NIH GRANT APPLICATION SCORING SYSTEM

All NIH scientific review groups are asked to use the nine-point numerical scale shown below for assigning Criterion and Overall Impact/Priority scores. A score of 1 indicates an exceptionally strong application (or exceptionally strong Significance, Investigators, Innovation, Approach, and Environment) with essentially no weaknesses. A score of 9 indicates serious and substantive weaknesses with very few strengths. Reviewers provide scores in whole numbers only, not decimals.

CRITERION SCORING

Criterion scores rate each of the five review criteria—Significance, Investigators, Innovation, Approach, and Environment—using the 1-9 scoring scale. For each criterion rating, the strengths and weaknesses within that review criterion should be considered. In considering strengths and weaknesses, reviewers should consider the relative importance of the strengths and weaknesses noted, not simply the number of strengths and weaknesses. These criterion scores are included in the summary statement to give applicants of both discussed and non-discussed (i.e. streamlined) applications a sense of how consideration of the review criteria influenced the overall evaluation of the application. The individual criteria scores are not mathematically related to the Overall Impact/Priority score. This allows for an assessment of how the application as a whole comes together. Additionally, 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 un-weighted averages of the criterion scores across applications to determine the Overall Impact/Priority score.

OVERALL IMPACT/PRIORITY SCORE

When determining the Overall Impact/Priority score, the far left column in the table 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. In addition, unrated/un-scored 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 that is 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 that is critically important to the proposed research is not highly rated.
With nine possible scoring discriminations, it is imperative that reviewers distribute or spread their scores as widely as possible among applications. The descriptors (above) associated with each rating were designed to encourage the spreading of scores. Therefore, although score distributions may vary by review meeting, reviewers should use the full range of 1 to 9; the expectation, however, is that there will be few 1s and few 9s.

High scores for all or most of the assigned applications greatly diminishes the ability of a reviewer or review panel 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.

Some review meetings, due to the large number of applications, discuss only a percentage (usually 50 percent) of applications that were submitted. Typically, these discussed applications have preliminary scores in the better half of the scoring range. Following discussion of an application, 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 2, 3 and 3, 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.

After discussion of an application, the assigned reviewers state their final Overall Impact/Priority scores, defining the score range. Based on the discussion, all non-conflicted reviewers also score the application. Unassigned reviewers may score outside this range; however, they must announce their intent to do so during scoring and briefly describe the reason to ensure that the
reasons have been discussed adequately. Any score outside the range suggested by 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.