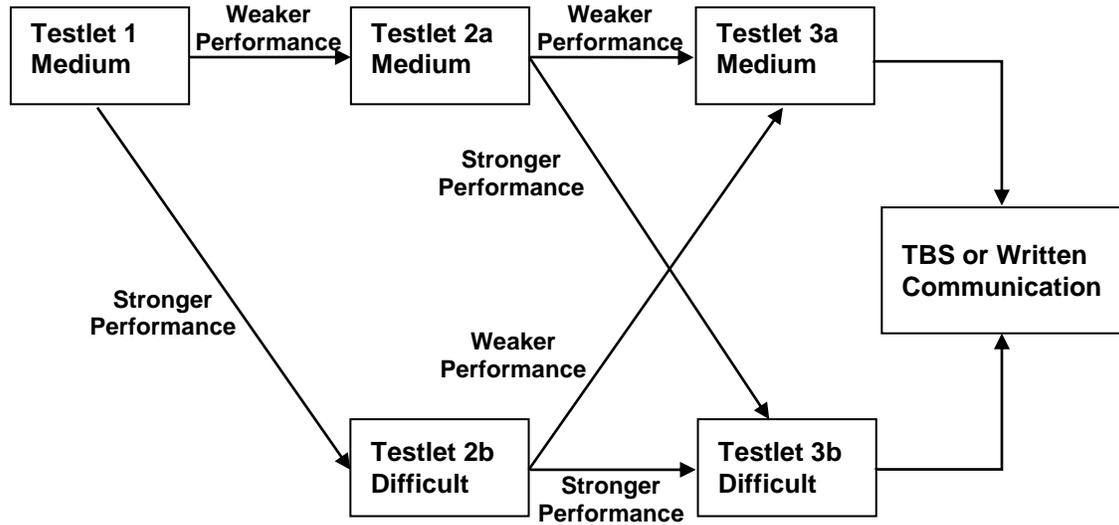
2. Are some administrations of the Exam more difficult than others?

There may be minor differences among different administrations; however these differences are accounted for during scoring. Since we use different questions for different administrations to enhance the security of the Exam, we create multiple forms of the Exam. Each form is comparable but not identical. Great care is taken to match the forms in terms of content and item difficulty. Remember, you may be given questions of varying difficulty depending on your performance (see The Structure of the CPA Examination section, above). The difficulties of the questions are accounted for during scoring. Therefore, it does not mean that it is easier to get a higher score simply because you receive easier questions.

3. When are easier or more difficult questions given?

Candidates take three multiple-choice testlets. A testlet is a group of multiple-choice questions. The first testlet is always a medium testlet. If you perform well on the first testlet, you will get a more difficult second testlet while those who do not perform well on the first testlet will receive a second medium difficulty testlet. Similarly, the third testlet can be a medium or a difficult and assignment of the third testlet is based upon your performance on the first two testlets. This testing process is called Multi-Stage Testing (MST) or Computer Assisted Sequential Testing (CAST). Please note that the task-based simulations are not chosen based on your performance on the multiple-choice testlets. They are pre-assigned. The following diagram shows how it works. If you have more questions about Multi-Stage Testing please see FAQ #9.

Testlet Selection on the Uniform CPA Examination



4. Can I compute my score from the number of questions I answered correctly?

No. The total reported score is a scaled value that takes into account both the response to and the statistical characteristics of each question administered. If you have more questions about the statistical characteristics see FAQ #14.

5. In college some of my professors gave tests that had questions that were worth one point and others that were worth two points. If one student got five of the one-point questions right, he got five points. If another student answered five two-point questions right, he got 10 points. Is that what you're doing?

Conceptually, yes. But the professor assigned the weights based on judgment. In the CPA Exam, the weights are determined from candidate response data using Item Response Theory (IRT).

6. How do you score the written communications responses?

Most responses are scored by a computer grading program, which is calibrated using human scorers. In some cases, responses are scored by a network of human graders (all CPAs). If your score is close to the passing score, your written responses will be automatically re-graded by human graders. When there is more than one grader for a response, the average of the scores is used as the final grade.

(The following FAQ's offer more detail about the exam. Read on if you would like to gain a better understanding of the scoring process.)

7. Can I pass the BEC section just by doing well on the multiple-choice questions?

Yes. However, it would be very difficult to do so as you would have to perform exceedingly well. It is advisable to be prepared for both the multiple-choice and the written communication questions.

8. Can I pass the AUD, FAR, or REG sections just by doing well on the multiple-choice questions?

No. The portion contribution from task-based simulations in those sections is too large. You would need to get some of the TBS questions correct in order to pass.

9. Is Multi-Stage Testing fair? Why are you using it?

Yes, it is fair. Since the characteristics of the test questions are taken into account in the scoring, there is no advantage or disadvantage to being assigned testlets of different difficulty. We use Multi-Stage Testing because the test questions are matched to proficiency levels, and therefore fewer questions are needed to obtain accurate estimates of proficiency.

10. How do you decide which questions are difficult and which are medium?

The difficulty levels of the test questions (and other statistics that are used to describe each test question) are determined through statistical analysis of candidate responses. At the question level, difficulty is not quantified as a category (e.g. moderate or difficult), but as a numeric value along a scale. Testlets are classified as either medium or difficult based on the average difficulty of the questions within that testlet.

11. Does that mean difficult testlets can have easier questions and medium testlets can have difficult questions?

Yes. All testlets have questions ranging in difficulty. Questions in difficult testlets just have a higher average level of difficulty than those in medium testlets.

12. What if I'm well prepared but do poorly on the first testlet? Can I still pass?

Yes, you can still pass the exam but you will need to do better on the second and third testlets.

13. Can I get all medium testlets and still pass?

Yes, but for this to happen, you would have good but not excellent performance on the first two testlets, and then excellent performance on the last testlet.

14. What do you mean when you say “statistical characteristics” of test questions?

There are three statistics used to describe the questions: Difficulty – whether the question is generally easier or more difficult for candidates, Discrimination – how well the question differentiates between more able and less able candidates, and Guessing – the chances of candidates answering the question correctly just by guessing. The statistics are generated when the questions are administered as pretest questions and used in the scoring when the questions are operational. The formulas for generating the statistics and scoring the Exam come from a scoring approach commonly referred to as “Item Response Theory.” Item Response Theory is currently being used or has recently been adopted by nearly all of the large licensing examination programs in this country and also by many of the moderate-sized and smaller examination programs.

15. Can two candidates answer the same number of questions correctly, but get different scores?

Yes, because scores depend on the characteristics of the questions, not just how many are answered correctly. For example, all else being equal, a candidate who correctly answers 10 difficult questions would score more than a candidate who correctly answers 10 easy questions.

16. How was the passing score set for each section?

Volunteers who are licensed CPAs with recent experience supervising entry level CPAs participated in a passing score study. They reviewed test questions and how candidates performed on those questions in order to make judgments of what test performance is required to ensure the protection of the public interest. The results of this study were used by the Board of Examiners (BOE) as a guide when it established the passing scores for each section. These passing scores were then mapped to a score of 75 on the scale used to report scores. This process, known as Standard Setting, is common practice within high stakes testing.

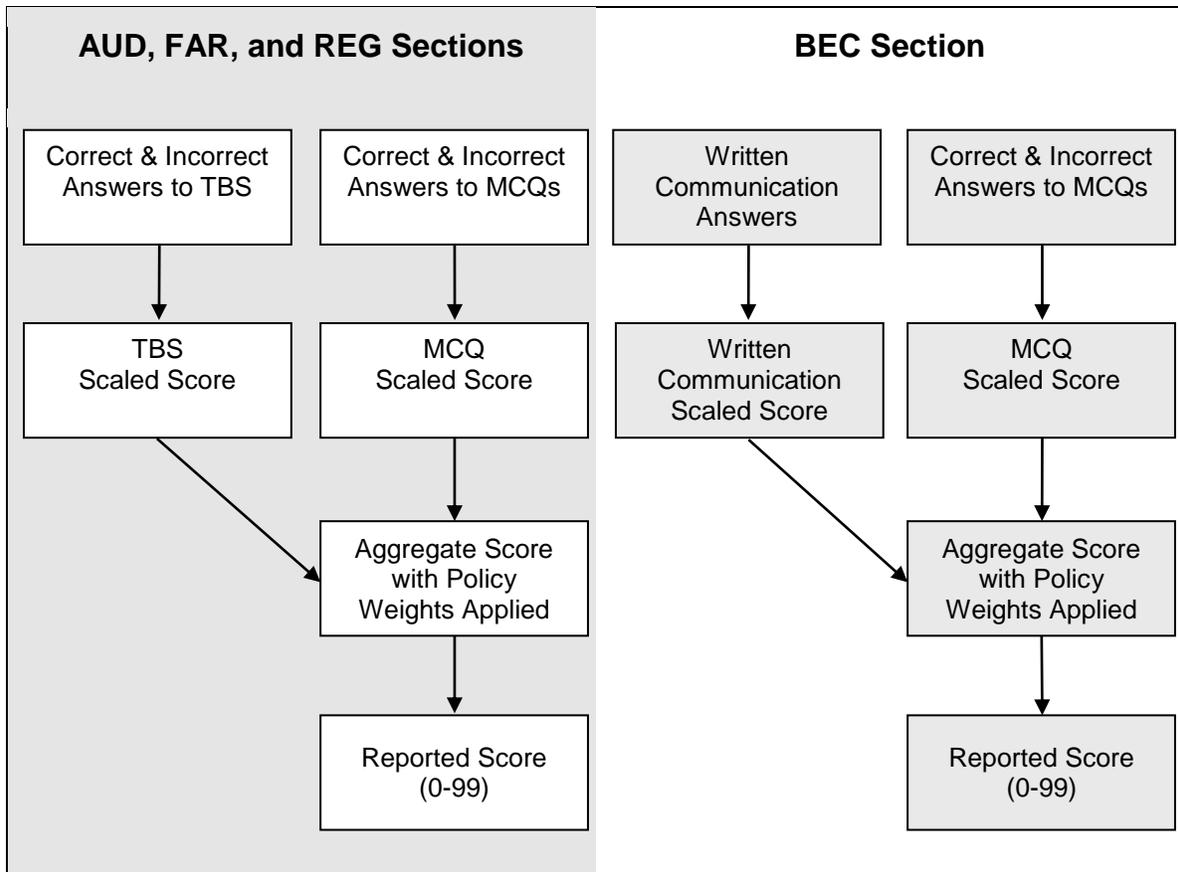
17. How do I find out my scores for each content area of the exam?

We do not release subscores by content area. We do, however, report categories of performance. You should use caution in interpreting your content area performance. The subscores are calculated on fewer items and therefore not as reliable as the final score. The performance comparisons of *weaker*, *comparable* and *stronger* are provided to candidates as a general indicator of performance. The subscores are not disclosed because we do not want to convey a sense of precision in those scores that is unwarranted. That’s why, in the event you decide to retake any section of the Exam, you should study all content areas. If you study only the areas in which you are weak, you might do better on those areas, but worse on others when you take the Exam again.

18. In general terms, what are the steps taken to produce the reported score?

Initially, for purposes of score reporting, each component (multiple-choice questions, task-based simulations, and written communication) is treated separately. For the multiple-choice and task-based simulation components, Item Response Theory is used to obtain the scaled score for each type of question. The multiple-choice estimate is then

mapped to a score on a multiple-choice scale. Similarly, the task-based simulation estimate is mapped to a score on a task-based simulation scale. The total written communication raw score is mapped to a written communication scaled score. The scores are then combined with the policy weights (60% multiple-choice and 40% simulations for AUD, FAR and REG, 85% multiple-choice and 15% written communications for BEC). The final step involves mapping the aggregate score to the 0 to 99 scale used for score reporting. The process can be viewed in the following diagram.



19. Is there more that I can read about the Exam?

Yes. There are many technical reports related to the psychometric characteristics of the Exam. All of these reports are posted to the [Psychometrics section of the Exam website](#). Other useful publications can also be found under the Research and Reports section of the site.