Distributed peer review: A DIFFERENT APPROACH TO APPLICATION REVIEW

Lakshmi Matukumalli
National Program Leader

Daniel Schmoldt
National Program Leader

Robbin Shoemaker
National Program Leader

Background

- Competitive grant applications submitted to NIFA are subject to external merit review (advisory)
- NIFA regulations
  - *Scientific peer review means an evaluation of the technical quality of a proposed project and its relevance to regional or national goals, performed by experts with the scientific knowledge and technical skills to conduct the proposed research work.*
  - *Peer reviewers means experts or consultants qualified by training and experience to give expert advice on the scientific and technical merit of applications or the relevance of those applications to one or more of the application evaluation criteria. Peer reviewers may be ad-hoc or convened as a panel.*
- NIFA application review—three mechanisms
  - Ad hoc review (electronic only)
  - On-site panels
  - Virtual panels

Background

- Scientific journal peer review
  - *Ad-hoc reviews are submitted by 2-4 experts.*
  - *Editor referees those independent reviews to arrive at an "acceptance" decision.*
  - *Principles of this peer-review process are: (1) fully independent review and (2) complete anonymity.*
- At one time, these same principles and process were applied to grant application review
- Of course, as the number of applications grew, this process readily became cumbersome and inefficient
- Convening a panel of experts to deliberate:
  - Requires relaxing fully independent reviews and complete anonymity
  - More efficient
  - Introduces "panel funding recommendation" to replace referee

Background

- Number of competitive programs and number of applications for funding have increased dramatically over the past decade
- The current review system is strained and much less efficient than it could be
- The reviewer community is strained
- The time and cost expended to select 10% (in many cases) of submitted projects for funding is not realistic (20-25%+ worthy of funding)
- We are looking for peer-review options
- There is no "gold standard" for peer review

What is distributed peer review?

- DPR is an implementation of *mechanism design*, which is a branch of mathematics, sometimes called reverse game theory
- In mathematics, a "game" is any interaction among two or more people (decision-making entities), it deals with groups
- Groups don't make decisions, they have emergent behavior that depends on both the decisions of individuals in the group *and* the rules by which they interact (including social dynamics)

What is mechanism design?

- The question posed in mechanism design is, what should be the rules of the game (the interaction) so that the result achieves a desired goal
- Application review is, mathematically, a "game"
- Goals include:
  - Quality application review and decision making
  - Efficiency
  - Timeliness
  - Etc.

Proposed process in a nutshell

- A pilot study will be conducted using applications submitted to:
  - Agriculture Systems & Technology – Bioprocessing & Bioengineering
  - Animal Health and Production and Animal Products - Tools and Resources for Animal Breeding, Genetics, and Genomics
  - Agriculture Economics and Rural Communities - Economics, Markets and Trade
- Applications are grouped much like panels
- PDs/reviewers declare conflicts with other applications in their group

Proposed process in a nutshell

- PDs/reviewers are assigned 7 non-conflicted applications from their group for written review
- PDs/reviewers provide an ordinal rank for the 7 applications they review
- Group composite ranking is computed
- Reviewing consensus incentives are calculated and applied to create a revised composite ranking
- NPLs make funding recommendations given reviews and ranking

But, here's the twist that makes it work

- PDs/reviewers do NOT rank applications according to their preference, rather they rank according to what they think will be the composite ranking (→ consensus)
- PDs/reviewers get bonus points depending upon how closely their ranking compares to the composite ranking—this is their incentive to provide fair and objective reviews of the other applications (moderate, localized reordering)
- PDs/reviewers who do not submit their reviews and rankings <u>on time</u> are disqualified

What do we mean…

- PDs do NOT rank applications according to their preference, rather they rank according to what they think will be the composite ranking
  - A PD's typical preference is to get an award
  - A PD might reason that assigning a low rank to very good applications could raise his/her ranking
  - We discourage this behavior by encouraging consensus ranking, and offering bonus points
  - A reviewer's best estimate of the group ranking is their unbiased assessment of the worth of each application—thus we encourage a "collective" ranking of the applications

Soooo---

- The mechanism design approach to application review is an ad hoc review approach on steroids
  - It simplifies the reviewer selection process
  - It promises to accelerate the review process
  - It offers an opportunity to provide applicants more comprehensive feedback (7 reviews, versus 3)
  - It maintains a modicum of anonymity
  - It asks applicants to contribute to the review process in proportion to the burden they impose on the community
  - Engages a larger segment of the science community in peer review, enhancing grantsmanship and scientific advancement

Let’s look at the details

After the submission deadline

- Both analytical and manual methods will be used to cluster all standard applications into relatively homogenous groups of 25-40 for each program.
- NIFA will email a welcome letter to PDs that outlines the review process for standard grants.
- Conference, seed, sabbatical, and equipment grant apps are excluded from this process.
- There will be a short deadline for each project to:
  - Identify someone on their team to be the sole reviewer (not a team review) and contact;
  - Supply NIFA with conflicts of interest for the other submitted applications in your grouping.

Reviewer assignment

- A list of "excluded" applications is generated for each reviewer.
- Based on these exclusions, each reviewer will be randomly assigned 7 projects for review.
- Assignments will be made using the same NIFA-developed software used by NSF in an earlier pilot of DPR.
- Each reviewer will receive their review assignments with instructions for completing reviews in PRS.
- DPR reviews are treated as ad-hoc reviews by PRS, changes allowed until the deadline.

Review completion

- A strict 4-week deadline will be enforced, after which, any reviewer with uncompleted reviews will have their application removed from peer review.
- Standard AFRI evaluation criteria will be used.
- Each review will consist of:
  - Written comments,
  - A scientific merit rating (E, VG, G, F, or P)—only used for analysis,
  - An ordinal ranking

Creating the composite ranking

- Modified Borda count is used.
  - Suppose a reviewer ranks applications, A>B>C>D>E>F>G, then assign ranks: G=0, F=1, E=2, D=3,…, A=6
  - Ties share the same rank, e.g., E=2.5, D=2.5; only "pair" ties are allowed.
  - The Borda count equals the sum of the scores across all reviewers
  - Since each application receives *m* reviews, the total score must lie between 0 and $m(m-1)$. Thus, the modified Borda score is the Borda count divided by $m(m-1)$.
- Applications are ranked by the modified Borda score

Incentivizing good reviewing

- To provide incentive for reviewers to be impartial, reviewers are given a reward that depends on how well they agree with other reviewers.
- The reward is proportional to how closely their ranking agrees with the composite ranking (determined by Borda scores).
- This encourages reviewers to consider opinions that others might have, and to address those viewpoints.

Computing the bonus points

- See details provided at: tinyurl.com/2016AFRI
- The final ranking is based upon the composite score plus the bonus points
- Bonus-related changes to the final ranking are moderate and local, and do not subvert the scientific-merit basis of peer review
- If everyone aligns well with the composite ranking, there will be no relative movement of position

So, what does this mean for proposal writing?

- As with any grants program, you should write your argument for funding (proposal), so that it can be understood and appreciated by a broad science audience.
- With seven reviewers for each application, not all reviewers will have the specialized expertise to understand the technical nuances of your science sub-field.
- There are no champions for your application, so you must convince each reviewer that your application is the best one they've read.

Evaluation plan

- Fairness and community acceptance—measured by feedback of reviewers
- Impact on application quality—measured by average scores
- Quality of the reviews—qualitative and quantitative measures
- Efficacy—measured by convergence to reasonable recommendations
- Cost—measured by comparison to panels
- Internal workload—measured by time spent

Summary

- This is a pilot, and it mirrors a similar pilot conducted by NSF
- There is no "gold standard" for peer review
- DPR has advantages, just as panel review has advantages
- For the reasons stated early on, we need to explore more efficient options for peer review
- DPR won't work for all programs and all peer-review scenarios

QUESTIONS
COMMENTS