# Standard Setting (Cut Score)

Introduction

In this section of the tutorial, our goal is to provide instruction on achieving a defensible passing point. This process is known as **standard setting** and allows us to appropriately determine a pass/fail cutoff score.The subsections are as follows:

1. Introduction to Standard Setting
2. **How to Set up a Standard Setting Meeting**
3. Different Standard Setting Techniques: **Absolute** and **Compromise** Methods
4. **Associated Legal Issues**
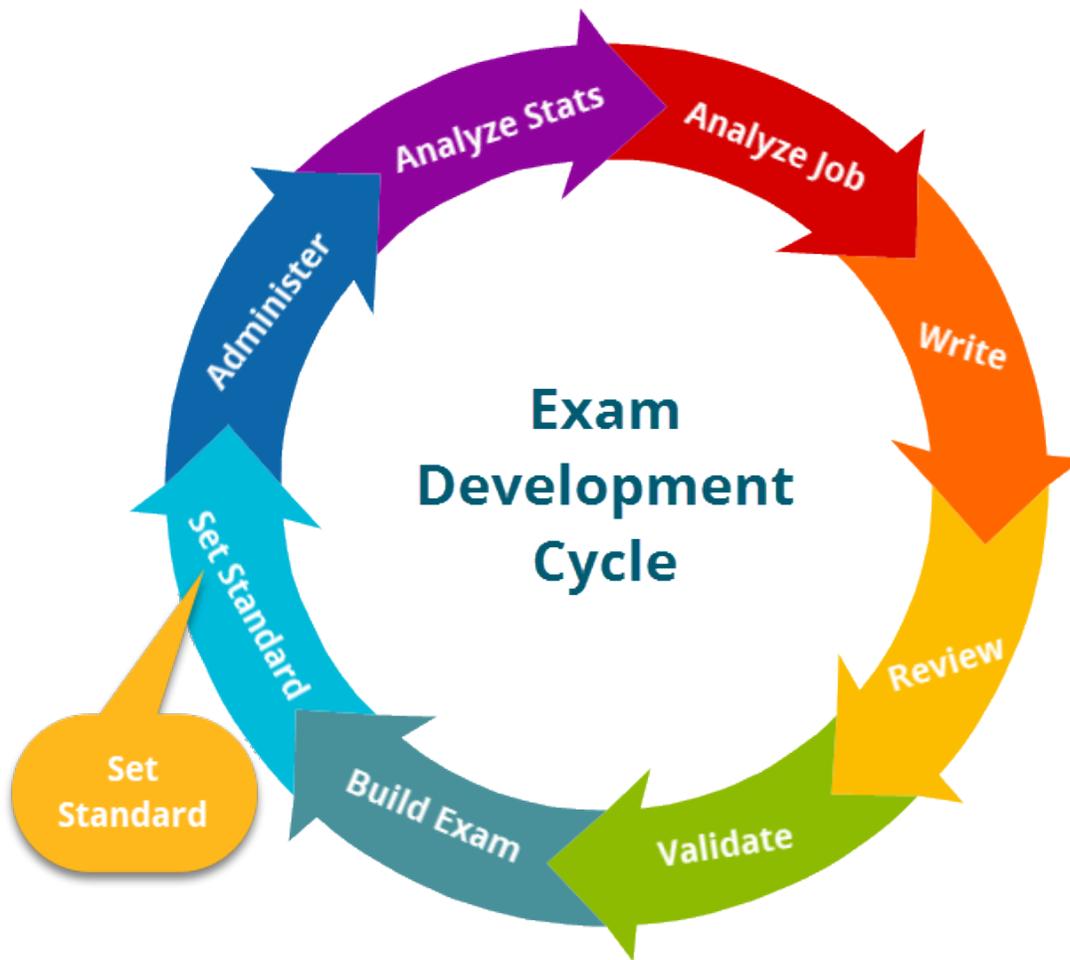
Here we revisit our cyclical model of exam development. Once an **exam has been assembled**, we need to set a standard to determine the appropriate level of difficulty.

> ⊙ **Misconception of Passing Points**
> 70% Cut Score is Arbitrary!

Recall your days in school when teachers set the universal "70%" as the minimum score needed in order to pass an exam. Although this percentage may be sufficient for a grade school science class, using this or any passing score without sound, research-backed methodology is the easiest way to invite legal challenges to your exam program. For all intents and purposes, 70% is just an arbitrary number. It is unlikely that your teachers were psychometrically determining that each test administered was equal in difficulty.

Standard Setting: The Basics

The purpose of a standard setting is to determine the exact point that separates those candidates who know the material and those who do not. We want to be confident that good performance on an exam correlates to high competence on a job, for example. This is a difficult process and is open to much legal scrutiny. In fact, it is one of the most common legal challenges for employment exams.
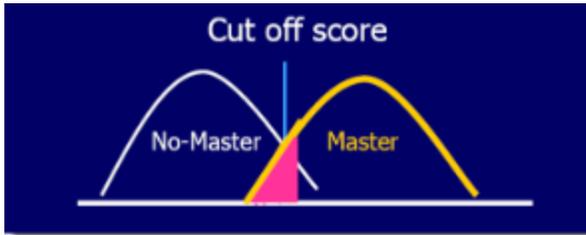
Norm-Referenced vs. Criterion-Referenced Exams

Exams can be categorized into two major groups: norm-referenced exams and criterion-referenced exams. These two types differ in their intended purposes, the way in which content is selected, and the scoring process which defines how the test results must be interpreted.

A norm-referenced exam is a type of exam, assessment, or evaluation which yields an estimate of the position of the tested individual in a predefined population, with respect to the trait being measured. This estimate is derived from the analysis of exam scores and possibly other relevant data from a sample drawn from the population. The term normative assessment refers to the process of comparing one test-taker to his or her peers. Norm referenced exams are designed to highlight achievement differences between and among students to produce a dependable rank order across a continuum of achievement from high achievers to low achievers. Schools might want to classify students in this way so that they can be properly placed in remedial or gifted programs. For this purpose, they are fine, but they are not appropriate for employment exams.

On the contrary, a criterion-referenced exam is one that provides for translating test scores into a statement about the behavior to be expected of a person with that score or their relationship to a specified subject matter. The objective is simply to see whether the student has learned the material. Criterion-referenced exams report how well candidates are doing relative to a pre-determined performance level on a specified set of educational goals or outcomes included in the school, district, or state curriculum.

ⓘ **What's the difference between Norm and Criterion-Referenced Exams?**
Norm Referenced vs. Criterion Referenced
- Norm Referenced exams are used to rank candidates (not recommended for employment exams)
- Criterion Referenced exams are used to determine whether or not the candidates know the subject matter


Norm-referenced methods are based on a comparison among the performances of candidates. Using this method, a set proportion of candidates fails regardless of how well they perform. For example, the top 70% pass. Here is a sample test score distribution that applies the norm referenced method. The red shaded portion that illustrates the lowest 30% of candidates are the pre-determined cutoff point.
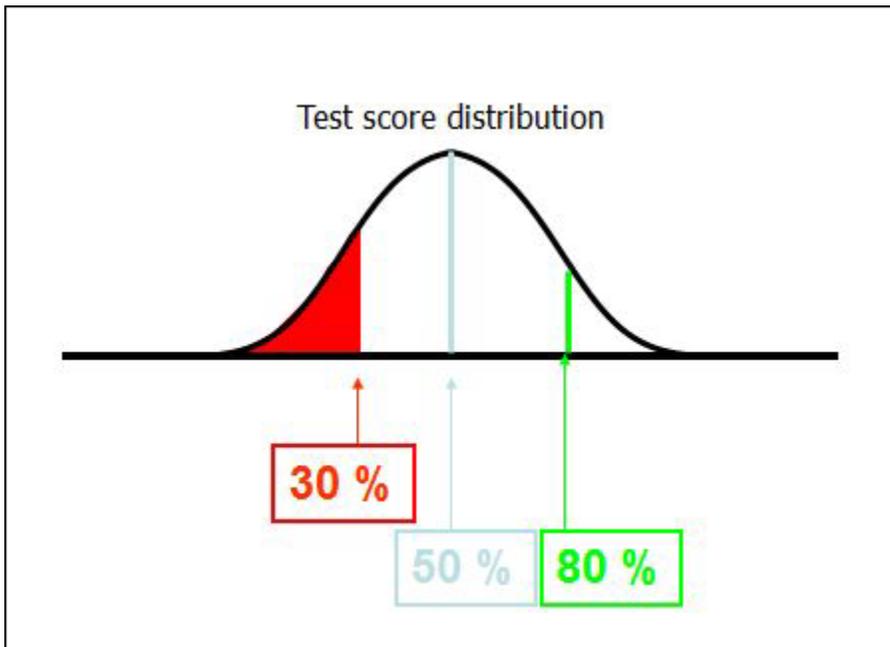
ⓘ **Norm-Referenced Methods**
- Based on a comparison among the performance of all candidates
- A set proportion of candidates fail regardless of how well they perform (e.g., the top 70% pass)

Criterion Referenced exams are used to determine whether candidates know the subject matter or they do not. Here is a test score distribution using the criterion-referenced method. You can see the exact intersection of the group that performed well and the group that performed poorly. This is designed to be the point of minimum competence.
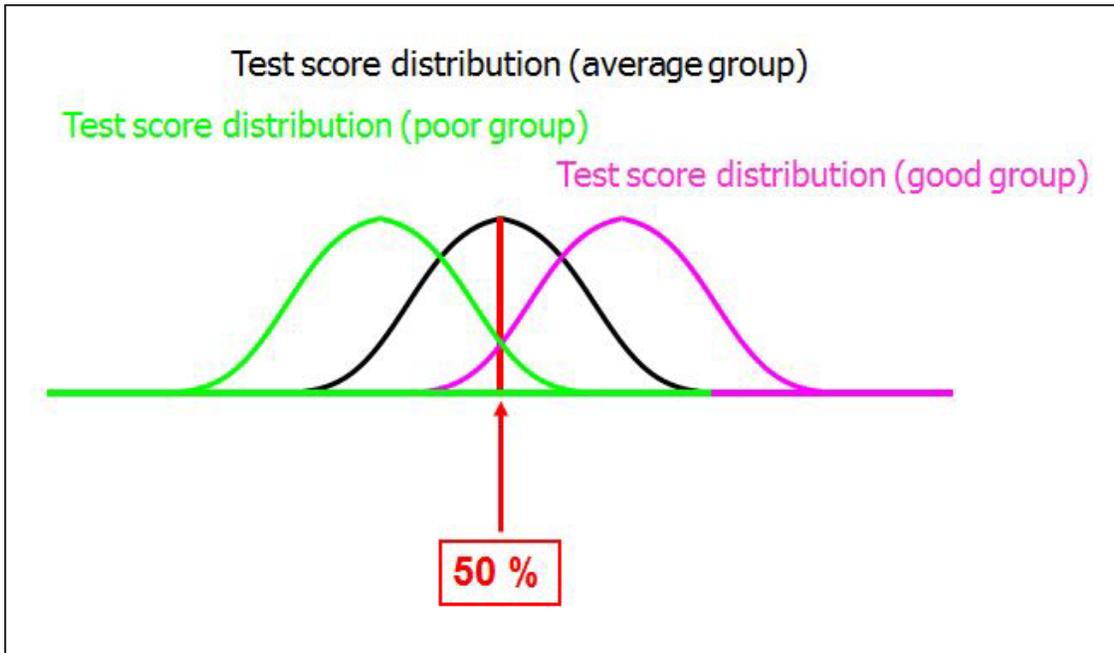
When organizations set a passing standard for their exam, there are several important factors they consider:

What is the difficulty of the exam?

Subject matter experts developing a high-stakes exam usually write items with a wide range of difficulty, from very easy to very difficult. As such, it is advisable to document a quantitative evaluation of the exam difficulty, even if it is based on opinions from subject matter experts.

What is the minimum level of performance that can be considered "competent"?

Observations of employees on the job typically yield a range of performances, from exceptional (the superstars) to inadequate (the incompetent). The goal is to establish the minimum set of knowledge and skills someone needs to possess in order to be labeled "competent" in regards to the job.

What is the experience level of the target population with the material being tested?

Another consideration is the experience those who will be taking the exam have with the exam content. If the exam contains concepts recently identified as important due to a change in job role or expectations, the target population may not have had the experience to effectively master them yet. As a result, the passing rate may be unreasonably low until proper training has occurred. Therefore, it is important to be aware of any new knowledge or skills being tested that would limit the performance of the target population due to experience or exposure.

What is the market demand for practitioners in the organization, industry, or public?

A passing standard set too high produces too few practitioners to adequately meet the needs of the organization, industry, or public. If too few people are meeting the performance standard, it may be advisable to revisit its appropriateness in regards to the exam's objectives.