



Strategies & Lessons
Learned from Implementing
External Peer Review Panels Online:
A Case Example from a National
Research Center

Presented by: Kelly Robertson

Team members: Daniela Schröter, Richard Zinser, Chris Coryn, &
Pedro Mateu

The Evaluation Center, Western Michigan University



Purpose of Project

- Government Performance and Results Act (GPRA) measures:
 - Relevance of the research to practice
 - Quality of disseminated products
- 5 panel studies will be held
 - 2 Face-to-face
 - 3 Online





THE EVALUATION CENTER

Panel Study

STRATEGIES & LESSONS LEARNED





Panelist Selection Initial

1. E-mailed letter to invite nominations for panel
2. Nominees contacted and sign conflict of interest statement
3. Selection based on criteria:
 - a) Expertise in field
 - b) Outside field (e.g., evaluator or researcher)
 - c) Prior experience on national panel





Panelist Selection

2nd Iteration and Beyond

1. Availability
2. Conflict of interest
3. Budget
4. Past performance
5. Participation in initial training





Panelist Selection

Lessons Learned

- Back-up panelists
- Diversity of panelists





Panel Logistics

Initial

- Face-to-face
- Individual ratings completed on-site
- All members rated all items ($n=2$)
- Panelist feedback survey
 - Logistics
 - Process
 - Strengths
 - Areas for improvement





Panel Logistics

2nd Iteration and Beyond

	2 ND	3 RD	4 TH
Items Reviewed	52	85	70
# Panel subgroups	2	3	3
# Panelists per group	4	3-4	3





Panel Logistics

Lessons Learned

- Face-to-face meeting good for building rapport
- Included picture cheat sheet during virtual panel study





Training of Judges

Initial

- Introduced nature and intent of study
- Provided detailed instructions for using evaluation instruments
- Worked through hypothetical case





Training of Judges

2nd Iteration and Beyond/Lessons Learned

- Instructions provided but no training on rating
- Training on virtual software annually





Calibration of Judges

Initial

- Training: Hypothetical case
- Panel study:
 - Independent ratings
 - Discussion
 - Re-rating





Calibration of Judges

2nd Iteration and Beyond/Lessons Learned

- Clarification of rating scales by developing precise categories

NA	Inadequate; fatal deficiencies		Barely adequate; some serious deficiencies		Adequate; minor, but nonfatal, deficiencies			Strong overall, but not exemplary; no deficiencies of consequence		Excellent overall; no deficiencies	
NA	①	②	③	④	⑤	⑥	⑦	⑧	⑨	⑩	

- Discussion of individual ratings to overcome discrepancies vs group average
 - Diversity of raters
 - Error (e.g., panelists may miss information)





Independent Ratings

- Initial:
 - Conducted on-site at EC
- 2nd Iteration and Beyond:
 - Online





Independent Ratings

Lessons Learned

- Checklist
- Requiring justifications
- Timelines: 1-2 weeks between submission of ratings & panel
 - Multiple deadlines and reminders





Consensus Seeking Initial

1. Group split into 2 subpanels and asked to rerate one study.
2. 2 subpanels came together to discuss new ratings and rationales.
3. Group deliberated on subpanel ratings and determined final ratings for each study.





Consensus Seeking

2nd Iteration and Beyond

Group review all independent ratings during virtual panel meeting

Item Reviewed Title
Author 1 & Author 2

● = Panelist Name ● = Panelist Name ● = Panelist Name

	NA	No merit; inadequate; fatal deficiencies		Limited levels of merit; weak; some serious deficiencies		Reasonably good; minor, but nonfatal, deficiencies			Very good, but not exemplary; no deficiencies of consequence		Excellent; exemplary; no deficiencies	
1. Relevance: Relevance of product for intended audiences.	NA	①	②	③	④	⑤	⑥	⑦	⑧	⑨	⑩	
Justification (Please indicate salient features, strengths, and weaknesses): There obviously is a need for common data collection methods. So this has relevance for administrators and policymakers The report is very relevant and addresses an issue that has the potential to cause harm to field. Strong need for commonality and longitudinal study											8.7	
2. Purpose: The degree to which the report has a clear purpose.	NA	①	②	③	④	⑤	⑥	⑦	⑧	⑨	⑩	
Justification (Please indicate salient features, strengths, and weaknesses): It was an exploratory study, and the purpose was clear - to see what states are up to with respect to current data systems. The purpose was well-stated Yes, the report is clear about the danger of incomplete data and is clear about the proposed solution to this problem I felt the purpose was very clear.											8.7	

▲ =Average Per Item / Average for Proposal Overall=8.1





Consensus Seeking Lessons Learned

Automated data visualization is time saving over alternative means

Item Reviewed Title

Author 1 & Author 2

● = Panelist Name ● = Panelist Name ● = Panelist Name

	NA	No merit; inadequate; fatal deficiencies	Limited levels of merit; weak; some serious deficiencies	Reasonably good; minor, but nonfatal, deficiencies	Very good, but not exemplary; no deficiencies of consequence	Excellent; exemplary; no deficiencies					
1. Relevance: Relevance of product for intended audiences.	NA	①	②	③	④	⑤	⑥	⑦	⑧	⑨	⑩
Justification (Please indicate salient features, strengths, and weaknesses):										8.7	
There obviously is a need for common data collection methods. So this has relevance for administrators and policymakers											
The report is very relevant and addresses an issue that has the potential to cause harm to field.											
Strong need for commonality and longitudinal study											
2. Purpose: The degree to which the report has a clear purpose.	NA	①	②	③	④	⑤	⑥	⑦	⑧	⑨	⑩
Justification (Please indicate salient features, strengths, and weaknesses):										8.7	
It was an exploratory study, and the purpose was clear - to see what states are up to with respect to current data systems. The purpose was well-stated.											
Yes, the report is clear about the danger of incomplete data and is clear about the proposed solution to this problem											
I felt the purpose was very clear.											

▲ =Average Per Item / Average for Proposal Overall=8.1





Technology Used

- Planning:
 - WhenisGood

- Ratings:
 - Hosted Survey
 - Survey Monkey

- Virtual Panel Study
 - WebEx





Reporting of Findings

- Government wanted summative information but provided formative as well, as part of the panel.
 - Present results in aggregate across research proposals and product categories (summatively).
 - Presented results for each unique proposal and product (formatively) to allow researchers to improve their materials.





Audience

OTHER STRATEGIES?

