

The logo for CIMA, consisting of the letters 'CIMA' in a white, stylized font inside a red rectangular box.

Chartered Institute of  
Management Accountants

# Chartered Institute of Management Accountants

---

## CIMA Objective Tests: Understanding Scaled Scores

Last updated November 2015

# Understanding Scaled Scores

---

From **1 November 2015**, every CIMA Objective Test candidate will be provided with a scaled score in addition to the overall pass/fail grade and section level feedback. A scaled score is used to ensure that all candidates receive a score that is comparable across multiple versions of the same exam, potentially taken at different times.

Before scaled scores can be calculated, it is necessary to first collect and then analyse data on how candidates respond to each item; information from these analyses is used to determine how the test scores are scaled for each exam. Item response data for all the CIMA Objective Exam data were collected from the 2nd January 2015 testing window and will be used to determine the scaled scores for candidates who test on or after 1 November 2015. Note that scaled scores will not be issued retrospectively for candidates testing prior to 1 November 2015.

## Why not use raw scores?

A candidate's raw score on an exam (i.e., the sum of the points earned on that exam) reflects both (1) the ability, or knowledge level, of that candidate, and (2) the difficulty of the exam questions. For any single exam, a high-ability candidate would be expected to obtain a higher raw score than would a low-ability candidate. Equally important is the fact that a single candidate would be expected to obtain a higher raw score on an easy exam than he or she would obtain on a more difficult exam. Thus, raw scores have limited meaning for reporting exam performance because the effects of candidate ability and exam difficulty are confounded.

To offer greater flexibility and exam security, there are multiple versions of each of the nine CIMA Objective Tests available at any one time; thus, it is likely that two candidates will see two different sets of questions. When the questions are selected for each exam version, great care is taken to ensure that the level of difficulty is comparable. Nevertheless, minor differences in exam difficulty may remain. If this is the case then a raw score of 35 on one version may not be comparable to a raw score of 35 on a slightly easier version; these two (identical) raw scores do not therefore represent equal levels of underlying candidate ability. It is for this reason that candidates are provided with a scaled score, which does take into account variations in difficulty, rather than a raw score.

## What are scaled scores?

A scaled score always reflects the same performance level or underlying ability, regardless of the exam version that was taken. Thus, when scores from different exam versions are scaled, or transformed, differences in item difficulty between the versions are taken into consideration. The process of scaling is no different from the transformations that are done to scale (or to measure) temperature. Note the following example.

Suppose Monday's afternoon temperature was measured at 35 degrees Celsius (i.e., its "score" is 35). If Tuesday's temperature was measured at 77 degrees, can we conclude that Tuesday was warmer than Monday? No. If different types of thermometers were used (i.e., if the days' scores were based on different "versions" of the measuring instrument), then we cannot make any direct comparison between the scores or temperatures on those two days.

However, that situation changes dramatically if we express the temperatures on the same scale. If we determined that Tuesday's thermometer was measuring temperature in degrees Fahrenheit, then we could easily transform or scale that 77-degree score to the Celsius scale and arrive at a new scaled score of 25. And only now can we conclude that Monday was actually warmer than Tuesday.

It is intuitively obvious that scaling the temperature does not in any way affect the warmth of the day – it only assigns a different value to the temperature as originally measured. Similarly (but perhaps less obviously), scaling exam scores merely assigns a different value to the exam performance of the candidates. Importantly, when scores from different exam versions are scaled, differences in item difficulty between the two versions are taken into consideration. That is, scaled scores are defined so that the same scaled score reflects a similar level of underlying ability and skill, no matter which exam version is administered. For example, if candidates A and B both have a raw score of 35, but candidate B took an easier version of the test, then candidate B would most likely receive a lower scaled score than candidate A. Scaled scores therefore provide a fair representation of how a candidate performs and are comparable across different versions of the exam and also across time.

Note that as scaled scores take into account both the difficulty of the questions and the candidate's raw score, it is not possible, or meaningful, for candidates to calculate their raw scores from their scaled scores and this should not be attempted.

## What does this mean for the CIMA Objective Tests?

From 1 November 2015 CIMA Objective Test candidates will be awarded a scaled score between 0 and 150 with the pass mark set at 100 for each exam. Therefore in order to pass each exam it is necessary to achieve a scaled score of 100 or higher. There are nine Objective Tests and the scaling is determined independently for each of them and therefore scaled scores are not comparable across the exams.

The passing standard, corresponding to a scaled score of 100, for each of the nine exams has been set independently using an industry recognised standard setting procedure where a panel of experts determine the minimum level of competence, or ability, required to pass each exam.

The level of ability corresponding to the passing standard varies across the exams; i.e. the underlying ability level associated with the passing standard for E1 is slightly different than the underlying ability level associated with the passing standard for F1, because different content is being examined and therefore the level of minimal competence to be demonstrated is different.

The standard required to pass each exam (scaled score of 100), however, is the same across all versions of a specific exam regardless of the difficulty of an individual exam version.

Although it is possible that the raw score pass mark may vary slightly depending on the level of difficulty of each version of the exam, the scaled score, and therefore the ability, corresponding to the passing standard will remain consistent for any single exam. Thus, scaling the test scores and setting a passing standard ensure that all candidates are treated fairly.