

AN ADAPTIVE ANGOFF METHOD FOR COMPUTER ADAPTIVE TESTS

Tilman Sheets, Danielle B. Allen, Drake Doumit, Reagan Girardot
Louisiana Tech University



ABSTRACT

A revised Angoff method loosely based on the Ordered Item Booklet Angoff method with some inspiration from computer adaptive testing is proposed. The new method was used to speed up the rating process and reduce fatigue on raters. To the authors' knowledge, this method has not been used before.

THE PROBLEM

Our I-O psychology graduate program's consulting group, AROS, contracted with a large energy and chemical company to assist in **reviewing their selection process** for four technician positions. The goal was to create a multiple hurdle process that placed the least expensive selection devices early in the process, followed by the more expensive selection devices such that the **candidate pool is potentially reduced to only include those who are qualified for the position.**

In the testing phase of the process, applicants are required to take a **popular mechanical comprehension test (MCT)** and a position-specific knowledge test. **SMEs were recruited** to assist in identifying cutoff scores for these two tests that could be used to eliminate unqualified applicants from the selection process. SMEs participated in a focus group in which their task was **to rate items for each test** in regard to their belief that a minimally-qualified, newly-hired worker could successfully answer the item.

The MCT asks test takers to solve questions that are related to hydraulics, heat, electricity, acoustics, and other mechanical or physics-related phenomena. A score obtained on the MCT is not simply a total number of correct responses but rather a **calculated estimate of ability**, making the MCT a **computer adaptive, item-banked test**. The test bank is composed of a total of 300 possible test items, but testers are only presented with 55 items during each testing session. Each test item is assigned a difficulty level (easy to highly difficult), and an equal number of items from each difficulty level are presented to each tester to ensure fairness.

Due to the fact that the MCT is an item banked test, and because of the time required to review the large number of items, **asking each SME to review all 300 items was less than ideal**. Therefore, the traditional Angoff method was revised in order to establish a cutoff score.

THE SOLUTION

The test items were arranged and presented to SMEs in order from easiest to most difficult based on the amount of ability needed to answer each item (provided by the publisher).

1. The first item that was presented to the SMEs was at the point on the test at which **50% of test-takers respond correctly**.
2. For each item, the SMEs were asked to indicate whether a **minimally qualified employee** would have the ability to answer the item correctly (0 = no, 1 = yes).
3. Each SME first **rated the item independently**. For any items showing rating discrepancies (80% agreement or less), the SMEs were asked to **explain to the group** why they rated the item the way that they did.
4. After the discussion, each SME was asked to provide a **new rating for that item**.
5. SMEs were directed to move **backward** (to easier items) or **forward** (to more difficult items), based on their ratings of the previous item.
6. If a **"yes" response occurred for 3 consecutive items** at the 50th percentile, we then **skipped forward** and asked the SMEs to rate an item at the 60th percentile.
7. If a **"no" response occurred**, then we **skipped backward** such that the item directly before that item was presented.
8. This process was repeated until reaching the point (3 consecutive items) that a minimally qualified employee would not be able to answer the items correctly.

IMPLICATIONS AND NEXT STEPS

- The introduction of new technology over the years has changed employment testing and led to the development of methods such as item banked and computer adaptive tests
- Although these methods provide increased test security and allow for continuous improvement of test items, establishing cutoff scores for large item-banked tests can prove challenging
- The current method is proposed to reduce the number of items that must be rated by SMEs by considering the item's calculated difficulty scores (thus, reducing potential rater fatigue)
- SMEs had positive reactions to the method and the method reduced the number of items that had to be presented to SMEs in order to establish the cutoff score
 - For example, in one session SMEs only had to rate 45 out of 300 of the items to establish the cutoff score
- In the future, the established cutoff scores will need to be validated by examining the relationship between employee performance and test scores

Table 1

An example of a master spreadsheet with recordings from 7 different items from 5 raters. Difficulty levels are shown for each item in the form of percentiles.

Item #	Rater A	Rater B	Rater C	Agreement	%tile
227	1	1	1	1.000	50
228	1	1	1	1.000	51
229	1	1	1	1.000	52
↓ Items 230-240 are skipped ↓					
241	1	1	1	1.000	57
242	1	0	1	0.750	58
243	0	0	0	0.250	58

Green:
 Many "Yes" responses in a row.
 Skip to more difficult items.

Yellow:
 Raters disagree.
 A discussion is needed.

Red:
 Many "No" responses in a row.
 Skip back to less difficult items.