

What's the Score?

Perhaps one of the most consistently confusing and misunderstood elements of certification programs is actually understanding the meaning of the test score itself. Like many professional licensure and certification examinations, most certification bodies report scores as a “scaled” score. In order to pass the most certification examinations, candidates must achieve a scaled score of some specific number. But what does that number mean and why do certification bodies use scaled scores?

Certification Programs normally adhere to certain best practices that include the development of multiple forms of the examination that are in circulation at any given time. All of these test forms are developed to be as equal as possible, following the same coverage of content topics and level of difficulty. However, despite these efforts what we know about tests is that no two are exactly alike. The level of difficulty on each test form may vary slightly because questions of different difficulty may appear on different forms. Therefore, the passing scores on examinations are adjusted to account for differences in the “degree of difficulty” on each test form.

Whether you realize it or not, the number of questions a candidate needs to answer correctly in order to pass may vary slightly depending on the difficulty of the examination form they are administered. The statistical process by which these adjustments are made is known as *equating* and without getting bogged down in a description of that process, it is mentioned only for purposes of identifying that there is a formal process for adjusting the passing score (score needed to pass) for each exam form in order to make them equal in difficulty. This methodology results in a passing score that is slightly lower for an examination containing more difficult questions and a passing score that is slightly higher for an examination containing easier items. This means that if we only reported a *raw score* (the actual number of questions a candidate got correct) or if we only reported scores as the percentage of questions a candidate got correct, a candidate who was administered a difficult exam form might have passed having answered 70 percent of the questions correct and another candidate who was administered a less difficult form of the examination may have failed having answered 72 percent of the questions correct. As you might imagine the candidate who failed with a score of 72 percent might very well be upset to find out that a friend taking the more difficult exam passed with a score of 70 percent. As more forms are administered, it becomes very difficult to keep track of all the different passing scores required for all the test forms in circulation. So how do we report passing scores in such a way that does not change from form to form? We use a process known as *scaling* to convert each test form to a common scale with a common passing score.

So what does that mean? Scaling a score is simply mathematically transforming a raw test score (the number of questions answered correctly) to any other measurement value. We could use any scale to report the scores. Many of us are familiar with the SAT scaled scoring which ranges from 200-800 on a section or the ACT scaled scoring that ranges from 0-36.

In the case of the National Environmental Health Association (NEHA) we have a possible raw score range of 0 to 225 (225 being the total number of questions) and we

mathematically transform that to a scaled score that ranges from 0-900 with the passing point always set at 650.

Exam	Exam Form A	Exam Form B	Scaled Score
Characteristics	More difficult exam	Less difficult (easier) exam	
Number of Items	225	225	
Failing Scores	Raw scores between 0 and 148	Raw scores between 0 and 154	Scaled scores between 0 and 646
Passing Score	Raw score of 149 (66.22%)	Raw score of 155 (68.89%)	Scaled score of 650
Passing Scores	Raw scores between 149 and 225	Raw scores between 155 and 225	Scaled scores between 650 and 900

Let's assume a difficult exam form (Form A in the table above) has a raw passing score of 149 items or 66.22% and an easier examination form (Form B in the table above) has a raw passing score of 155 (68.89%). For the difficult examination form (Form A) we would then transform the *raw passing* score of 149 to be equal to the *scaled passing* score of 650 (third column in the table above). This results in a failing raw score of 0-148 being transformed to scaled score range between 0 and 646 and a passing raw score of 149-225 being transformed to a scaled score range between 650 and 900.

For the less difficult examination form (Form B) we would then transform the raw passing score of 155 to be equal to a *scaled passing* score of 650 (third column in the table above). This results in a failing raw score of 0-154 being transformed to scaled score range between 0 and 646 and a passing raw score of 155-225 being transformed to a scaled score range between 650 and 900. The end result is that a raw score of 149 on the difficult test and a 155 on the less difficult test are *BOTH equal to a scaled score of 650* because they all have the same meaning in terms of the level of demonstrated knowledge needed to be deemed competent and awarded the certification.

We deal with these types of scaling conversions in our everyday life in areas such as miles to kilometers or perhaps more commonly temperature scales. On the Fahrenheit scale the freezing point of water is 32°, and on the Kelvin scale the freezing point of water is 273.15°. While these are both different numbers they mean the same thing when it comes to the freezing point of water. Now we can add a common scale such as Celsius which uses 0° to indicate the freezing point of water. Much like different passing scores on two exam forms, the temperature of 32° Fahrenheit and 273.15° Kelvin are both equal to 0° Celsius because they all have the same meaning.

In summary a scaled score is not a percentage score but rather a simple transformation of a raw score in order to report comparable test results when each examination form varies in difficulty over time.