



**AMERICAN
PSYCHOLOGICAL
ASSOCIATION**

October 8, 2026

Raymond Jacobson, PhD
Director
Center for Scientific Review
National Institutes of Health
6701 Rockledge Drive MSC 7768
Bethesda, MD 20817

**Re: NIH Request for Information: Proposed Changes to Reporting Outcomes from NIH Peer Review
(NOT-OD-26-088)**

Submitted electronically via email: peerreview@nih.gov and through the NIH response portal

Dear Dr. Jacobson:

The American Psychological Association (APA) appreciates the opportunity to respond to the National Institutes of Health's (NIH) Request for Information on Proposed Changes to Reporting Outcomes from NIH Peer Review ([NOT-OD-26-088](#)). APA is the nation's largest scientific and professional nonprofit organization representing psychology, with 190,000 members and affiliates, including clinicians, researchers, educators, consultants, and students in psychological science.

Under the proposal, NIH would retain its numerical scoring process but report application outcomes only as "Most Competitive," "Competitive," or "Not Discussed." Final overall impact scores and percentiles would no longer be shared with applicants, institutions, NIH program staff, Institute and Center leadership, or Advisory Councils. APA supports NIH's goal of ensuring that funding decisions reflect scientific merit, mission priorities, strategic plans, and responsible stewardship rather than an overly mechanical reliance on paylines or slight differences in numerical scores. Such scores can imply a level of precision that peer review cannot support. However, replacing them with broad categories would eliminate valuable information without addressing variability or potential bias in review. **APA therefore urges NIH not to adopt the proposal in its current form.**

Preserve Transparency and the Value of Peer-Review Feedback

Peer review is both an evaluative process and a means of strengthening science. Written critiques, criterion scores, panel discussion summaries, and overall impact scores provide complementary information that helps applicants understand a panel's judgment, improve future submissions, and assess the fairness of review procedures. Although overall impact scores are imperfect and should not be treated as precise measures of scientific value, removing them would not address the underlying sources of variability or bias in the review process (Recio-Saucedo et al., 2022). Instead, broad categorical labels would discard useful information, weaken accountability, and diminish confidence in NIH funding decisions.

As Mietchen (2014) explains, transparency in research funding serves purposes far beyond oversight, helping to foster learning, collaboration, and public engagement while strengthening confidence in the stewardship of public resources. Removing an existing source of information would represent a step backward. This transparency is particularly important when funding decisions diverge from peer-review rankings because it allows stakeholders to

APA.ORG
APASERVICES.ORG

Advocating for APA members and psychology

750 First Street, NE
Washington, DC 20002-4242
202.336.5800
202.336.6123 TDD

understand how and why such decisions were made. When funding decisions depart from peer-review outcomes without a clear, auditable explanation, confidence in the primacy of scientific merit is weakened. NIH can reduce overreliance on paylines without reducing transparency or weakening accountability.

Information Needed for Applicant Decisions and Accountability

Investigators use scores and percentiles to gauge an application's competitiveness, assess whether reviewers' concerns can be addressed, decide whether to resubmit or redirect a project, and plan staffing, budgets, and study preparation. Scores also indicate whether the concerns raised by assigned reviewers were shared by the full panel. Because narrative critiques are not calibrated to a common scale, eliminating the panel-level score would make it harder to distinguish an application that could become competitive through targeted revisions from one that requires substantial redesign.

Page | 2

Converting scores into categories that each span 25 percentage points would obscure meaningful differences among applications and make the relationship between peer-review judgments and funding decisions more difficult to assess. Scores and percentiles help applicants, Advisory Councils, and the scientific community understand whether an outcome reflects scientific merit, programmatic priorities, portfolio balance, or budget constraints. Withholding this information would make departures from peer-review standing harder to identify and weaken the auditable record needed for public accountability. Even when funding decisions are well founded, limited visibility into their basis could create the appearance that criteria are insufficiently documented or inconsistently applied. APA recommends that NIH continue providing final overall impact scores and percentile rankings, where applicable, to program staff, Institute and Center leadership, and Advisory Councils. NIH may supplement these measures with categorical designations, but categories should not replace the more detailed information currently available. NIH should also clearly explain that numerical metrics inform, but do not determine, award decisions.

Design Risks of Broad Categories

The proposed categories would create new cliff effects at the 25th- and 50th-percentile boundaries. Applications on opposite sides of a boundary could receive different labels despite nearly indistinguishable panel judgments, while applications with substantially different evaluations could receive the same label. The proposal could also obscure variation across review panels, funding mechanisms, Institutes and Centers, and scientific fields. This concern is particularly significant for interdisciplinary, behavioral, prevention, implementation, and population research, for which review context and portfolio fit may vary across study sections.

Efficiency and Administrative Burden

The proposal could also create unintended administrative burdens. Applicants who cannot determine where they fall within a category may resubmit applications that are already likely to be funded or continue revising proposals with little realistic prospect of success. These avoidable submissions could increase workloads for investigators, reviewers, scientific review officers, and program staff while generating additional requests for individualized guidance from program officers.

The "Competitive" category would encompass applications in the 26th through 50th percentile range, masking meaningful distinctions in panel assessments. Because reviewers would still enter numerical scores and NIH could still calculate final overall impact scores, the proposal would not reduce scoring responsibilities; it would simply make the resulting information less accessible and useful.

If NIH adopts broader categories, it should publicly document how applications are selected for funding within each category, how programmatic considerations are applied, and how often award decisions differ from the relative standing established through peer review. Aggregated reporting would support public accountability and help NIH identify uneven effects across applicants and institutions.

Equity Implications for Early-Career Investigators and Behavioral and Psychological Science

Reduced access to review information would not affect all applicants equally. Established investigators and well-resourced institutions may be better positioned to interpret limited feedback because they have experienced research administrators, robust mentoring networks, and established relationships with NIH staff. Early-career investigators and researchers at under-resourced institutions are less likely to have comparable support and may rely more heavily on scores and percentiles to learn from the review process and make informed strategic decisions.

Page | 3

Broad categories may also disadvantage behavioral and psychological science if assessments of scientific impact rely primarily on conventions associated with clinical trials or laboratory-based biomedical research. Before implementation, NIH should examine these potential distributional effects and ensure that a revised system does not compound disparities in access to mentorship, institutional support, or research funding.

The study by Ginther et al. (2011) illustrates the importance of preserving sufficiently granular review and award data. Drawing on NIH R01 records from fiscal years 2000 through 2006, supplemented by publication and related data, the study found that Black applicants were 13 percentage points less likely than White applicants to receive funding and remained 10 percentage points less likely after adjustment for measured applicant and institutional characteristics. Although this observational study neither examined the proposed categorical system nor established the causes of the disparity, it demonstrates the value of detailed data for identifying where disparities arise. If NIH proceeds with categorical reporting, it should retain the underlying scores and percentiles and assess outcomes by career stage, institution type, scientific field, applicant demographics, and other relevant factors.

APA Recommendations

- **Retain overall impact scores and percentiles.** If NIH introduces categories, it should report them alongside existing metrics while evaluating effects on applicant behavior, administrative workload, equity, and confidence in peer review. Overall impact scores and percentiles should be retained at least during a transition period.
- **Use a more informative alternative to three broad categories.** Score ranges or decile bands could reduce false precision while preserving the information applicants need to interpret outcomes and plan resubmissions.
- **Strengthen narrative feedback.** Summary statements should distinguish major from minor weaknesses, reflect the substance of panel discussion, and identify the factors that most influenced the overall assessment.
- **Provide actionable funding context.** NIH should use standardized language to indicate whether an application is highly competitive, may be fundable depending on Institute or Center priorities and available funds, or is unlikely to be funded without substantial revision.
- **Maintain an auditable decision record.** Institutes and Centers should document how peer-review outcomes, strategic priorities, portfolio balance, and budget considerations informed each funding decision, particularly when an award decision departs from peer-review standing.
- **Evaluate equity and administrative burden before full implementation.** NIH should pilot the approach, publish its evaluation criteria and findings, and assess differential effects on early-career investigators, under-resourced institutions, and scientific fields, including behavioral and psychological science.

Conclusion

APA urges NIH to withdraw the proposal in its current form and pursue reforms that clearly communicate uncertainty while preserving useful review information and accountable decision-making. APA welcomes continued engagement with NIH on approaches that strengthen scientific merit review and public trust. Please do not hesitate to contact Angela L. Sharpe, MG, Senior Director, Congressional & Federal Relations (asharpe@apa.org) if you have any questions.

APA.ORG

750 First Street, NE
Washington, DC 20002-4242

202.336.5500
202.336.6123 TDD

Sincerely,

Adrienne Stith

Adrienne Stith, Ph.D.
Executive Lead Psychologist, Science
American Psychological Association

Page | 4

References

Ginther, D. K., Schaffer, W. T., Schnell, J., Masimore, B., Liu, F., Haak, L. L., & Kington, R. (2011). Race, ethnicity, and NIH research awards. *Science*, 333(6045), 1015–1019. <https://doi.org/10.1126/science.1196783>

Mietchen, D. (2014). The transformative nature of transparency in research funding. *PLoS Biology*, 12(12), e1002027. <https://doi.org/10.1371/journal.pbio.1002027>

Recio-Saucedo, A., Crane, K., Meadmore, K., Fackrell, K., Church, H., Fraser, S., & Blatch-Jones, A. (2022). "What works for peer review and decision-making in research funding: A realist synthesis." ("Amanda Blatch-Jones, Simon Fraser, Hazel Church, Kathryn Fackrell ...") *Research Integrity and Peer Review*, 7(1), Article 2. <https://doi.org/10.1186/s41073-022-00120-2>