

ARRT certifies and registers more than a quarter-of-a-million radiologic technologists in a range of modalities

Primary Eligibility Pathway

Radiography

Nuclear Medicine Technology

Radiation Therapy

Sonography

Magnetic Resonance Imaging

Post-Primary Eligibility Pathway

Cardiovascular-Interventional Radiography

(test replaced 1/03 with CI and VI)

Mammography

Computed Tomography

Magnetic Resonance Imaging

Quality Management

Sonography

Bone Densitometry

Vascular Sonography

Cardiac-Interventional Radiography

Vascular-Interventional Radiography

Breast Sonography

Radiologist Assistant



ARRT

American Registry of
Radiologic Technologists

1255 Northland Drive
St. Paul, MN 55120-1155
Phone (651) 687-0048

www.arrt.org

© ARRT 2006



American
Registry of
Radiologic
Technologists

*You've Taken
Your ARRT Exam.
Now It's Time to...*

Settle the Score

*For more information on
ARRT, visit our web site
at www.arrt.org*

*Promoting high standards of
patient care by recognizing
qualified individuals in medical
imaging, interventional procedures,
and radiation therapy*

Whew! You've completed the ARRT exam. Now you can settle back and wait for the results to arrive sometime in the next four weeks. In order to fully appreciate what goes into determining your score, check out all the science that's going on in the background.



Leveling the Playing Field

ARRT regularly revises its exams; so different examinees at different times may be taking slightly different tests. What does this mean for your score? Might you have done better (or worse!) on another form?

Correct answer: the test form doesn't matter. ARRT scoring is structured to compensate for any differences in difficulty between forms of a test.

Questions for each version of a given test are all drawn from the same large pool. But the underlying content of the tests is virtually identical. Content specifications — which determine the topics and the number of questions on each topic — ensure that all exams test essentially the same knowledge.

Still, in some cases, “different” might indeed translate into slightly more or less difficult. But the science of testing a person's knowledge — called psychometrics — enables ARRT scores to factor in any differences in difficulty.

How does it work? “Sophisticated software” is the short answer. Through comparing the two tests, specific levels of difficulty are determined, and scoring is adjusted to take those differences into account. The net result: a level playing field.

Why are the exams always changing? There are two primary reasons. First, the “state of the art” is constantly evolving through ongoing innovation and invention. Those influences are reflected in the job analysis conducted by ARRT. Second, security is an important factor in the integrity of ARRT exams. Constantly adjusting the order and content of test questions ensures that unscrupulous persons cannot cheat.

‘Scaling’ the Heights, Making the ‘Cut’

Your test performance is reported as a scaled score — somewhere on a range from 1 to 99. These aren't percentage points; these are scaled scores. Nor are they built on “curves,” where a certain percentage would pass and a certain percentage would fail.

A scaled score of 75 is required to pass all ARRT exams. This pass-fail point, called the “cut score,” is established by ARRT's Board of Trustees through a process called standard setting. During the standard-setting process, the Board receives input from a

panel representing a broad constituency of radiologic technologists. The cut score represents the standard of performance required to obtain certification. Those who exceed the standard pass the exam.

In recent years, a scaled score of 75 has required correct answers on 65 to 70 percent of the questions. The exact number for each form is based on the difficulty of that form.

Balancing the ‘Scale’

The process ARRT uses to adjust for differences in exam difficulty is known as “calibration and scaling.” The process results in scores that are comparable regardless of which form was taken. This process is well documented and used in a number of different testing programs. Here's an overview of how the process is applied.

Assume that two different forms of an exam are being taken — one by Group A and one by Group B. It turns out that Group A scored an average of 3 points higher than Group B. The question is, “Why?”

Perhaps Group A is more knowledgeable than Group B? Possibly the form taken by Group A is easier than that taken by Group B? Or, maybe it's both?! Without more information, it's impossible to know how to adjust the scores to take these influences into account.

By design, both exam forms are constructed to have a set of questions in common, called “anchor questions.” By analyzing how candidates from both groups respond to the anchor questions, we can evaluate the level of candidate knowledge for each group. After determining each group's knowledge level, it is possible to correctly calculate the difficulty of each question that appears on either form. This is the calibration stage of the process.

How can tests with different questions be scored fairly? Different forms of an exam have a set of questions in common, called an “anchor set.” Performance on the anchor set indicates how each test form needs to be adjusted in order for scores to be comparable across forms.

Once all the items on different forms have been calibrated, we can evaluate the difficulty of each form. At this point, we are able to determine how best to adjust scores in order to compensate for differences in difficulty between forms. Any necessary adjustments are made during the scaling stage.

The result is two different forms of a test that can be used to place candidates on a comparable scale. It really doesn't matter which test form an examinee takes.

Scoring Section-by-Section

Your total score is above the cut: you pass.
Your total score is below the cut: you don't.

But your score isn't just a question of pass or fail. It is also structured to reveal, by content area, your strengths and weaknesses. That's information you can use in devising your own continuing education program.

Section scores range from 1 to 10, at one-tenth point intervals. On one section of a test, you might get 8.1; on another section, 8.6. This 1 to 10 point range is intentionally narrow, because the section scores are often based on a small number of test questions. Consequently, they're not as reliable as total scores and should be interpreted with caution.

If you add up your section scores, will the total be the same as your total scaled score? Yes, if you calculate a weighted average to balance the number of items in each section, allowing more weight to larger sections in determining the total score.