

# Determination of cut-offs for passing and mastery scores

## Background

The Critical Thinking about Health (CTH) Test includes two multiple-choice questions (MCQs) for each of the nine IHC Key Concepts included in the IHC secondary school resources. There are three response options for each question.

## Objective

To determine cut-off scores for passing (having at least a borderline ability to apply the Key Concepts) and mastery (being at least on the border of having mastered the concepts).

## Methods

Eight people judged the likelihood that someone on the border between passing and not passing, and someone on the border between having mastered and not having mastered the concepts would answer each MCQ correctly. The judges included two curriculum specialists (from Kenya and Rwanda), three secondary school teachers (from Kenya, Rwanda, and Uganda), three health service researchers (from Croatia, Norway, and Uganda). They determined cut-off scores by summing up the probability of answering each MCQ correctly. They were provided with persona describing borderline students, the correct answers, the difficulty of each question based on the Rasch analysis that was conducted to validate the test (Additional 8), and instructions based on a combination of two widely used methods: Nedelsky's and Angoff's.<sup>1, 2</sup>

The Nedelsky method allows judges to eliminate response options that a borderline learner would be able to eliminate. The chances of getting each question correct is then equal to one divided by the number of remaining response options, for example, if there are two remaining response options (one of which is the correct option), the chances of a borderline individual answering the question correct is  $\frac{1}{2}$  or 50%. The resulting cut-off score is then determined by adding up the probabilities for all the questions. With

Angoff's method, which is one of the most widely used, the judges assess the difficulty of each question as a whole.

Using a combination of Nedelsky and Angoff's methods, the judges started with the Nedelsky method, then increased or decreased the assigned probability for each question based on an overall assessment. This gave the judges a logical approach to making an initial judgement about the difficulty of each question. It then allowed them to adjust for uncertainty about the number of response options a borderline individual would eliminate, the difficulty of the stem (scenario) for the question, the difficulty of the concept, anything else that might make a question more or less difficult, and the difficulty of the question based on the Rasch analysis.

The judges answered the 18 questions without looking at the answers to get a sense of the difficulty of the questions. They discussed the instructions, undertook a practice round of assessing some of the MCQs, and then discussed their judgements and agreed on guidance for assessing the MCQs. They then independently assessed all 18 MCQs. Their assessments were summarised and discussed before reaching a consensus on the cut-off scores using a nominal group technique.

## **Results**

The consensus judgements varied from 40 to 55% for the percentage of students on the border between passing and not passing who would answer each question correctly. They varied from 60% to 85% for students on the border between having mastered and not having mastered the concepts. The judges agreed that 9 out of 18 questions needed to be answered correctly to pass and 14 out of 18 questions needed to be answered correctly to demonstrate mastery.

## **Conclusion**

Although there was wide variation in many of the individual judgements, it was possible to reach a consensus on the cut-off scores for passing and mastery in an online meeting that lasted less than 90 minutes.

## **References**

1. Livingston SA, Zieky MJ. Passing scores: A manual for setting standards of performance on educational and occupational tests. Princeton, NJ: Educational Testing Service; 1982. <https://eric.ed.gov/?id=ED227113>
2. Angoff WH. Scales, Norms, and Equivalent Scores. Princeton, NJ: Educational Testing Service; 1984.

[https://www.ets.org/research/policy\\_research\\_reports/publications/book/1984/ibmt](https://www.ets.org/research/policy_research_reports/publications/book/1984/ibmt)

3. Jones J, Hunter D. Consensus methods for medical and health services research. *BMJ*. 1995;311(7001):376-80. <https://doi.org/10.1136/bmj.311.7001.376>