



Center for
Scientific Review

CSR Advisory Council Reviewer Evaluation Working Group Update

Lynn M. Yee, MD, MPH

Thomas J. Watkins Memorial Professor
of Obstetrics and Gynecology

Northwestern University Feinberg
School of Medicine

Delia Olufokunbi Sam, Ph.D.

DNDA Division Director
March 25, 2024



Acknowledgment: Reviewer Evaluation Working Group

CSR Advisory Council



Karen Anderson, M.D. Ph.D.
Arizona State University



Anton Bennett, Ph.D.
Yale University



Matthew Carpenter, Ph.D.
Medical University of South
Carolina



Narasimhan Rajaram, Ph.D.
University of Arkansas at
Fayetteville



Lynn Yee, M.D. M.P.H.
Northwestern University
Co-Chair

Ad-hoc



Catherine Fox, Ph.D.
University of Wisconsin–Madison



James Moody, Ph.D.
Duke University



Daniel Woo, M.D.
University of Cincinnati

NIH Staff



Stephanie Constant, Ph.D.
OD



Delia Olufokunbi Sam, Ph.D.
CSR
Co-Chair



Miriam Mintzer, Ph.D.
CSR



Elena Smirnova, Ph.D.
CSR

Charge and Guiding Principles

Charge: Develop a practical framework for evaluating reviewer performance – one that allows for a **uniform, structured, consistent, and transparent** approach to performance assessment and improvement.

Principles:

- Key words – **efficient, effective, and transparent**
- Should be tied to reviewer training and reviewer recognition – but focused on evaluation.
- Goal is to consider the characteristics of a strong review
- Not designed to compare reviewers to each other
- Not designed to penalize reviewers – focus is on helping reviewers improve
- Framework should supplement (and incorporate) what SROs are currently doing.
- Use conceptual framework developed by internal committee as a starting point to develop an actionable, measurable evaluation process

CSR Internal Evaluation Committee Conceptual Framework

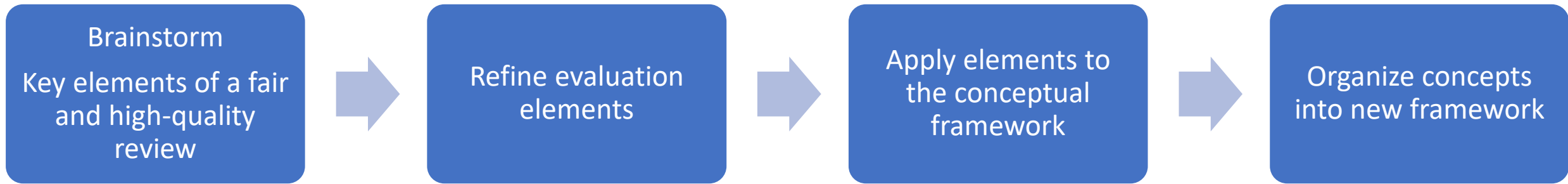
Domains that characterize a reviewer's performance



- 1 Fair
- 2 Knowledgeable
- 3 Evaluative

The mission of the Center for Scientific Review (CSR) is to see that NIH grant applications receive **fair, independent, expert, and timely scientific reviews — free from inappropriate influences —** so NIH can fund the most promising research.

Evaluation Framework Development Process



Brainstorming - Challenges to a fair and high-quality review

Examples generated:

- Scores without aligned and evaluative comments or not reflecting the relative strengths and weaknesses of the proposal
- Reviewers not abiding by NIH guidance
- Not understanding and appropriately evaluating each of the criteria
- Reviews excessively focus on editing or suggestions to improve the science rather than substantive evaluation
- Written or oral reviews may lack respectful tone or reviewer lacks flexibility

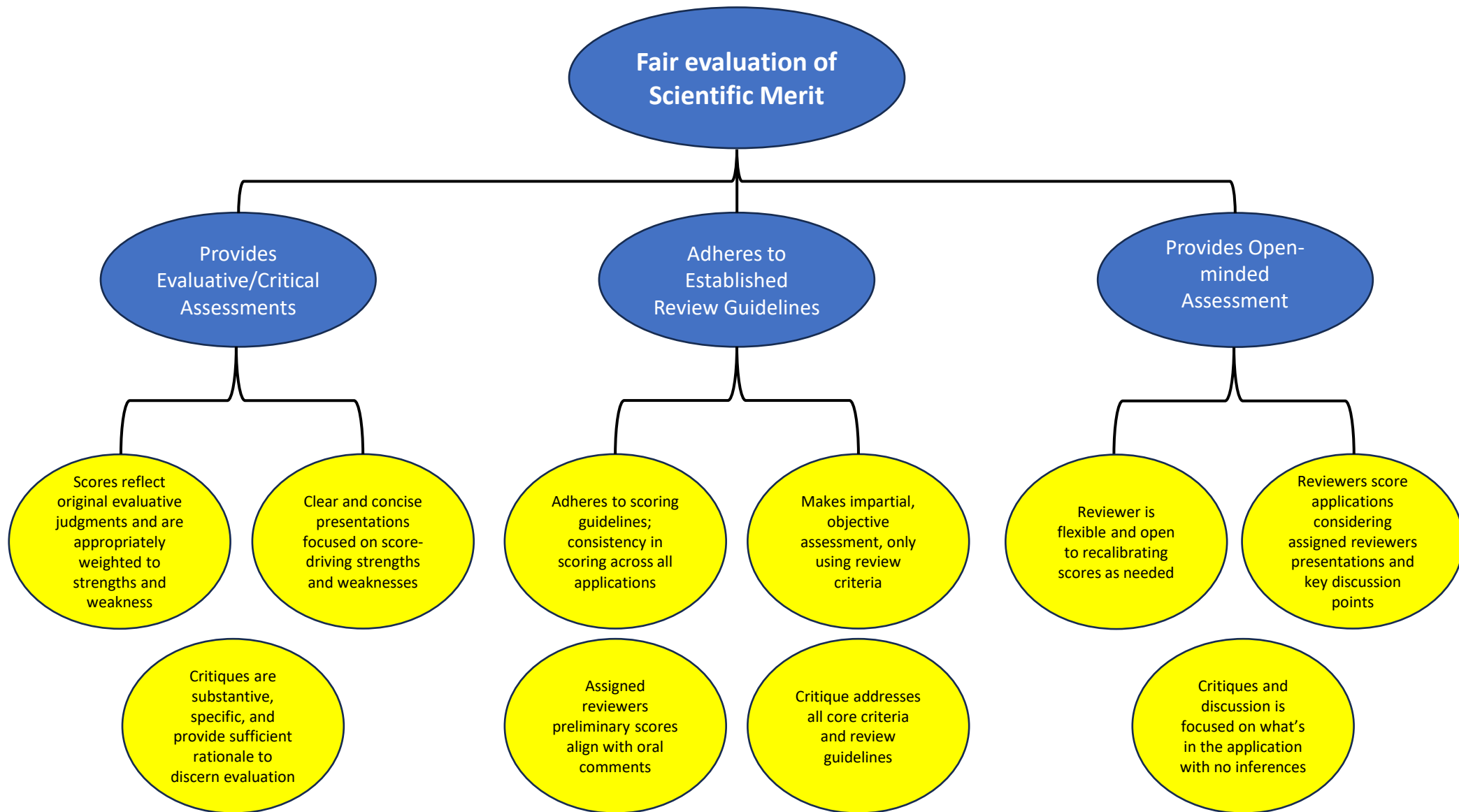
Evaluation Framework Development – Refining evaluation elements

- **Administrative Tasks vs. Conceptual Variables**
- **Identifying Key Variables for Evaluation**
- **Flexibility and Recalibration**
- **Confidentiality and Professionalism**

Applying elements to existing conceptual framework

- Initial approach developed was time/task based
 - Pre-meeting (critiques and preliminary scores)
 - At the meeting (discussion, final impact scores)
 - Post-meeting (amending critiques, aligning scores)
- Overlap in indicators for “knowledgeable” and “Evaluative”
- Worked through various iterations and ended up reorganizing conceptual domains using a graphic model
 - Goal was simplicity and practicality – discussed the option of sub-groups within each domain of “critique”, “scoring”, and “presentation” but opted against this for simplicity

Reviewer Evaluation Framework



Discussion and Questions

- Feedback on current framework
- Concerns and questions?

(Original conceptual model slides available if needed for Q&A)

		Critiques	Scoring	Meeting Discussion	Potential measures/metrics	Examples
Fair						
Adheres to Review Criteria and scoring guidelines		Critique addresses all core criteria and ONLY review criteria	Adheres to scoring guidelines; consistency in scoring across all applications	Assigned reviewers preliminary scores align with oral evaluation comments	Scoring analysis (within reviewer cluster; consistent outlier compared to other assigned reviewers); scores and comments are aligned	Example of not adhering to review criteria: Requests preliminary data for R21 or considering things that are not in the criteria (e.g., investigator must have prior funding, must be hypothesis driven etc.)
Open-minded, not limited to preconceived ideas, free of bias		Amends critiques as necessary to reflect any change in scoring and/or corrected commentary; professional tone?	Reviewer is flexible and open to recalibrating scores as needed	Open to other perspectives; compromise isn't necessarily the goal; discrepant opinions are fine but not to the point of obstructionism - respectful tone; free of bias	scoring analysis (pre-post score movement -- consistent lack of score movement compared to other reviewers?)	Need to be careful with this one -- an unintended consequence could be encouraging/promoting consensus.
Objective assessment of Scientific merit		Uses only information presented in the application	Scores reflect merit assessment	Discussion is focused on what's present in the application with no inferences or assumptions		does not infer "investigators can be trusted to figure out the details"; does not refer to information not provided in the application
Original assessment		critiques reflect individual reviewer assessment		Reviewers score applications considering assigned reviewer presentations and key discussion points	reviewers don't average assigned reviewers scores when they are divergent; scores out of range when warranted	significance/Impact assessment is not cut and pasted directly from application
Scientific						
Makes evaluative/critical judgments of strengths and weaknesses		Critiques are substantive, specific, and provide sufficient detail/rationale to discern meaning	Scores reflect evaluative judgments only and are appropriately weighted to the identified strengths and weaknesses (i.e., clear articulation of why an application was given a particular score	Clear and concise presentations focused on score-driving strengths and weaknesses. If changing a score based on discussion, rationale for this change is clearly stated verbally for the benefit of other study section members to understand. As unassigned reviewer, is engaged and adds value to panel discussions	comparing critique to application, significance to application, OI to application - percentage of descriptive vs evaluative comments	statements are supported by rationale (this is a weakness because.....). Does not try to improve the application, redesign project, or mentor applicant; does not include primarily descriptive information. Does not score to provide a range or to send the PI message; Significance/ impact assessment goes beyond restating Aims/goals,
Distinguishes between applications by scientific merit		Assessments are discerning, including regarding Significance	Scores differentiate between applications; clear articulation of why an application was given a particular score		can look at range of scores of a reviewer across meetings to see differentiation?	

Fair

The Conceptual Framework

Reviewer

- Adheres only to review criteria
- Open-minded, not limited to preconceived ideas, free of bias
- Uses only information presented in the application



Critiques

- Original evaluation of the application
- Focused on assessment of scientific merit, objective assessment
- Professional in tone

Scoring

- Scores and comments are aligned
- Consistency in scoring of all applications
- Adherence to scoring guidance

Meeting

- Clear presentations, focus on score-drivers
- Open to other opinions or viewpoints
- Respectful of other reviewers and applicants

Knowledgeable

The Conceptual Framework

Reviewer

- Applies scientific knowledge, expertise and experience to evaluating applications
- Communicates their expert assessment in critiques and presentation to a broad audience



Critiques

- Critiques are supported by scientifically rigorous judgments
- Critiques convey reasons for evaluation
- Critiques are substantive and insightful

Scoring

- Scores are explained - why an application was given a particular score is clearly articulated

Meeting

- Clear presentation of the scientific assessment to the panel
- Engaged and adding value to panel discussions

Evaluative

The Conceptual Framework

Reviewer

- Uses evidence based declarative statements of strengths and weaknesses
- Distinguishes between applications by scientific merit



Critiques

- Thorough and discerning assessments
- Specific and thoughtful statements
- Original assessments of significance

Scoring

- Scores differentiate between applications
- Clear articulation of why an application was given a particular score

Meeting

- Making own assessments – casting informed scores on all applications
- Not averaging assigned reviewers' scores when they are divergent