

# Scoring System and Procedure

The [NIH scoring system](#) was designed to encourage reliable scoring of applications. Reviewers or study sections who assign high ratings to all applications diminish their ability to communicate the scientific impact of an individual application. Therefore, reviewers who carefully consider the rating guidance below can improve the reliability of their scores as well as 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](#)

## SCORING

### Summary

- The NIH grant application scoring system uses a 9-point scale for both overall impact scores and scores for individual review criteria.
  - For both types of score, ratings are in whole numbers only (no decimal ratings).
  - NIH expects that scores of 1 or 9 to be used less frequently than the other scores.
- For the overall impact score,
  - the scale is used by all eligible (without conflict of interest) SRG (Scientific Review Group) members
  - 5 is considered an average score.
- For criterion scores,
  - the scale is used by the assigned reviewers to evaluate (at least) five individual criteria (e.g., Significance, Investigator(s), Innovation, Approach, Environment).
  - reviewers should consider the strengths and weaknesses within each criterion. For example, a major strength may outweigh many minor and correctable weaknesses.
- For information about using the critique template, see [Critique Template Instructions](#)

### Preliminary Scores

- Before the review meeting, assigned reviewers determine preliminary scores for each of the scored review criteria and a preliminary score for the overall impact
- The impact 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 (IAR) 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, SRG members assigned as discussants 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
- Final scores are given by private scoring and are based on the outcome of the deliberations at the peer review meeting

## **Criterion Scoring**

- In most cases, five individual criteria are scored, but certain Funding Opportunity Announcements may include more than five scored criteria
- Criterion scores are provided for all applications
- Criterion scores are intended to convey how each assigned reviewer weighed the strengths and weaknesses of each section
- Providing scores without providing comments in the review critique is discouraged
- The impact score for the application is not intended to 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 score as part of the EDIT phase
- The criterion scores appear in a table at the beginning of each critique in the summary statement

## **Impact Score**

- Discussed applications receive numerical impact scores from all eligible reviewers (e.g., without conflicts of interest)
- The impact score for an application is based on each individual reviewer's assessment of the scored criteria plus additional criteria regarding the protection and inclusion of human subjects; vertebrate animal care and welfare; biohazards, and criteria specific to the funding opportunity
- 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

- Reviewers will score an application as presented in its entirety, and may not modify their scores on the assumption that a portion of the work proposed will be deleted or modified according to the SRG's recommendations
- After the meeting, individual reviewer scores will be averaged and the result multiplied by 10 to determine the final impact score
- The range of the final application scores is 10 through 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 score
  - These applications do receive individual criterion scores
  - Not all meetings use the "Not Discussed" option
- Not Recommended for Further Consideration (NRFC)
  - NR for an application occurs by majority vote of the SRG members
  - NR occurs in the following scenarios:
    - Application lacks significant and substantial merit
    - Application presents serious ethical problems in the protection of human subjects from research risks
    - Application presents serious ethical problems in the use of vertebrate animals, biohazards, and/or select agents
  - NR-scored applications do not proceed to the second level of peer review (National Advisory Council/Board) because they cannot be funded
  - The NR is a serious committee recommendation that is substantially different from Not Discussed (ND)
- Other Non-numeric Scores
  - DF: Deferred (usually due to lack of sufficient information or quorum, allegations of research misconduct)
  - AB: Abstention (used rarely)
  - CF: Conflict (score put in by a reviewer who is in conflict with the application)
  - NP: Not Present

## Reviewer Guidance

- The table below provides a guide for reviewers in assigning overall impact scores and individual criterion scores.
- Overall impact, for a research project, is the project's likelihood to have a sustained, powerful influence on the research field(s) involved, but may be defined differently for different types of applications.
- Each review criterion should be assessed based on the strength of that criterion in the context of 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 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 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.
- A score of 5 is a good, medium-impact application.
- The entire scale (1-9) should always be considered.

Overall Impact or Criterion Strength	Score	Descriptor
<b>High</b>	<b>1</b>	<b>Exceptional</b>
	<b>2</b>	<b>Outstanding</b>
	<b>3</b>	<b>Excellent</b>
<b>Medium</b>	<b>4</b>	<b>Very Good</b>
	<b>5</b>	<b>Good</b>
	<b>6</b>	<b>Satisfactory</b>
<b>Low</b>	<b>7</b>	<b>Fair</b>
	<b>8</b>	<b>Marginal</b>
	<b>9</b>	<b>Poor</b>
<b>Other Designations for Final Outcome</b>		
AB	Abstention	
CF	Conflict of Interest	
DF	Deferred	
ND	Not Discussed	
NP	Not Present	
NR	Not Recommended for Further Consideration	

See specific guidance for [Research Applications](#) and [Training Applications](#).