INTRODUCTION

Standard setting is the process used to select a passing score for an exam. Of all the steps in the test development process, the standard setting phase may be the one most like art, rather than science; while statistical methods are often used in conducting a standard setting, the process is also greatly impacted by judgment and policy. The passing score (also known as the passing point, the cutoff score, or the cut-score) is used to classify examinees as either masters or non-masters. An examinee's score must be equal to or greater than the passing point, in order for that examinee to be classified as a master, or to pass the test. If an examinee is misclassified, that is referred to as a classification error. Typically, the passing score is set at a score point on the exam that the judges determine reflects the minimum level of competency to protect the public from harm or to provide minimal competency at the occupational level being assessed. For the standard setting to be conducted successfully, the panel of judges should be carefully selected and then thoroughly prepared and trained to their task. There are a number of approaches to standard setting, including: informed judgment, conjectural, and contrasting groups methods.

METHODS FOR STANDARD SETTING

Types of Classification Error

The passing score for a test should be set in accordance with the purposes of the exam. It should not be set arbitrarily, but rather should be carefully determined by a panel of judges who are familiar with the content of the exam as well as the characteristics of the occupation concerned. Two types of classification error can occur when the passing score is applied. One type of misclassification is termed a false-positive (i.e., an error of acceptance). An example of a false-positive error would be an examinee who was not minimally competent, but who passed the test. The second type of misclassification is termed a false-negative (i.e., an error of rejection). In this type of misclassification, an examinee who actually has the level of competence fails the test. Depending upon the nature of the exam program, one of these types of errors may be far more problematic than the other. Awareness of these potential consequences may be used to influence the determination of the final passing score, after the panel of judges has made their recommendation. Policy makers at the exam program may adjust that recommended passing point based on other factors, possibly including operational test score data when it becomes available.

Informed Judgment Method

The informed judgment method is a test-based approach. A panel of judges, or stakeholders, reviews the overall test and its content. Based on their holistic reviews, the judges then each suggest a percentage of items on the test that he or she believes ought to be correctly answered by a minimally competent examinee. This percent-correct score on the total test can be viewed as the individual judges' recommended passing score. Once you have these recommended passing scores from the panel you can use them, perhaps along with additional information, to set the final passing score. The informed judgment method might be difficult to rationally defend when it is used in isolation. However, it may be a very appropriate method for use in combination with other methods, particularly the contrasting groups method.

Conjectural (Modified Angoff) Method

The modified Angoff method is the most commonly used of the conjectural methods, all of which are item-based approaches to standard setting. A panel of judges is assembled and is asked to review the test, one item at a time. For each item, each judge gives an estimate of the probability that a minimally competent examinee would be likely to respond correctly. (Alternatively, the judges may be asked to imagine a hypothetical group of minimally competent examinees and then to indicate the percentage of that group who would be likely to respond to the given item correctly.) These item probabilities, or percentages, are averaged across the set of test items for each judge, forming that judge's recommended passing score. Typically, one or more additional rounds of review are then undertaken in order to allow the panel to talk with one another about the reasons for their assigned percentages, and to hopefully move towards consensus. These passing scores are then averaged across the individual judges to arrive at the full panel's recommended final passing score.

Contrasting Groups Method

The contrasting groups method is an examinee-based approach to standard setting. This method in particular requires that the panel of judges be highly familiar with the target test population. The panel of judges identifies a set of examinees who are clearly non-masters and another set of examinees who are clearly masters; borderline examinees are not included. It is especially important that the non-masters be carefully selected. While these non-master examinees would not yet be considered minimally competent in the occupational area, they should nevertheless be members of the target test population. If, instead, the examinees who are identified as non-masters are completely

unknowledgeable in the exam's content area, the passing score may be set at an artificially low point. After the two groups of examinees have been identified, they are then administered the test. The two resulting test score frequency distributions are plotted on the same continuum. The passing score can be set at the intersection point of the two distributions; or, alternatively, the final passing score can be adjusted somewhat, based on the relative cost of false-positive and false-negative classification errors. While the contrasting groups method can be used independently, it may also be used as a complement to the informed judgment or other standard setting method.

Summary

Standard setting is the phase of test development in which the critical decision of setting the passing score is made. This decision will have direct and obvious consequences for examinees, as it will form the basis of defining which examinees pass the test and which do not. Several approaches to standard setting can be used, but in all cases human judgment must be applied. Careful selection and training of the panel of judges is essential for the success of the standard setting.