AICPA®

How is the Uniform CPA Examination scored?

Effective April 1, 2017

Contents

A passing score is 75

Anyone who has taken the Uniform CPA Examination, prepared for the Examination, or been involved in the CPA licensure process knows that the passing score is 75. But very few understand what that 75 means. The AICPA regularly hears from candidates, state board representatives, educators, and others who wonder how the Examination is scored. This non-technical overview of the scoring process will help answer the most frequently asked questions.

Uniform CPA Examination structure

The Examination comprises 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 (MCQ) and task-based simulations (TBS). BEC has also has a portion for written communication tasks.

Testing within each Examination section is administered in blocks called testlets. Testlets contain operational and pretest questions. Operational questions are scored, while pretest questions are not scored. Pretest questions are indistinguishable from operational questions. Pretest questions that meet certain statistical criteria are used as operational questions on future examinations. This strategy for pretesting questions is common practice in high-stakes testing.

MCQ testlets vary in difficulty — there are two levels labeled "medium" and "difficult." Within the testlets, items often vary 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. The succeeding testlet can be either medium or difficult, depending on a candidate's performance (*for more information about multi-stage testing, see FAQ No. 3*). 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.

TBSs are case studies that allow candidates to demonstrate their knowledge and skills by generating responses to questions rather than simply selecting the correct answer. TBSs typically require candidates to use spreadsheets and/or research authoritative literature provided in the Examination.

In addition, the BEC section contains three written communication tasks. Each requires the candidate to write a short passage, letter, or memo.

Examination structure by section

Section	Item types*	Testlet
Auditing and Attestation (AUD) Time to complete: 4 hours	72 MCQs 8 TBSs	No. 1: 36 MCQs No. 2: 36 MCQs No. 3: 2 TBSs No. 4: 3 TBSs No. 5: 3 TBSs
Business Environment and Concepts (BEC) Time to complete: 4 hours	62 MCQs 4 TBSs 3 Written communication	No. 1: 31 MCQs No. 2: 31 MCQs No. 3: 2 TBSs No. 4: 2 TBSs No. 5: 3 Written communication
Financial Accounting and Reporting (FAR) Time to complete: 4 hours	66 MCQs 8 TBSs	No. 1: 33 MCQs No. 2: 33 MCQs No. 3: 2 TBSs No. 4: 3 TBSs No. 5: 3 TBSs
Regulation (REG) Time to complete: 4 hours	76 MCQs 8 TBSs	No. 1: 38 MCQs No. 2: 38 MCQs No. 3: 2 TBSs No. 4: 3 TBSs No. 5: 3 TBSs

* Each section includes 12 pretest MCQs and 1 pretest TBS. BEC also includes 1 pretest written communication task.

Before appearing on the Examination, all operational and pretest questions have passed through several extensive and rigorous subject matter reviews to ensure they are technically correct, have a single, best answer, are current, and measure relevant content knowledge and skill as specified in the Examination Blueprints. The blueprints specify the percentage of questions that should be devoted to each content area, and to each skill level. Operational questions have been statistically evaluated to ensure they meet the psychometric requirements of the Examination.

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 percentage correct score nor can it be interpreted as a percentage. The total score in the AUD, FAR and REG sections is a weighted combination of scaled scores from the MCQs and TBSs. For the BEC section, the total score is a weighted combination of the scaled scores from the MCQs, TBSs, and written communication tasks. Scaled scores on the MCQ and TBS portions of the Examination 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 type

Item type	AUD	BEC	FAR	REG
Multiple choice questions (MCQs)	50%	50%	50%	50%
Task-based simulations (TBSs)	50%	35%	50%	50%
Written communication tasks	n/a	15%	n/a	n/a

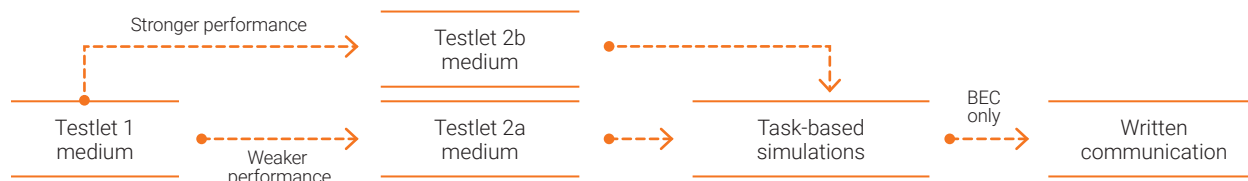# Score frequently asked questions and answers

1. **Can you compare scores of candidates who are administered Examinations with variable difficulty and different questions?**
    Yes. Scores from the different administrations of the Examination 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 Examination 0 to 99 scale. Given the significant changes made to the Examination (effective April 2017), a new passing score will be established via the standard setting process (*see FAQ No. 14*). It should be noted that scores for tests administered after April 1, 2017, are not comparable with scores from prior administrations except that 75 is the passing score. The use of a standard reporting scale is common practice in the testing industry. Similar examples are the SAT scale or the ACT scale.

2. **Are some administrations of the Examination more difficult than others?**
    There may be minor differences among different administrations; however these differences are accounted for during scoring. The AICPA enhances test security by creating multiple forms of the Examination with different questions for different administrations. Each form is comparable but not identical. Great care is taken to match the forms in terms of content and item difficulty. Remember, a candidate may be given questions of varying difficulty depending on his/her performance (see page 2, Uniform CPA Examination structure). 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 a candidate receives easier questions.

CPA examination testlet selection

```
        Stronger performance
     ┌─────────────────────────►     ┌──────────────┐        ●┄┄┄┄┄┄┄┄┄┐
     ┊                                │  Testlet 2b  │                 ▼
     ┊                                │   medium     │
 ┌───┴────────┐                       ├──────────────┤                          BEC
 │            │                                                                  only    ┌──────────────┐
 │ Testlet 1  │          ●┄┄┄┄┄┄►     ┌──────────────┐    ●┄┄►  ┌──────────────┐   ●┄┄┄►  │   Written    │
 │  medium    │        Weaker         │  Testlet 2a  │          │  Task-based  │          │ communication│
 │            │       performance     │   medium     │          │ simulations  │          │              │
 └────────────┘                       └──────────────┘          └──────────────┘          └──────────────┘
```

3. **When are easier or more difficult questions given?**
   Candidates take two multiple-choice testlets. The first testlet is always a medium testlet. If a candidate performs well on the first testlet, he/she 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. This testing process is called multi-stage testing (MST). Please note that the task-based simulations are not chosen based on candidate performance on the multiple-choice testlets. They are pre-assigned. The diagram on page 6 shows how it works. (*see FAQ No. 9 for further explanation of multi-stage testing*).

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. (see FAQs No. 12 and No. 16 for further information about the statistical characteristics and how scores are produced, respectively)

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 Examination, the weights are determined from candidate response data using item response theory (IRT). (*see FAQ No. 12 for further information about IRT*)

6. **How do you score the written communication tasks?**
   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 a candidate score is close to the passing score, his/her written communication tasks 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.

   Note: The remainder of the FAQs offer more detail about the Examination. Read on to gain a better understanding of the scoring process.

7. Can I pass by only doing well on the multiple-choice questions?

No. The portion contribution from task-based simulations and written communication tasks in BEC is too large. A candidate would need to get some of the TBS questions correct in order to pass.

8. 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. The AICPA uses multi-stage testing because the test questions are matched to proficiency levels, and therefore fewer questions are needed to obtain accurate estimates of proficiency.

9. 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.

10. 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.

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

Yes, a candidate can still pass the Examination but will need to do better on the second testlet; and that is just the MCQ portion. Fifty percent of a candidate's score comes from the TBSs in the AUD, FAR and REG sections while 35% comes from TBS and 15% from written communication tasks in BEC.

12. 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 Examination come from a scoring approach commonly referred to as — Item Response Theory (IRT). IRT is currently being used or has recently been adopted by nearly all of the large licensing examination programs in this country.

13. 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.

14. How is the passing score set for each section? Volunteers who are licensed CPAs with recent experience supervising entry-level CPAs participate in a passing score study. They review 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 the study are used by the Board of Examiners (BOE) as a guide when it establishes the passing scores for each section. These passing scores are 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.

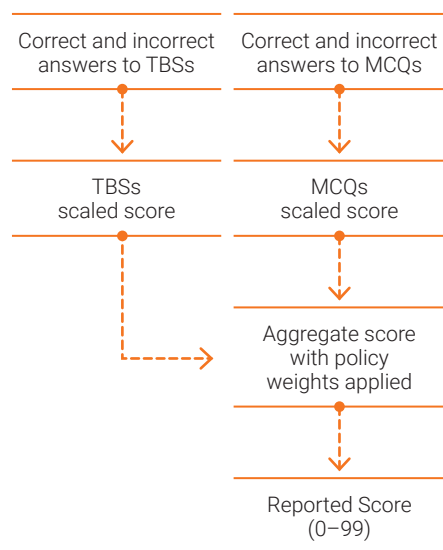15. How do I find out my scores for each content area of the Examination? The AICPA does not release subscores by content area, but does report categories of performance. Candidates should use caution in interpreting content area performance. The subscores are calculated on fewer items and therefore 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 the AICPA does not want to convey a sense of precision in those scores that is unwarranted. That's why, in the event a candidate decides to retake any section of the Examination, he/she should study all content areas. If a candidate studies only the areas where there is weakness, he/she might do better on those areas, but worse on others when taking the Examination again.

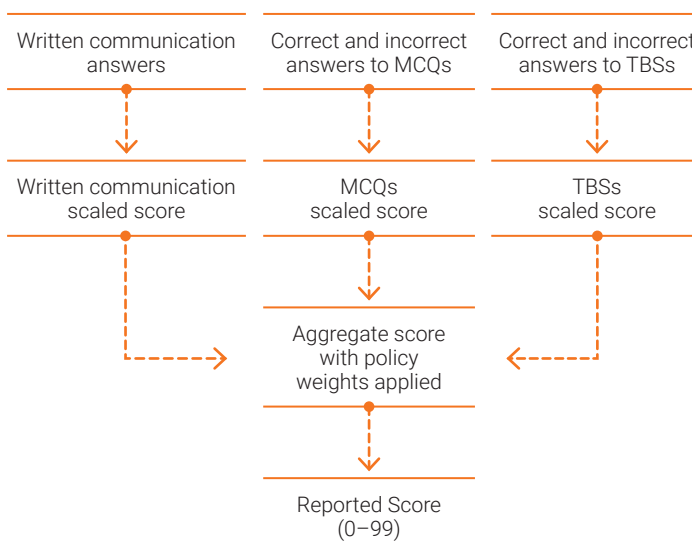16. **In general terms, what are the steps taken to produce the reported score?**
Initially, for purposes of score reporting, each component (multiple-choice questions, task-based simulations and written communication tasks) is treated separately. For the multiple-choice and task-based simulation components, item response theory (IRT) 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 task raw score is mapped to a written communication task scaled score. The scores are then combined with the policy weights 50% multiple-choice and 50% task-based simulations for AUD, FAR and REG, and 50% multiple-choice, 35% simulations, and 15% written communications for BEC. The final step involves mapping the aggregate score to the 0–99 scale with 75 equal to the passing score.

AUD, FAR and REG sections

Correct and incorrect answers to TBSs	Correct and incorrect answers to MCQs
↓	↓
TBSs scaled score	MCQs scaled score
	↓
	Aggregate score with policy weights applied
	↓
	Reported Score (0–99)

BEC Section

Written communication answers	Correct and incorrect answers to MCQs	Correct and incorrect answers to TBSs
↓	↓	↓
Written communication scaled score	MCQs scaled score	TBSs scaled score
	↓	
	Aggregate score with policy weights applied	
	↓	
	Reported Score (0–99)	

Is there more that I can read about the Examination?

Yes. There are many technical reports related to the psychometric characteristics of the Examination. All of these reports are posted to the CPA Examination Scoring section of the AICPA website. Other useful publications can also be found under the Research and Reports area as well.

For more information about the Uniform CPA Examination,
visit **aicpa.org/cpaexam** or email us at **cpaexam@aicpa.org**.

AICPA®