



Standard Setting Report: Magnetic Resonance Imaging - Effective January 2020

Background

The mission of the American Registry of Radiologic Technologists (ARRT) is to “promote high standards of patient care by recognizing qualified individuals in medical imaging, interventional procedures, and radiation therapy.” The ARRT’s equation for excellence states that excellence equals education plus ethics plus examination; standard setting is one of many processes within the examination component that ensure it is an accurate reflection of the knowledge, skills, and abilities required of entry level technologists.

This report details a standard setting conducted in September 2019 for magnetic resonance imaging, including committee composition, methods, results, recommendations, and any changes to the exam cut score. It is ARRT’s primary goal for the exam to reflect the current state of practice and expectations for entry-level technologist. Therefore, this meeting served to update those expectations from the previous standard setting in 1995.

The ARRT utilizes experts in standard setting, called psychometricians, to train and facilitate a committee of subject matter experts from the field to define expectations, collect data, and make recommendations before presenting the results to the ARRT Board of Trustees. The facilitators provided training throughout the meeting to ensure that the committee was prepared to hold productive discussions, make well-reasoned judgments, and provide suitable recommendations at the meeting’s conclusion.

Facilitators:

- Ben Babcock, Ph.D., Supervising Senior Psychometrician at ARRT
- Tim Walker, Ph.D., Psychometrician at ARRT
- Zachary Siegel, Ph.D., Cognitive Scientist for Assessment Design at ARRT

The ARRT Board of Trustees reviewed the results of the standard setting meeting and committee recommendations before approving the final standard. ARRT psychometrics staff will ensure the passing threshold for all exams administered on or after the effective date reflect that prescribed level of performance.

Committee Composition

ARRT staff selected individuals from the volunteer database with the goal of maximizing diversity in role, geography, and experience in magnetic resonance imaging. When possible, ARRT will bias the volunteer pool towards individuals early in their career as the exam is designed to assess candidates at entry level. In addition, the radiologist assigned to the exam committee by the American College of Radiology is invited to attend. In total, 16 subject matter experts participated in the standard setting meeting. Please refer to the following table for specific demographics of this group.



Table 1. Committee Demographics

Rater	Role	Location	Credentials
A	Technologist	CO	R.T.(R)(CT)(MR)(BD)(ARRT)
B	Technologist	TX	R.T.(R)(MR)(ARRT)
C	Technologist	NE	R.T.(R)(MR)(CT)(ARRT)
D	Technologist	MS	R.T.(R)(MR)(ARRT)
E	Technologist	NC	R.T.(R)(MR)(ARRT)
F	Technologist	IL	R.T.(R)(MR)(ARRT)
G	Technologist	PA	R.T.(R)(MR)(ARRT)
H	Technologist	WY	R.T.(R)(CT)(MR)(BD)(ARRT)
I	Technologist	CA	R.T.(R)(MR)(ARRT)
J	Technologist	MO	R.T.(MR)(ARRT)
K	Technologist	FL	R.T.(R)(MR)(ARRT)
L	Technologist	IL	R.T.(MR)(ARRT)
M	Technologist	OK	R.T.(R)(MR)(ARRT)
N*	Technologist	AK	R.T.(R)(CT)(MR) (QM)(ARRT), MRSO
O	Physicist	MA	Ph.D.
P	Radiologist	MN	MD

*Committee member N was also serving as an ARRT Trustee at the time of the meeting

Minimally Qualified Candidate

After training regarding the purpose and implications of standard setting, the committee discussed the knowledge, skills, and abilities expected of an entry level technologist with primary focus on the minimum qualifications that should be demonstrated to earn an ARRT credential. This discussion of the “minimally qualified candidate,” who possesses only the knowledge, skills, and abilities required for certification, is important because it allows the committee to come to a common understanding of what is required for the role prior to any data collection activities. Note that “entry level” and “minimally qualified” are not interchangeable terms. Entry level individuals are early in their career with limited clinical experience irrespective of their level of qualification.

Through their discussion, the committee created a list of generic and discipline-specific examples of knowledge, skills, and abilities that are representative of well qualified, minimally qualified, and not yet qualified candidates.

Modified Angoff

The committee performed a modified Angoff activity (Angoff, 1971) using a recently retired exam form. The facilitator provided training to explain the function and intent of the Angoff to the committee and the committee performed a practice activity with a few items to familiarize themselves with the software.

In the first round of the full activity, committee members read each item on the form and provided their judgment for the percentage of minimally qualified candidates that should answer the item correctly. No additional information was provided during this round.

After the first round, committee members were each provided with feedback regarding their own ratings. Specifically, the facilitator determined each individual’s cut score based on the first round before comparing each judgment to the expected percent correct for the individual’s cut. The facilitator then provided each committee member with a personalized mix of approximately twenty items that were either rated too high, too low, or close to the expected percent correct for their personal cut score.



The facilitator then provided training for the committee regarding the next round of ratings. Namely, that committee would be able to review the items again with their individual feedback as well as the overall proportion correct for first-time candidates. This training also included a discussion of first-time candidate score distribution and the differing implications of common academic scores (e.g., A, B, C, D, F) and binary pass/fail certification exam results.

Finally, the committee performed a second round of the activity with their first-round judgements provided for them in the response window. Committee members could keep or modify any judgements they desired during this round.

Table 2. Modified Angoff Results

<u>Percent Correct Cut</u>	<u>Round 1</u>	<u>Round 2</u>
Mean	63.7	64.0
Minimum	28.5	33.7
Maximum	81.4	77.7
Standard Deviation	12.9	11.9

Hofstee

After the Angoff activity, the facilitator provided training for the Hofstee activity (Hofstee, 1983), directing the committee away from item-level decisions and encouraging them to think more globally about the form they had just reviewed. This training included an explanation of statistical error, impact data, and reasons why the committee may wish to adjust their results higher or lower.

The Hofstee activity consisted of four questions with additional explanations that mirrored the training.

1. What is the lowest acceptable cut score? If a majority of candidates failed the exam, and you were pressured to lower the cut score to permit more candidates to pass, what would be the lowest cut score that you would feel comfortable with? In other words, how low could the cut score be without doing a disservice to the public and the profession (0% - 100%)?
2. What is the highest acceptable cut score? If nearly every candidate passed the exam, and you were pressured to raise the cut score to prevent the exam from being criticized as a trivial standard, what would be the highest cut score that you could live with? In other words, how high could the cut score be without being unfair to candidates and the programs that prepare them (0% - 100%)?
3. What is the lowest acceptable pass rate? Irrespective of the cut score, what is the lowest pass rate that you would be comfortable with (0% - 100%)?
4. What is the highest acceptable pass rate? Irrespective of the cut score, what is the highest pass rate that you would be comfortable with?

Table 3. Mean Hofstee Responses

Minimum cut score	65.8%
Maximum cut score	75.0%
Lowest pass rate	68.4%
Highest pass rate	81.1%



Beuk

The global ratings activity included the Beuk method (Beuk, 1984) alongside the Hofstee method. Committee members answered two questions related to their global expectations for the exam. These questions included additional explanation that mirrored the training.

1. Given everything you know about the content of the exam, and the level of knowledge required for competent performance in a clinical setting, what percentage of the questions should a candidate answer correctly in order to pass the exam (0% - 100%)?
2. Given your knowledge of candidates taking this exam, what percentage of them should pass (0% - 100%)?

Table 4. Beuk Results

Mean cut score	70.3%
SD cut score	4.1%
Mean pass rate	75.2%
SD pass rate	4.9%

Item Plat

Prior to the meeting, ARRT psychometric staff grouped similarly difficult items from different sections of the exam and placed those groups (plats) into a binder in difficulty order. 15 plats of 6 items were necessary to cover the difficulty range likely to be deemed acceptable by the ARRT Board of Trustees. During the meeting, the facilitator provided each committee member with a copy of the plat binder as well as instruction regarding the goals and methods of the activity.

Committee members individually reviewed the plats and selected two; the first being the plat where they believed that 50% of minimally qualified candidates would answer the items correctly and the second was to be the point where they believed that 66% of minimally qualified candidates would answer correctly. The facilitator used the flag (50% or 66%) and plat difficulty to determine two potential cut scores per individual. The following table states the activity results as a percent correct cut for the form used in the Angoff activity to allow comparisons between activities.

Table 5. Item Plat Activity Results

<u>Percent Correct Cut</u>	<u>50% Flag</u>	<u>66% Flag</u>
Mean	65.3%	65.4%
Minimum	54.6%	61.8%
Maximum	72.2%	72.3%

Final Discussion

After all data collection was complete, the standard setting committee reviewed the activity results along with an estimated pass rate for potential cut scores within that range. The committee discussed the results, their impressions of the activities, and the cut score they wished to submit to the ARRT Board of Trustees for review. After the discussion, each committee member submitted their final recommended cut score, and the mean of those recommendations was submitted as the overall committee recommendation.

Table 6. Final Recommended Cut Score

Median	67.0%
Minimum	65.0%
Maximum	70.0%



New Standard and Implementation

The ARRT Board of Trustees reviewed the results and discussed the impact of potential new standards before approving a final standard for the Magnetic Resonance Imaging exam.

The board elected to renew the current standard (equivalent to 133 out of 200 items on the exam form used for this meeting). This standard will remain in place until at least 2026, when the next standard setting is scheduled to take place.

References

Angoff, W.H. (1971). Scales, norms, and equivalent scores. In R.L. Thorndike (Ed.), *Educational Measurement* (2nd ed., pp. 508-600). American Council on Education.

Beuk, C. H. (1984). A method for reaching a compromise between absolute and relative standards in examinations. *Journal of Educational Measurement*, 21(2), 147-152.

Hofstee, W. K. (1983). The case for compromise in educational selection and grading. In S.B. Anderson & J.S. Helmick (Eds.), *On educational testing*, (pp.109-127). Jossey-Bass.



Appendix A

Generic Candidate Description Worksheet

At a high level and applicable to all disciplines, this document describes the knowledge, skills, and abilities of clearly passing, clearly failing, and just barely passing candidates. It includes generic examples of items that they would likely get correct or incorrect.

Individual Standard Setting Committees should refine the criteria for the Just Barely Passing candidate specific to their discipline. In doing so, committee members may want to think of entry level people that they know from the workplace or students that they have taught.

Type of Candidate	Typical Knowledge, Skills, and Abilities*	Examples of items that they should answer correctly	Examples of items that they may answer incorrectly
Clearly Passing (way above cut score)	<ul style="list-style-type: none"> • Knows more than what is on the exam • Obtains high quality images on all types of patients • Readily identifies common pathologies • Demonstrates problem-solving (critical thinking) skills • Demonstrates strong oral and written communication skills • Adapts positioning or projections as needed for difficult patients • Anticipates needs of the radiologist • Collaborates with healthcare team to provide optimal patient care 	<ul style="list-style-type: none"> • Fundamental concepts of imaging and patient care • Detailed anatomy (e.g., parts of organs vs whole organs) • Common pathology and physiology • Concrete understanding of how equipment works • Presence and cause of artifacts 	<ul style="list-style-type: none"> • Advanced physics • Unusual anatomic variants • Pharmacology • Advanced calculations/formulas • Difficult item types (multiselect, hot spot, ordered list) • Items with multi-level thinking • May blank out on a simple question (random error)



Type of Candidate	Typical Knowledge, Skills, and Abilities*	Examples of items that they should answer correctly	Examples of items that they may answer incorrectly
Just Barely Passing (just above cut score)	<ul style="list-style-type: none"> • May exhibit test anxiety • Clinical skills greater than or equal to classroom/book knowledge • Possesses limited clinical or work experience (esp. post-primary exams) • Understands basic anatomy and physiology • Can discriminate between normal and abnormal anatomy • Limited knowledge of study indications • Practices quality patient care and safety (not dangerous) • Self-evaluates abilities and knows when to get help • Collaborates with the healthcare team • Unable to explain the “why” (equipment, instrumentation, tube angulation, etc.) 	<ul style="list-style-type: none"> • Patient care • General patient safety • Modality-specific safety guidelines and procedures (radiation, MR, etc.) • Basic image analysis • Definitions • Basic physics • Presence of artifacts • Basic image production items (depending on clinical or lab experience) 	<ul style="list-style-type: none"> • Physics beyond basics • Detailed anatomy (e.g., parts of organs vs. whole organs) • Calculations • Physiology • Cause of artifacts • Correlation with other modalities • Regulatory requirements • QA/QC requirements • Negatively worded items
Failing (below the cut score)	<ul style="list-style-type: none"> • Demonstrates poor test taking skills • Barely passed other exams (post-primary candidates) • Unable to modify protocols from anatomical programming (button pusher) • Unable to adapt positioning or projection based on patient needs or equipment locations • May not understand indications for exams • Weak anatomy and pathology knowledge • Does just enough to meet requirements • Does not know their own limitations • Does not take responsibility for actions • Language barrier 	<ul style="list-style-type: none"> • Basic positioning • Patient care (general concepts) • Consent for exams • Basic modality-specific protection (radiation, MR, etc.) • Basic anatomy only 	<ul style="list-style-type: none"> • Almost anything • Alternate ways to do a study • How the equipment works

*Certain items in the list of knowledge, skills, and abilities are not construct-relevant (e.g., test anxiety, test taking skills, language barrier) but are never-the-less factors that may impact an examinee’s performance on the examination.

