

Scoring System and Procedure

This scoring system was designed to encourage more effective spreading of scores. Highly rating all applications greatly diminishes the ability of a reviewer or study section to communicate the scientific impact of an application. Therefore, reviewers who carefully consider the rating guidance provided in determining their scores improve not only the reliability of their scores, but also improve their ability to communicate the scientific impact of the applications reviewed.

Contents

[Scoring](#)

- [Summary](#)
- [Preliminary Scores](#)
- [Criterion Scoring](#)
- [Impact/Priority Score](#)
- [Non-Numeric Scores](#)
- [Reviewer Guidance and Chart](#)

[Percentiling](#)

SCORING

Summary

- The NIH grant application scoring system uses a 9-point scale
- A score of 1 indicates an exceptionally strong application with essentially no weaknesses. A score of 9 indicates an application with serious and substantive weaknesses with very few strengths; 5 is considered an average score
- Ratings are in whole numbers only (no decimal ratings)
- This scale is used by all reviewers to provide an overall impact/priority score and for assigned reviewers to score five individual criteria (e.g., Significance, Investigator(s), Innovation, Approach, Environment)
- For the impact/priority score rating, strengths and weaknesses across all of the review criteria should be considered. For each criterion rating, the strengths and weaknesses within that review criterion should be considered.
- Reviewers should consider not only the relative number of strengths and weaknesses noted, but also the importance of these strengths and weaknesses to the criteria or to the overall impact when determining a score. For example, a major strength may outweigh many minor and correctable weaknesses.
- For information about using the critique template, see Critique Template Instructions
- For information about using the critique template, see [Critique Template Instructions](#).
- NIH expects that scores of 1 or 9 would be used less frequently than the other scores.

Preliminary Scores

- Before the review meeting, assigned reviewers will determine preliminary scores for each of the five scored review criteria and a preliminary score for the overall impact/priority
- The impact/priority score should reflect the reviewer's overall evaluation, not a numerical average of individual criterion scores.

- Reviewers should consider the full range of the rating scale and the scoring descriptors in assigning preliminary and final scores. However, a reviewer should not assume that the applications assigned to him/her necessarily cover that entire range of scores, and should assign scores as appropriate for the work or science proposed.
- An application does not need to be strong in all categories to be judged likely to have major impact. For example, a project that by its nature is not innovative may be essential to advance a field.
- Reviewers must enter the criterion scores into the Internet Assisted Review site in the NIH Commons for them to appear in the summary statement. If entered in IAR, the scores will be transferred to a table at the beginning of the reviewer's critique.
- Assigned reviewers may submit criterion scores only after their critiques have been uploaded. At the SRO's discretion, discussants who are assigned to the application and SRG members who are not assigned to the application may submit criterion scores without critiques.
- In the READ phase of the meeting reviewers may submit their scores and critiques, but may not edit them.
- These scores are not retained, but will be replaced by final scores that are given by private scoring and are based on the outcome of the deliberations at the peer review meeting

Criterion Scoring

- Up to five, individual criteria are scored
- Criterion scores are provided for both discussed and not discussed applications
- Criterion scores are intended to provide additional information on how each assigned reviewer weighed that particular section so that the reader has a better idea of strengths and aspects that need improvement
- It is not intended that an Impact/Priority score for the application be an average of criterion scores
- Criterion scores are entered into the Internet Assisted Review site for the meeting; the same screen also allows uploading of the written critique at the same time
- If the reviewer's opinion changed as a result of discussion at the meeting, the reviewer should change his/her criterion scores to match his/her critiques and overall impact/priority score The criterion scores appear in a table at the beginning of each critique in the summary statement

Impact/Priority Score

- Discussed applications will receive impact/priority scores from all eligible reviewers (e.g., without conflicts of interest)
- The Impact/Priority score for an application is based on each individual reviewer's assessment based on the five scored criteria plus additional criteria regarding the protection and inclusion of human subjects; vertebrate animal care and welfare; biohazards, and RFA-specific criteria, as applicable to the application
- Reviewers are guided to use the full range of the rating scale and spread their scores to better discriminate among applications
- Reviewers whose evaluations or opinions of an application fall outside the range of those presented by the assigned reviewers and discussant(s) should ensure that their opinions are brought to the attention of the entire committee. In addition, the SRO and Chairperson should ensure that all opinions are voiced before final scoring is conducted.

- Reviewers should feel free to assign the score that they believe best represents the impact of the application, and not feel constrained to limit their scores to the upper half of the score range if they do not feel such a score is warranted.
- After the meeting, individual reviewer scores will be averaged and the result multiplied by 10 to determine the final impact/priority score
- The range of the final application scores is from 10 to 90

Non-Numeric Scores

- Not Discussed (ND)
 - Applications unanimously judged by the peer review committee to be less competitive are not discussed at the peer review meeting
 - These applications do not receive a numerical Impact/Priority score
 - These applications do receive individual criterion scores
 - There is no set number of applications that are not discussed; up to 60% may not be discussed depending on the number received and programmatic funding expectations; in some meetings, the “not discussed” option may not be used at all
- Not Recommended for Further Consideration (NRFC)
 - Score for an application NRFC is by majority vote of the peer reviewers
 - Based on lack of significant and substantial merit; OR it presents serious ethical problems in the protection of human subjects from research risks; OR it presents serious ethical problems in the use of vertebrate animals, biohazards, and/or select agents.
 - NRFC-scored applications do not proceed to the second level of peer review (National Advisory Council/Board) because they cannot be funded.
 - Typically, NRFC scores are rare
- Other Non-numeric Scores
 - Deferred (usually due to lack of sufficient information, quorum, allegations of research misconduct)
 - Abstention (used rarely)
 - Conflict (score put in by a reviewer who is in conflict with the application)
 - Not Present

Reviewer Guidance and Chart

- For the impact/priority score and for the individual criterion scores, the far right column provides a descriptive guide of how strengths and weaknesses are considered in assigning a rating.
 - Weakness descriptions: **Minor weakness:** easily addressable weakness, does not substantially lessen impact; **Moderate weakness:** lessens impact; **Major weakness:** Severely limits impact
- Impact (far left column) is the project’s likelihood to have a sustained, powerful influence on the research field(s) involved: 1 – 3 = High Impact; 4 – 6 = Moderate Impact; 7 – 9 = Low Impact
- Each review criterion should be assessed based on how important each review criterion is to the work being proposed. As a result, a reviewer may give only moderate scores to some of the review criteria but still give a high overall impact/priority score because the one review criterion critically important to the research is rated highly; or a reviewer could give mostly high criterion ratings but rate

the overall impact/priority score lower because the one criterion critically important to the research being proposed is not highly rated.

- An application does not need to be strong in all categories to be judged likely to have major impact, e.g., a project that by its nature is not innovative may be essential to advance a field.

Impact	Score	Descriptor	Additional Guidance on Strengths/Weaknesses
High	1	Exceptional	Exceptionally strong with essentially no weaknesses
	2	Outstanding	Extremely strong with negligible weaknesses
	3	Excellent	Very strong with only some minor weaknesses
Medium	4	Very Good	Strong but with numerous minor weaknesses
	5	Good	Strong but with at least one moderate weakness
	6	Satisfactory	Some strengths but also some moderate weaknesses
Low	7	Fair	Some strengths but with at least one major weakness
	8	Marginal	A few strengths and a few major weaknesses
	9	Poor	Very few strengths and numerous major weaknesses

Additional Information for Scoring Guidance Table
<p>Non-numeric score options: NR = Not Recommended for Further Consideration, DF = Deferred, AB = Abstention, CF = Conflict, NP = Not Present, ND = Not Discussed</p>
<p>Minor Weakness: An easily addressable weakness that does not substantially lessen impact</p>
<p>Moderate Weakness: A weakness that lessens impact</p>
<p>Major Weakness: A weakness that severely limits impact</p>

PERCENTILING

- For the appropriate applications (certain activity codes or RFAs), scores will be percentiled to the appropriate base (e.g. study section base if the number of R01 applications ≥ 25 ; CSR-all or IC-all base if <25)
- All percentiles are rounded to a whole number
- Until a base has been established from three rounds of review (i.e., May 2010 Council), percentiles are based on less than three application rounds