

Standard Setting Process in Educational Assessment

Educational assessments are used to make a range of decisions which categorize test takers based on their performance. This may involve classifying examinees as “pass” or “fail” or may place them into one of several achievement levels. For example, students taking NYS Grades 3-8 ELA and Mathematics exams are classified as either Level I (Below Standards), Level II (Meets Basic Standards), Level III (Meets Proficiency Standards), or Level IV (Exceeds Proficiency Standards) to establish and communicate achievement goals. The achievement level descriptions define what students should know and be able to do when they have reached each level. Standard setting is the process where the descriptions of these achievement levels and the minimum scores necessary to be classified into each level (i.e., cutscores) are established.

Standard setting panels are made up of NYS educators and are geographically representative of the state. Additionally, educators with diverse roles (e.g., special education, bilingual, etc.) are sought to ensure that all student interests are represented. The panelists are tasked with recommending the minimally acceptable examinee performances for classification into each achievement level. These cutscores, or *standards*, serve as the minimum threshold for classifying examinees into a specific performance level, rather than the next lowest level.

While there are numerous methods for setting performance standards, all include a combination of technical considerations and expert judgment. The standard setting process allows panelists to incorporate their professional knowledge and experience in accordance with the *Standards for Educational and Psychological Testing* (AERA, APA, & NCME, 2014). Panelists receive training that outlines the standard setting process, their specific task, and the materials that they will be using. Multiple rounds of discussion allow panelists to revisit the standards they have previously set and the panelists’ judgments are informed using empirical data to give feedback (e.g., percentage of students classified into each achievement level given a proposed cutscore and minimum, maximum, mean, and median cutscores proposed by the group). Once the predetermined number of rounds is completed, the final cutscores recommendations are established. Based on the recommendations from the standard setting panel and technical advisors, the Commissioner of Education selects the final cutscores for the operational examinations.

Throughout the standard setting process, all rationale and procedures for establishing cutscores are documented as required by the Standards for Educational and Psychological Testing (AERA, APA, & NCME, 1999). This includes information on panelist qualifications, procedures followed for setting performance standards, and the impact of the proposed cutscores on student achievement. This documentation, along with standardized procedures and expert panelists, provides evidence for the defensibility of the final cutscores.

Reference:

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (2014). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.