The Magic Number: A Look at What Goes Into an ARRT Scaled Score of 75

Ben Babcock, Ph.D.
Psychometrician, ARRT

Nance Cavallin, B.A., R.T.(R)(T)(ARRT)
Senior Exam Development Coordinator, ARRT

June 28, 2012
Where Can I Find This Presentation?

www.arrt.org/research
Where Are We Going?

1. Practice analysis
2. Exam construction
3. Exam scoring / scaling
4. After the test
Do You Do This Study Anymore?

OF COURSE YOU DON’T!
Practice Analysis

• Helps ARRT keep up with modern medical imaging

• Goal 1: Create an inventory of “typical” R.T. tasks

• Goal 2:
  - Exam Content Specifications
  - Task Inventory
  - Clinical Requirements

• How does ARRT decide on the tasks?
  - PA Advisory Committee + Survey Data
PA Surveys: Who, and How Often?

- Task survey primarily targets entry-level R.T.s.
  - Primary: 1-3 years experience
  - Post-Primary: 1-5 years experience

```
Primary   ≈ 3-Year Interim
          ≈ 6-Year Comprehensive

Post-Primary ≈ 3-Year Interim
              ≈ 9-Year Comprehensive
```
PA Data and the PA Committee Notes

- PA Advisory Committee = R.T.s, educators, physicians, and physicists
- ≥40% responsible tasks make the task inventory
- Committee may include infrequent but critical tasks
- The R.T. community comments on the draft content specifications and clinical requirements (arrt.org)
- The ARRT Board of Trustees has the final approval
Next Step: Compiling the Exam

- Exams come from a large pool of questions (items)

- Item writers: ARRT-trained R.T.s who write questions

- Only “tried and true” items compose the score
What Does the Exam Look Like?

- Exam content closely follows the content specifications from the PA.
### What Does the Exam Look Like?

<table>
<thead>
<tr>
<th>CONTENT CATEGORY</th>
<th>PERCENT OF TEST</th>
<th>NUMBER OF QUESTIONS</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Radiation Protection</td>
<td>22.5%</td>
<td>45</td>
</tr>
<tr>
<td>B. Equipment Operation and Quality Control</td>
<td>11.0%</td>
<td>22</td>
</tr>
<tr>
<td>C. Image Acquisition and Evaluation</td>
<td>22.5%</td>
<td>45</td>
</tr>
<tr>
<td>D. Imaging Procedures</td>
<td>29.0%</td>
<td>58</td>
</tr>
<tr>
<td>E. Patient Care and Education</td>
<td>15.0%</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>100%</td>
<td>200</td>
</tr>
</tbody>
</table>
B. EQUIPMENT OPERATION AND QUALITY CONTROL (22)

1. Principles of Radiation Physics (9)
   A. X-Ray Production
      1. source of free electrons (e.g., thermionic emission)
      2. acceleration of electrons
      3. focusing of electrons
      4. deceleration of electrons
   B. Target Interactions
      1. bremsstrahlung
      2. characteristic
   C. X-Ray Beam
      1. frequency and wavelength
      2. beam characteristics
         a. quality
         b. quantity
         c. primary versus remnant (exit)
      3. inverse square law
      4. fundamental properties (e.g., travel in straight lines, ionize matter)

D. Components of Digital Imaging (CR and DR)
   1. PSP - photo-stimulable phosphor
   2. flat panel detectors - direct and indirect
   3. start up and shut down
   4. CR plate erasure
   5. equipment cleanliness (imaging plates, CR plates)

E. Types of Units
   1. dedicated chest unit
   2. tomography unit

F. Accessories
   1. stationary grids
   2. Bucky assembly
   3. image receptors

3. Quality Control of Imaging Equipment and Accessories (4)
   A. Beam Restriction
      1. light field to radiation field alignment
      2. vertical alignment
What Does the Exam Look Like?

• Exam content closely follows the content specifications from the PA

• Why is this important? VALIDITY!

• ARRT staff maintains comparable exam difficulty from year to year

• ARRT launches “fresh” new exam forms often
How many items compose a Radiography Exam score?

A. 100
B. 120
C. 200
D. 220
Pop Quiz!

How many items compose a Radiography Exam score?

A. 100  
B. 120  
C. 200  
D. 220

There are 220 total items, but 20 are unscored, untried pilot test items designated before the exam launch.

The pilot items do not influence an individual’s score in any way!
Why Have Pilot Items?

Make sure items have good stats before they “count”
• Too easy / too difficult?
• High scorers answer the item correctly?

Examinees may comment on items before they “count”
• Non-statistical check of item functioning
• Exam committees review all relevant comments
Who Looks at the Exams?

• Staff R.T.s and psychometricians compile the initial exam forms

• Exam committees of R.T.s, educators, physicians, and physicists review every item on every exam... EVERY TIME!

• Staff R.T.s and psychometricians QC all final forms

• ARRT’s exam security team monitors the exams to protect those valuable items

1 operational item ≈ $1,000
The Cost of Operational Items

- **PRE-EXAM COMMITTEE COSTS**
  - 342 items
  - $23,384

- **EXAM COMMITTEE COSTS**
  - 172 items
  - $9,554

- **SELECTED TO PILOT**
  - 42 items
  - $1,734

- **PSYCHOMETRIC REVIEW**
  - 34 items
  - TOTAL: $34,672

Though we started with 342 items, only 34 survived!
Exam Scoring and Scaling

- ARRT must determine the minimum number of correct responses required to pass

- Standard setting: experts rate items to determine the minimum passing score
  
  *Infrequent, but occasionally necessary*

- Equating: making statistical adjustments to keep test scores equitable over time
  
  *Frequent*
Raw Scores to Scaled Scores

- Scale: a numbering system
- Raw Score: total number correct
- Scaled Score: a raw score placed on a numbering system that is consistent over time
Tea Time!

Sir Isaac Newton:
“My water is 36 degrees Newton.”

Lord William Thomson, 1st Baron Kelvin:
“My water is 366 Kelvin.”

Whose water, if any, is boiling?

Let’s SCALE the measurements to °C.
36 °N ≈ 109 °C
366 K ≈ 93 °C
Raw Scores to Scaled Scores

- Scale: a numbering system
- Raw Score: total number correct
- Scaled Score: a raw score placed on a numbering system that is consistent over time

- Why does ARRT use scaled scores?  
  Scaled scores → *fair scores among different forms*

  1 = minimum, *75 = passing*, 99 = maximum
A scaled score of 75 is NOT 75% correct!!!!
Raw to Scaled: The Hinge
Anchor #2: Equated raw passing score and scaled score of 75

Anchor #1: Max raw score of 200 and scaled score of 99
The second form is easier, so the raw score required to pass is higher.
After the Test Site

- Weekly score QC
  - Pearson VUE and ARRT independently score every testing event
  - Scores at the test site are preliminary. Why?

- Review of exam and item comments
- Review of item statistics
- Annual statistical reports
What Does This Mean For Me, an R.T.?

• Use the content specifications to guide study efforts for ARRT exams.

• You will properly interpret what scaled scores mean.
  ☺ Your score $\geq$ 75
  ☹ Your score $<$ 75

Scaled scores are NOT percentage correct scores!
What Does This Mean For Me, an R.T.?

WE NEED R.T. PARTICIPATION!!!

• PAs: Fill out and return PA surveys, comment on the draft documents, volunteer for the committee

• Examination: volunteer to be an item writer, volunteer to be an exam committee member, create an environment of exam ethics

• Standard setting: requires a large number of R.T. volunteers to set or re-set a new base standard
Some Useful Websites

https://www.arrt.org/Content-Specifications

https://www.arrt.org/Publications/Brochures

https://www.arrt.org/Practice-Analisis/

https://www.arrt.org/Examination/Statistics

https://www.arrt.org/Examination/Volunteers
Questions?
Thank You!