



Standard Setting Report: Sonography - Effective August 2026

Background

The mission of the American Registry of Radiologic Technologists (ARRT) is to “to promote safe, high-quality patient care through credentialing, collaboration, and advocacy.” Within ARRT’s examination program, standard setting is a formal, systematic process that links examination performance to a defensible determination of minimum competence for certification. The previous standard setting workshop for the Sonography (SON) examination was conducted in June 2021.

This report details a standard setting workshop conducted in May 15-16, 2025 for the SON examination. It describes (a) committee composition, (b) the definition of the minimally qualified candidate (MQC), (c) the standard setting methods and training provided, (d) the results from each method, and (e) the ARRT Board of Trustees’ decision and implementation plan.

The SON examination is designed to assess entry-level competence. The passing standard is intended to support public protection by distinguishing candidates who have achieved the minimum required knowledge and skill from those who have not.

Panelists completed required confidentiality and conflict-of-interest attestations consistent with credentialing best practices for accreditation review.

Facilitator:

- Zachary Siegel, Ph.D., Manager of Psychometrics at ARRT

Committee Composition

ARRT selected subject matter experts (SMEs) from the volunteer database to maximize diversity in geography, work setting, and professional experience in Sonography. When feasible, ARRT biases selection toward practitioners earlier in their career because the examination is designed to assess entry-level competence. ARRT also invites the radiologist assigned by the American College of Radiology to the relevant examination committee.

For this workshop, nine SMEs were invited. Of those, six SMEs participated as panelists. The panel included five female and one male. Four panelists reported hospital-based practice, one reported clinic-based practice, and one reported private practice. Three panelists were experienced sonographers and three were entry-level practitioners. Table 1 below summarizes demographic details.

Panel composition was reviewed to ensure appropriate representation across key practice characteristics, including experience level, work setting, and role.



Table 1. Committee Demographics

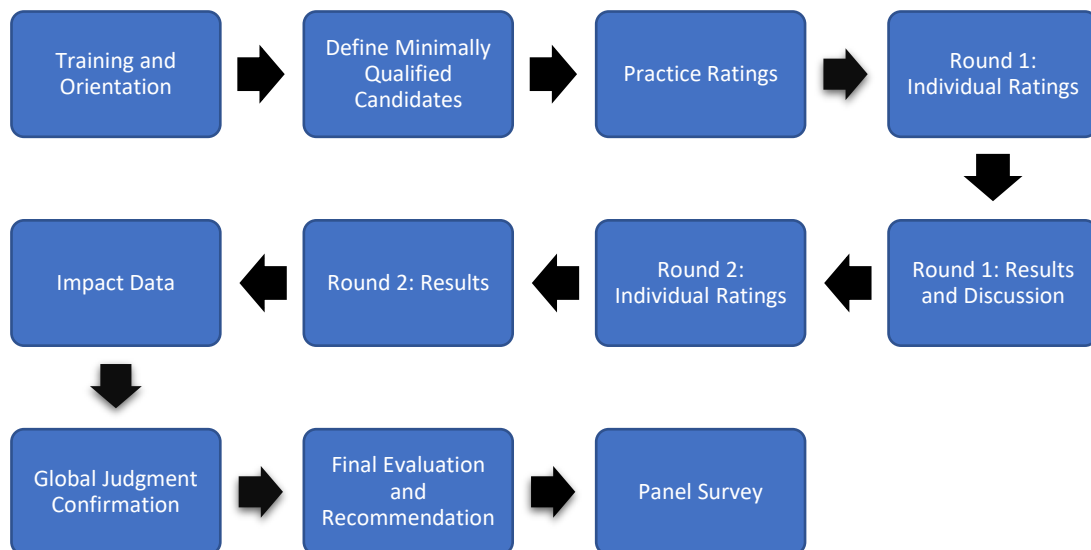
Member	Location	GENDER	INSTITUTION	EXPERIENCE	ROLE	CREDENTIALS
A	AL	F	Hospital	Experienced	Sonographer	R.T.(S)(ARRT)
B	OR	F	Hospital	Entry Level	Sonographer	R.T.(S)(ARRT)
C	VA	F	Private Office	Entry Level	Sonographer	R.T.(S)(ARRT)
D	TX	F	Hospital	Entry Level	Sonographer	R.T.(S)(ARRT)
E	FL	F	Clinic	Experienced	Sonographer	R.T.(R)(S)(VS)(ARRT)
F	TX	M	Hospital	Experienced	Sonographer	R.T.(R)(S)(VS)(ARRT)

Standard Setting Process

Prior to the workshop, panelists were sent the examination content specifications, which describes the domain areas assessed on the SON examination. Panelists were also provided with task inventory, a list of activities performed by practicing sonographers, and were asked to review both materials in advance.

Figure 1 summarizes the sequence of activities used during the workshop. The process began with an overview of standard setting and training on the modified Angoff method (Angoff, 1971). Panelists then used the content specifications and task inventory to define the minimally qualified candidate, completed practice ratings, and conducted two rounds of Angoff ratings with structured discussion and feedback. Panelists were instructed to base their judgments on the minimally qualified candidates, rather than on clearly passing (highly experienced) candidates or failing candidates. Finally, panelists completed global judgment activities (Beuk, 1984; and Hofstee, 1983) to evaluate the reasonableness of the Angoff-based results. The committee then made a final committee recommendation based on the Round 2 Modified Angoff results.

Figure 1. Procedures Followed during the Standard Setting Workshop



Minimally Qualified Candidate

Following training on the purpose of standard setting, its implications, and common sources of judgment biases, panelists developed a shared description of the minimally qualified candidate (MQC). The MQC is defined as a candidate who possesses the minimum knowledge, skills, and abilities required for safe and competent entry level practice in Sonography. This candidate represents the lower boundary of acceptable performance, demonstrating just sufficient competence to practice safely and effectively.

Panelists reviewed the SON content specifications and task inventory and discussed what MQCs can perform reliably and independently, including the minimum level of clinical reasoning and procedural knowledge expected under typical testing conditions. Although the term “entry level” refers to the stage of professional practice, the MQC represents a performance threshold rather than a typical practitioner. Standard setting judgments are therefore based on the MQC as the point at which a candidate’s performance is just sufficient to meet the requirements for certification.

Methods

The study used a modified Angoff procedure (Angoff, 1971) as the primary item-centered standard setting method on a SON examination form. Following completion of the modified Angoff ratings, panelists completed Hofstee and Beuk global judgment confirmation exercises (Beuk, 1984; Hofstee, 1983) to contextualize and evaluate the reasonableness of the Angoff-derived standard recommendation.

Standard Setting Examination

The operational Sonography examination consists of 360 scored items. To reduce panel burden while maintaining representative converge of the examination content, a 180-item standard setting form was assembled. The form reflected the operational content specifications, cognitive demands, and psychometric characteristics of the operational examination and served as the basis for all panel judgments during the modified Angoff exercise.

Modified Angoff Procedure

Prior to operational ratings, panelists completed a structured practice activity using a set of representative items. This exercise was designed to align their understanding of the MQC definition and the application of the rating scale, and to promote consistency in subsequent judgments. Panelists were instructed to make independent judgements, focusing on the MQC rather than on typical candidates or their own expectations.

In Round 1, panelists independently evaluated each of the 180 items and estimated the percentage of MQCs who would answer the item correctly. No performance statistics were provided during Round 1.

Before Round 2, facilitators provided structured feedback to support informed reconsideration of judgments, including summary statistics of Round 1 ratings and identification of items with higher rating variability. Where available, first-time candidate proportion correct (or p-value) information was provided as interpretive context, with training that emphasized that operational data may inform (but should not replace) MQC-based judgment.



In Round 2, panelists reviewed their initial ratings and completed a second round of independent judgments. Final Angoff results were computed by summing item-level probabilities to produce a recommended raw score cut and associated percent correct cut on the standard setting examination form.

Global Judgment Confirmation (Beuk and Hofstee) Procedure

Following completion of the modified Angoff procedure, panelists provided independent global judgments using the Beuk (1984) and Hofstee (1983) methods. These procedures complemented the item-level Angoff judgments by encouraging holistic evaluation of the examination form and expected candidate outcomes and are commonly used to contextualize Angoff results.

For the Beuk method, panelists responded to two questions designed to capture global expectations regarding expected percent correct and expected pass rate for the minimally qualified candidates. Panelists indicated (a) the percentage of items an MQC should answer correctly to pass the examination and (b) the percentage of candidates who should be expected to pass.

For the Hofstee method, panelists responded to four questions defining acceptable ranges for both cut scores and pass rates. These questions asked panelists to specify the lowest and highest acceptable cut scores and the lowest and highest acceptable pass rates that would remain defensible in terms of public protection and fairness to candidates.

These judgments were compared to the Angoff results to evaluate consistency and support the reasonableness of the recommended passing standard. This step was confirmatory and did not serve to adjust the cut score.

Results

Results from the standard setting study are presented separately for the modified Angoff procedure and the Hofstee and Beuk global judgment methods. The modified Angoff procedure served as the primary basis for the recommended passing standard, while the global judgment methods provided additional evidence regarding the reasonableness and practical implications of the recommendation. Tables 2 through 5 summarize the outcomes of each method and supporting evaluation information.

Table 2 presents the distribution of panelist Angoff judgments across Rounds 1 and 2, including measures of central tendency, variability, inter-rater reliability, and standard error of judgment. Specifically, Table 2 shows that the Round 1 mean Angoff recommendation corresponded to a raw score of 126.6 (70.4%), with substantial variability (13.8) across panelists. After feedback and discussion, the Round 2 judgments converged to a higher mean raw score of 128.4 (71.3%), with reduced variability (4.5) and increased inter-rater reliability (from 0.57 to 0.94), indicating greater consistency among panelist's judgments.



Table 2. Modified Angoff Results

First Round			Second Round		
Statistic	Raw Scale	% Scale	Statistic	Raw Scale	% Scale
Average:	126.6	70.4	Average:	128.4	71.3
SD:	13.8	7.7	SD:	4.5	2.5
Min:	106.1	59.0	Min:	123.6	68.7
Max:	143.0	79.4	Max:	134.2	74.5
Sample SE:	5.6	3.1	Sample SE:	1.9	1.0
IRR-based SE:	9.1	5.0	IRR-based SE:	1.1	0.6
Max possible:	180	100	Max possible:	180	100
Inter-rater Reliability:	0.57		Inter-rater Reliability:	0.94	
Number of raters:	6		Number of raters:	6	

Note: SEJ denotes the standard error of judgment and reflects uncertainty (i.e., variability) in the panel's judgment-based standard. Inter-rater reliability (IRR) estimates describe agreement among panelists.

Results from the Beuk method are summarized in Table 3, showing that panelists' Beuk judgments yielded a mean cut score of 75.5% correct, with an associated mean (expected) pass rate of 81.7%. The standard deviation reported in Table 3 reflects variability in panelists' holistic judgments regarding both the cut score and the resulting pass rate.

Table 3. Beuk Results

Mean cut score	75.5%
SD cut score	3.4%
Mean pass rate	81.7%
SD pass rate	5.6%

Table 4 displays the minimum and maximum acceptable cut scores (72.2% and 81.7%) and pass rates (63.3% and 80.0%) specified by panelists using the Hofstee method. These bounds define a region of acceptable outcomes within which panelists judge both the passing standard and the resulting pass rates to be reasonable for certification decisions.

Table 4. Hofstee Results

Minimum cut score	72.2%
Maximum cut score	81.7%
Lowest pass rate	63.3%
Highest pass rate	80.0%

To evaluate the potential impact of judgment uncertainty on candidate outcomes, projected passing rates were examined across a range of cut scores defined by ± 1 and ± 2 standard errors of judgment (SEJ). Table 5 presents projected passing rates associated with the Round 2 Angoff cut score, as well as Beuk- and Hofstee-adjusted scenarios, relative to the current operational passing standard.

As shown in Table 5, projected passing rates vary across methods and SEJ-adjusted cut scores, illustrating the sensitivity of candidate outcomes to modest changes in the passing



standard. The table also provides a reference comparison to the current operational cut score, supporting evaluation of continuity and potential impact.

Table 5. Panelist Cut Scores and Projected Passing Rate

Projected Cut and Passing Rate					
	Cut -2SEJ	Cut - 1SEJ	Cut	Cut + 1SEJ	Cut + 2SEJ
Angoff Cut	125	126	127	128	129
Pass rate	53%	52%	50%	47%	46%
Beuk Judgment	118	118.9	120	121	122
Pass rate	63%	62%	60%	59%	57%
Hofstee Judgment	121	122	123	124	125
Pass rate	59%	57%	56%	54%	53%
Recommendation			124		
Pass rate			54%		
Current Cut			122		
Pass rate			57%		

Note: Beuk and Hofstee adjustment values are presented for comparison purposes only and do not represent an alternative passing standard.

The Angoff recommendation together with supporting evidence from the Beuk and Hofstee exercises, projected candidate impact, and historical program data, was presented to the ARRT Board of Trustees for consideration.

Board of Trustees' Decision 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 Sonography examination.

After review and discussion, the Board approved a new passing standard for the Sonography examination. The approved operational passing standard is 248 on the 360-item operational examination, which corresponds to 124 on the 180-item standard setting form used during the study. This translation was accomplished using ARRT's established equating and score-scaling procedures, which use common anchor items to ensure that the approved passing standard represents the same level of candidate competence across operational examination forms.

The new standard will go into effect August 2026 and remain in place until the next standard setting is scheduled to take place. ARRT staff expect a future pass rate for first-time candidates around 54% based on the impact data provided to both the Board and standard setting committee.

References

Angoff, W. H. (1971). Scales, norms, and equivalent scores. In R. L. Thorndike (Ed.), *Educational measurement* (2nd ed, pp. 508-600). Washington, DC: 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.

