

Scoring System and Procedure

EXECUTIVE SUMMARY

- The NIH grant application scoring system uses a 9-point rating for the impact/priority score with 1 = Exceptional and 9 = Poor.
- Ratings are in whole numbers only (no decimal ratings).
- Assigned reviewers also provide ratings for each review criterion [e.g. Significance, Investigator(s), Innovation, Approach, Environment] using the same 9-point scale.
 - These criterion ratings are provided in the summary statement for applications, both discussed and not discussed.
 - Criterion ratings should be considered in determining the overall impact/priority score, but reviewers should determine the relative importance of each criterion for the science or work being proposed.
- Reviewers should use the full range of the rating scale and spread their scores to better discriminate among applications.
- Discussed applications will receive impact/priority scores from all eligible reviewers (e.g., without conflicts of interest). Individual reviewer scores will be averaged and the result multiplied by 10 to determine the final impact/priority score (range of 10 to 90).
- 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) and reported in whole number percentiles. Until a base has been established from three rounds of review (May 2010 Council), percentiles are based on less than 3 application rounds.
- For information about using the critique template, see [Critique Template Instructions](#).

PROCEDURE FOR PRELIMINARY SCORES

In scoring each of the core criteria and impact/priority, reviewers will use a scale of whole numbers, ranging from 1 to 9 (1= exceptional; 9= poor). The SRO will provide additional guidance on the use of this scoring scale.

Before the review meeting, determine a separate score for each of the core review criteria and a score for the impact/priority. The impact/priority score should reflect your overall evaluation rather than a weighted average applied to scores given to each criterion. 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.

- The criterion scores for the applications should be entered in the meeting Internet Assisted Review (IAR) site in NIH Commons before the review meeting using the same page that is used for submitting the preliminary impact/priority score and critique. Core criterion scores can be submitted only after your critique had been uploaded into IAR.

- You must enter the criterion scores into IAR for them to appear in the summary statement. If entered in IAR, the scores will be transferred to a table at the beginning of your 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.
- The criterion scores should be changed during FINAL SCORING on your electronic or paper Voter/Scoring Sheet, or following the review meeting during the EDIT phase if your opinion changed as a result of discussion.
- Each core review criterion should be given a score using the nine-point rating scale in accordance with the Enhanced Peer Review Criteria.

The NIH Grant Application Scoring System

The NIH scoring system uses a 9-point rating scale from 1 = Exceptional to 9 = Poor for the overall impact/priority score as well as the individual review criteria. Ratings are provided only in whole numbers, not decimals.

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
<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>			

- For the impact/priority score, the far left column provides guidance for assigning scores to applications based on the project's likelihood to have a sustained, powerful influence on the research field(s) involved:
 - 1 to 3 = high impact
 - 4 to 6 = moderate impact
 - 7 to 9 = low impact
- 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. A score of 1 indicates an exceptionally strong application (or exceptionally strong significance, investigators, innovation, approach, environment) with essentially no weaknesses. A score of 9 indicates serious and substantive weaknesses with very few strengths. 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.

Not Discussed and Not Recommended for Further Consideration

Applications judged unanimously by the peer reviewers as less competitive, based on preliminary impact/priority scores (roughly the bottom half of applications for that review meeting), will not be discussed and will not receive a final impact/priority score. Although the summary statement for such an application will indicate "ND" (not discussed), it will contain critiques and criteria scores from each of the assigned reviewers.

An application may be designated Not Recommended for Further Consideration (NRFC) by the Scientific Review Group if it: lacks significant and substantial merit; presents serious ethical problems in the protection of human subjects from research risks; or presents serious ethical problems in the use of vertebrate animals, biohazards, and/or select agents. Applications designated as NRFC do not proceed to the second level of peer review (National Advisory Council/Board) because they cannot be funded.

Following discussion, however, 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 score to the upper half of the score range if they do not feel such a score is justified. For example, if the assigned reviewers initially score an application as 4, 5, and 6, and subsequent discussion reveals a serious weakness that will substantially lessen the project's impact, then it is appropriate for reviewers to give a higher (worse) score.

Distribution of Scores

With 9 possible rating discriminations, it is imperative that reviewers distribute or spread their scores as widely as possible among applications. The descriptors associated with each rating were designed to encourage the spreading of scores. Therefore, although score distributions may vary by study section, reviewers should use the full range of 1 to 9; the expectation, however, is that there will be few 1s and few 9s.

This scoring system was designed to encourage greater spreading of scores. Highly rating all applications greatly diminishes the ability of a reviewer or study section to communicate the

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 impact of the applications reviewed.

Scoring Range

After discussion, the assigned reviewers state their final scores, defining the score range. Based on the discussion, all eligible reviewers also score the application. If reviewers wish to score outside the score range of the assigned reviewers, they should declare that they intend to score outside the range and briefly describe the reason. Any score outside the range of the assigned reviewers should be declared, even if the range is a single score (i.e. all assigned reviewers give the same final score). It is important that all points of view and opinions of reviewers are discussed; therefore, reviewers should feel free to score outside the range based on their determination of the overall impact of the application.

All scientific opinions concerning an application that is discussed at the SRG meeting should be raised during that discussion. Therefore, SRG members 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.

Additional Guidance on Criterion Scoring

Assigned reviewers provide both preliminary impact/priority scores and criterion scores (ratings of each review criteria). These criterion scores are included in the summary statement to give applicants of both discussed and not discussed (i.e. streamlined) applications a sense of how consideration of the review criteria influenced the overall evaluation of the application. However, because the relative importance of each individual criterion to the overall score differs for each application, reviewers should not use a formula of weighted or unweighted averages across applications to determine the overall impact/priority score. In addition, unrated criteria such as human subjects, vertebrate animal care, and RFA-specific criteria also should be considered in determining the overall impact/priority score. Therefore, each review criterion should be weighed differently for each application depending 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.

Final Impact/Priority Scores and Percentile Scores

Discussed applications will receive impact/priority scores from all eligible reviewers. Individual reviewer scores will be averaged and the result multiplied by 10 to determine the final impact/priority score (range of 10 to 90) reported in the summary statement.

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) and reported in whole number percentiles. Until a base has been established from three rounds of review (May 2010 Council), percentiles are based on less than 3 application rounds.