National Institutes of Health
Office of Extramural Research

Enhancing Peer Review Survey Results Report




Published September 2017

National Institutes of Health
Office of Extramural Research

Executive Summary

The Phase III Enhancing Peer Review surveys, initiated in the fall of 2015, assessed the opinions of recent NIH grant applicants, reviewers, Advisory Council members, Scientific Review Officers (SROs) and Program Officers (POs) on NIH's peer review process. The purpose of the surveys was to assess the opinions of NIH stakeholders on how the peer review system functions in relation to the following core peer review objectives: appropriate reviewer expertise, fairness, objectivity and transparency, including robust indicators of scientific merit, and peer review burden.

There was broad agreement among applicants, reviewers and POs that reviewers had the appropriate expertise to review their assigned applications. Applicants who were funded were significantly more likely than applicants who were scored but not funded to agree that their applications were evaluated by reviewers with appropriate expertise.

Most responded favorably to questions regarding the fairness of peer review. SROs responded most favorably to these questions, indicating that reviewers followed appropriate review guidelines and criteria for the applications assigned to them, and that reviewer's initial critiques are focused on the factors that influenced their preliminary scores. Most advisory council members agreed that information in summary statement resumes and critiques are helpful for making advisory council recommendations.

The Phase II surveys identified the nine point rating scale as an area where NIH could carry out further improvements. In 2014, new scoring guidance was introduced and reviewers were encouraged to better utilize the entire range of scores. SROs agreed more often in Phase III than in Phase II that reviewers use the full range of the nine-point rating scale in the preliminary review of applications. Most reviewers also agreed that the 1 to 9 rating scale had sufficient range for them to communicate meaningful differences in the quality of the applications. However, most POs continued to disagree that the range of overall impact scores used is consistent among study sections.

Fewer SROs responded favorably to questions regarding peer review burden in Phase III in comparison to Phase II. Fewer SROs agreed in Phase III than in Phase II that draft summary statements are easy to correct for problems with critique content and that assigned reviewers completed all required elements in the critique template.

There was broad agreement that information in Summary Statements has improved since Phase II, including the consistency between individual criterion scores and the strengths and weaknesses described in critiques and the composition of bulleted comments. Most POs also agreed that information in Summary Statements is helpful in advising applicants. However, fewer applicants than POs agreed that information contained in Summary Statements is helpful in deciding the appropriate next steps.

Overall, most respondents rated the peer review process at NIH as fair and rated themselves satisfied with NIH peer review. Applicants responded least favorably than the other respondent groups to questions about their overall satisfaction with the peer review process, with almost half rating themselves dissatisfied with the NIH peer review process.

National Institutes of Health
Office of Extramural Research

Table of Contents

Report on the Results of the Enhancing Peer Review Surveys: Phase III

NIH strives for a peer review process that is fair, equitable, timely, and free of bias. In 2007, NIH launched the Enhancing Peer Review initiative, an undertaking to formally review and modernize its peer review process. NIH conducted a comprehensive assessment of the peer review system. Among the implementation goals of the Enhancing Peer Review initiative was the continuous evaluation of the peer review system; beginning in 2009, NIH has periodically sought feedback from participants in the peer review process through web-based surveys.

The Phase III Enhancing Peer Review surveys were conducted in late 2015. Survey questions are directed to five stakeholder groups and were designed to assess the overall state of the NIH peer review system. The surveys were developed and hosted by Research Triangle Institutes, International (RTI). RTI also prepared randomized, stratified samples of applicants and reviewers, distributed invitations to potential respondents, and collected and analyzed the survey data.

Number of Respondents:

2,866 applicants (response rate 36%), 2,559 reviewers (45%), 159 Advisory Council members (46%), 244 SROs (56%), and 416 POs (36%) responded to the surveys. The sampling strategies for all stakeholder groups are described in the Appendix.

Results of the Phase III surveys for each stakeholder group are discussed in the following sections:

- Reviewer Expertise
- Fair and Balanced Reviews
- Robust Scoring
- Transparency of Reviews
- Review Burden
- Overall Satisfaction

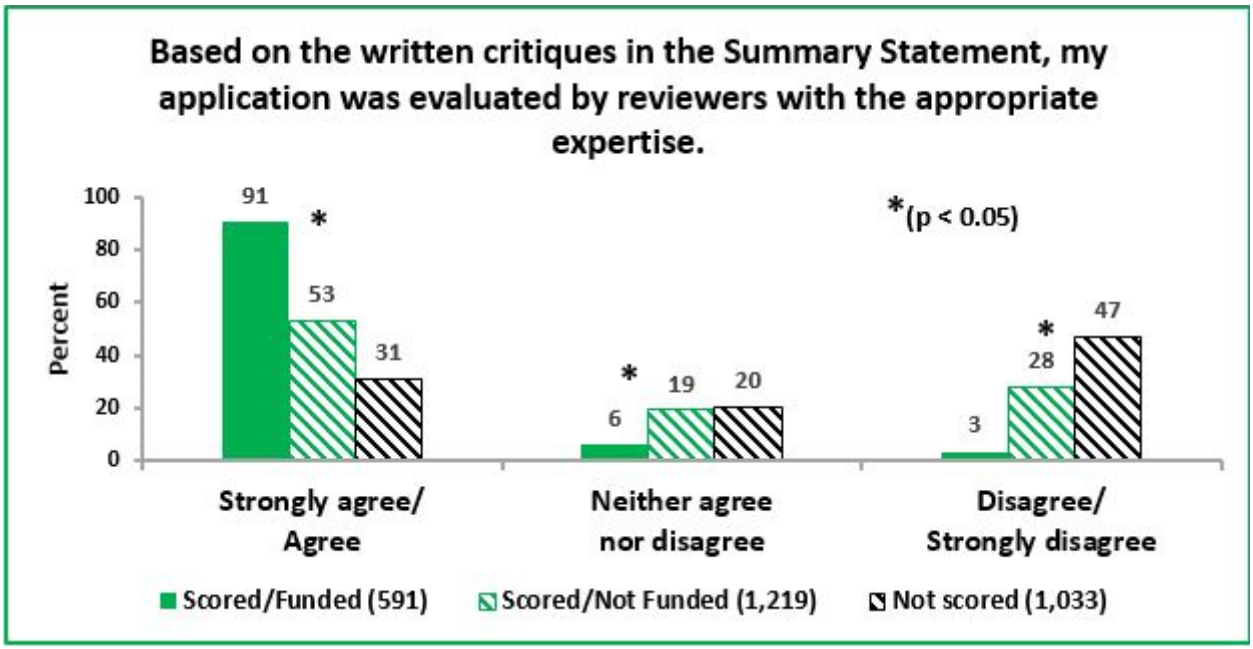
Reviewer Expertise

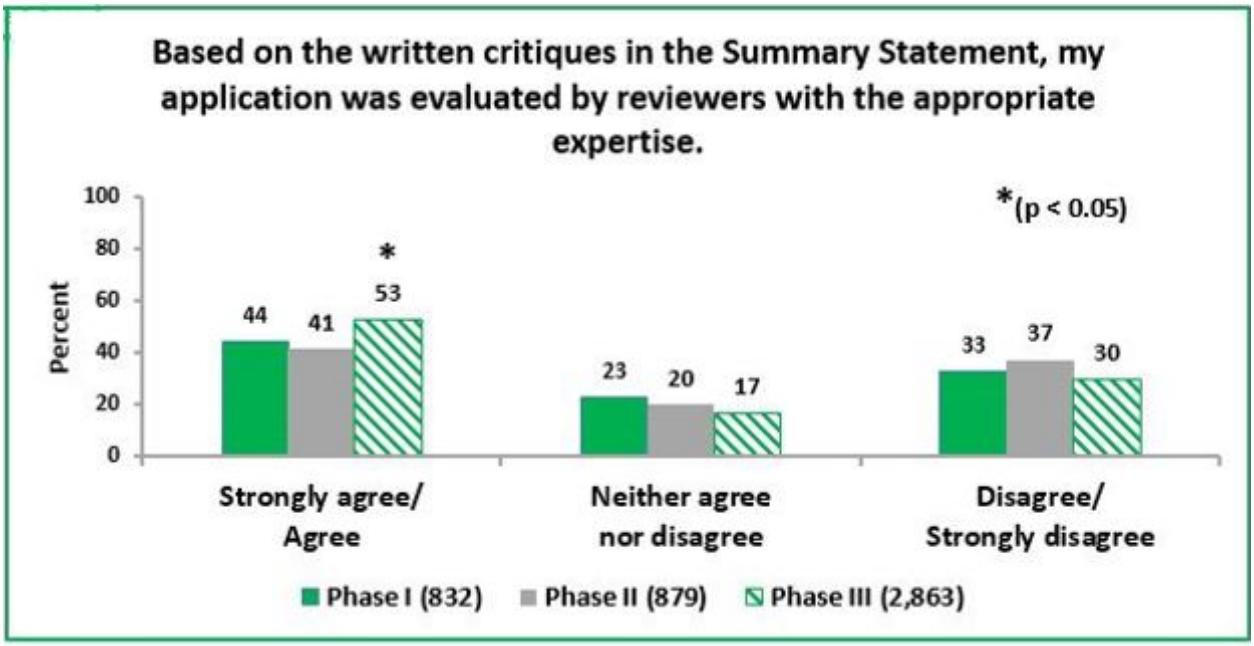
The ability to recruit and retain the most accomplished, broad-thinking and creative scientists to serve on study sections is an integral aspect of excellence in peer review. NIH policy specifies that the scientific expertise in the initial review panel be suitable for evaluating the potential impact of the proposed work. Peer reviewers donate an extraordinary volume of work to support NIH's peer review enterprise, and SROs tackle the expanding challenge of identifying reviewers with appropriate expertise who also are available to review applications on the designated date(s) and not in conflict with the applications that would be assigned to them. The questions shown below assessed stakeholders' opinions on reviewer expertise in the peer review process. More information is reported about annual peer review workload in the NIH Databook.

Applicants

Most applicants whose applications were funded (91%) agreed that their application was evaluated by reviewers with the appropriate expertise. Applicants whose applications were not scored (31%), or were scored but not funded (53%) were significantly less likely to agree (**Figure 1a, below)**.

- Overall, significantly more applicants agreed in Phase III (53%) than in Phase II (41%) that, based on the written critiques in the Summary Statement, their application was evaluated by reviewers with the appropriate expertise (**Figure 1b, opposite page**).
- Fifty-three percent of applicants agreed in Phase III that reviewers understood the significance of the proposed research, whereas 62% of applicants agreed in Phase II; however, the difference is not statistically significant. Significantly more applicants agreed that the reviewers understood their proposed approach in Phase III (57%) than in Phase II (38%) (**Figure 1c, opposite page**).





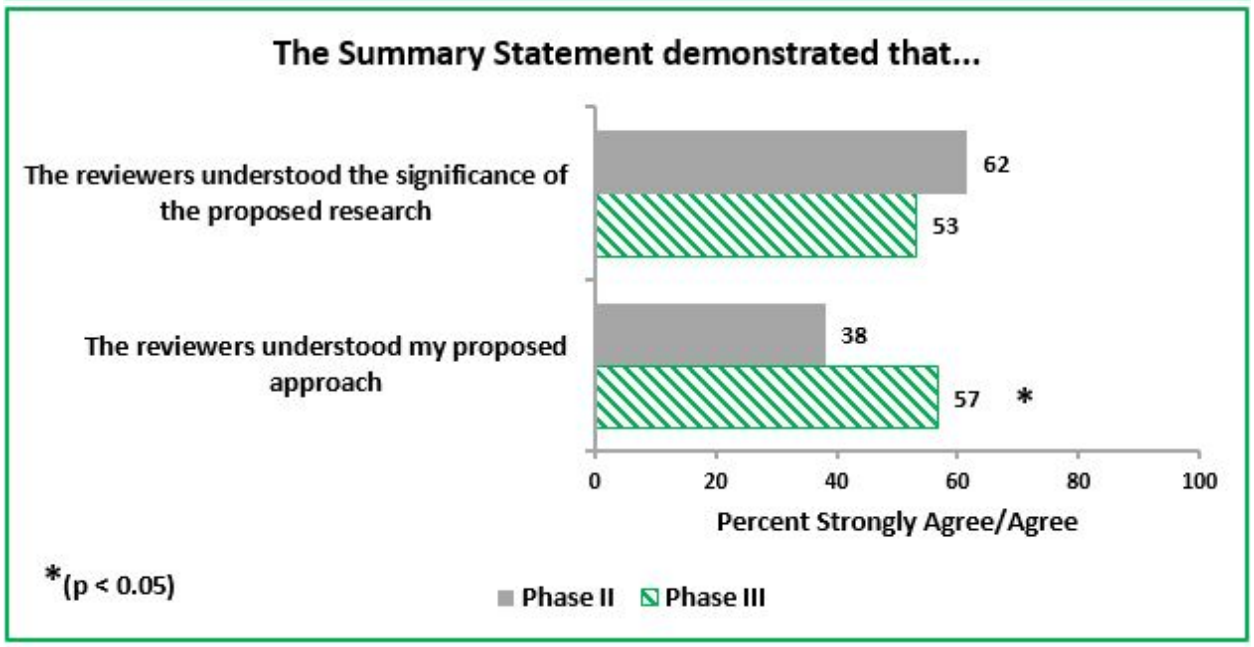


Figure 1a (opposite page): Applicants whose applications were funded agreed significantly more often and disagreed significantly less often than those whose applications were not funded that their application was evaluated by reviewers with the appropriate expertise. *Strongly Agree/Agree and Disagree/ Strongly Disagree groups with different review outcomes were significantly different in a pairwise fashion. **Applicants whose applications were funded were significantly less likely to select the neutral response (neither agree nor disagree) than applicants whose applications were not funded.
Figure 1b (top panel, above): Overall, applicants agreed significantly more often in Phase III than in Phase II that their application was evaluated by reviewers with appropriate expertise.
Figure 1c (bottom panel, above): Similar numbers of applicants in Phase II and Phase III agreed that reviewers understood the significance of the proposed research. However, significantly more applicants agreed that reviewers understood their proposed approach in Phase III in comparison to Phase II.

Reviewers and POs

Reviewers were asked to rate whether their own expertise was used appropriately in the peer review process and whether other reviewers were well qualified.

- Reviewers have consistently agreed (>90% across the three phases of surveys) that their own expertise is necessary and appropriately used in the review process.
- Reviewer' rates of agreement were similar across the three phases to a question about whether other group members seemed to be experts in their field. (**Figure 2**).
- POs' rates of agreement have steadily improved across the three survey phases to a question about whether the expertise of reviewers has been appropriate (**Figure 3**).

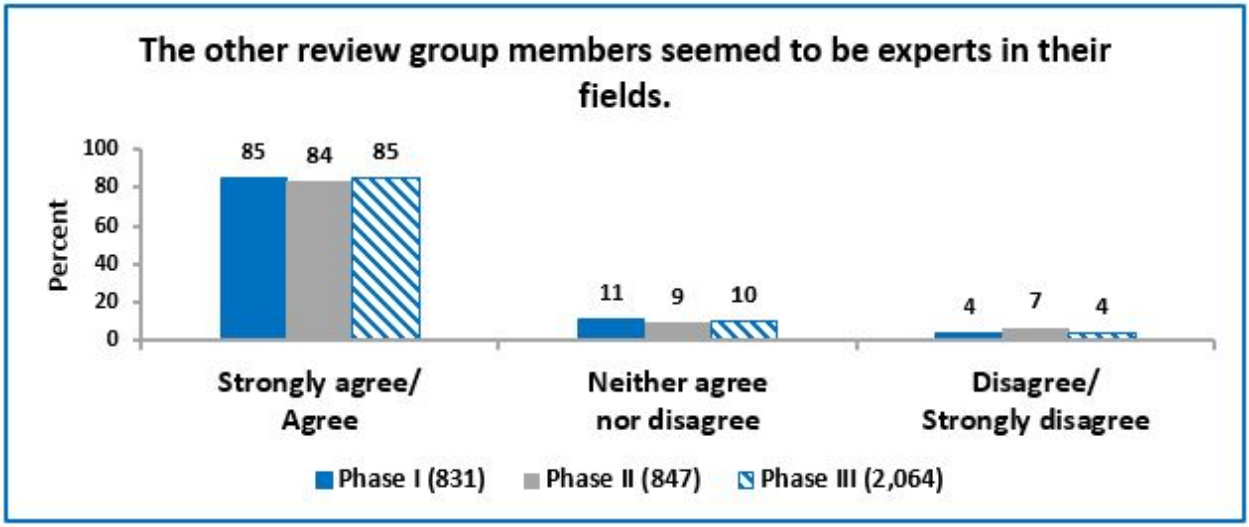


Figure 2: Most reviewers agreed that other review group members seemed to be experts in their field. There are no significant differences across phases.

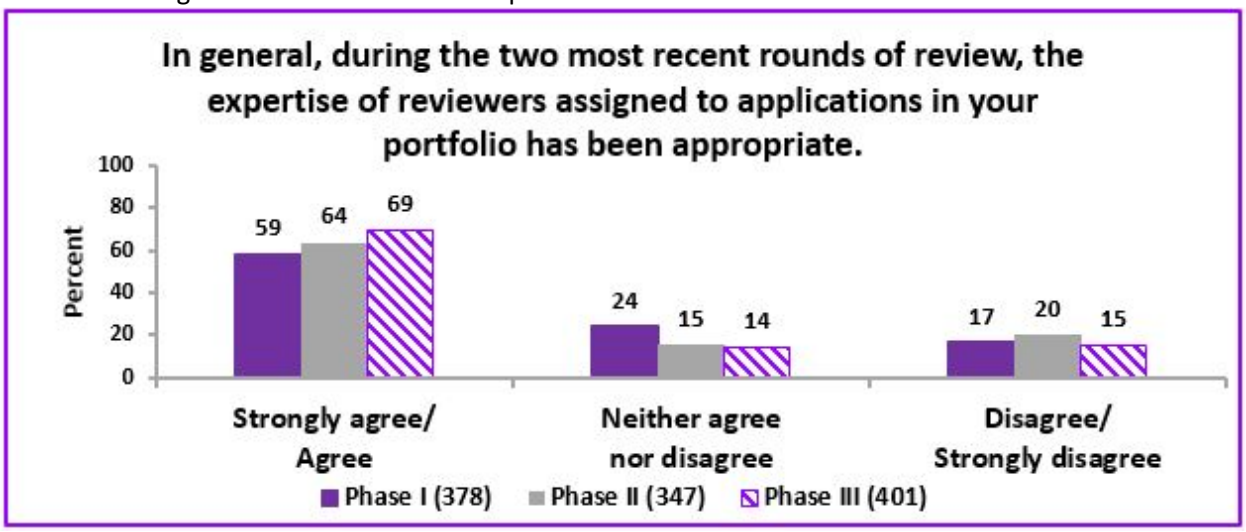


Figure 3: Most POs agreed that the expertise of reviewers assigned to applications in their portfolio has been appropriate. Although rates of agreement have consistently improved, adjacent phases did not significantly differ.

Fair and Balanced Reviews

The NIH peer review system strives to evaluate applications from all scientists in a fair and balanced manner. Grant applications submitted under the same funding opportunity announcement are evaluated using the same review criteria, reviewer guidelines, etc. Reviewers are directed to address in their critiques the most important aspects of each grant application, those that influence their ranking of the application's scientific merit. In addition, real or apparent conflicts of interest between reviewers and members of the research teams named on applications must be managed to avoid the possibility of inappropriate influences in the review process. A series of questions in the Enhancing Peer Review surveys assessed stakeholders' opinions about specific aspects of fairness in the NIH peer review process.

Reviewers

Significantly fewer reviewers agreed in Phase III (81%) than in Phase II (96%) that the information contained in grant applications is adequate for them to identify potential conflicts of interest in their assigned applications (**Figure 4**). This question did not appear on the Phase I surveys.

Reviewers may find it more difficult to identify conflicts in Phase III because of the increasingly collaborative nature of the research enterprise and the complexity of the conflict of interest rules.

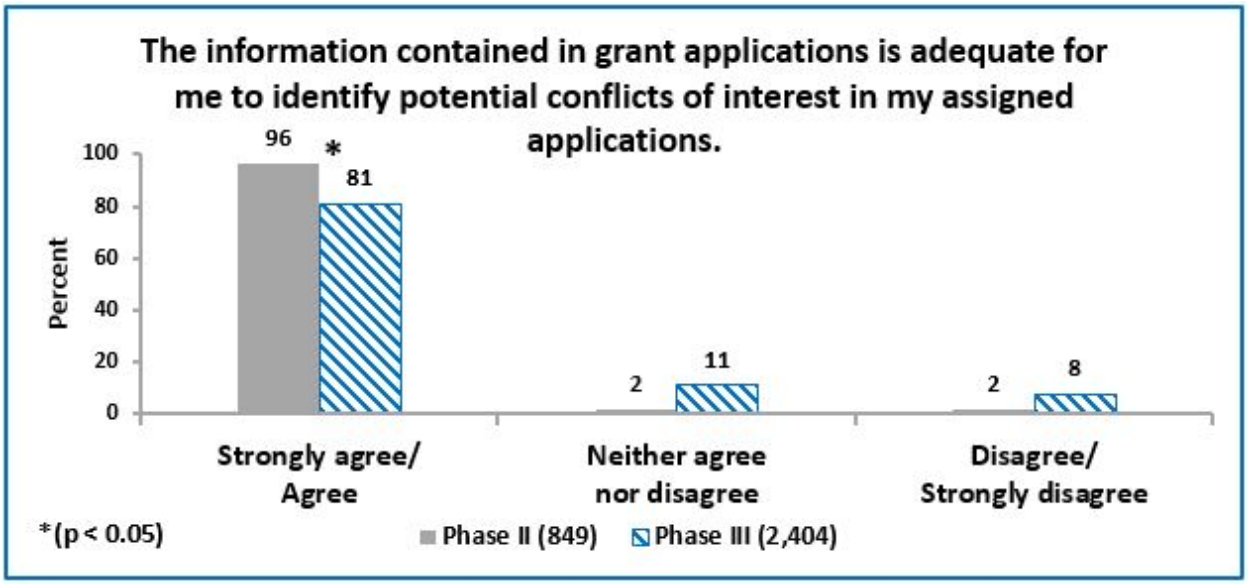
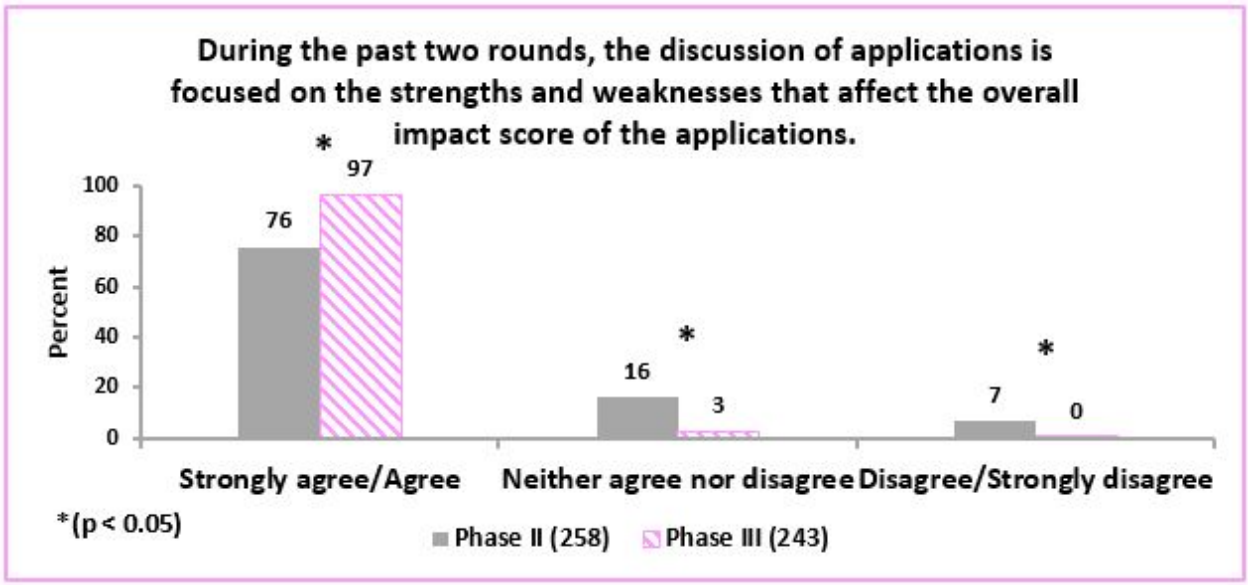


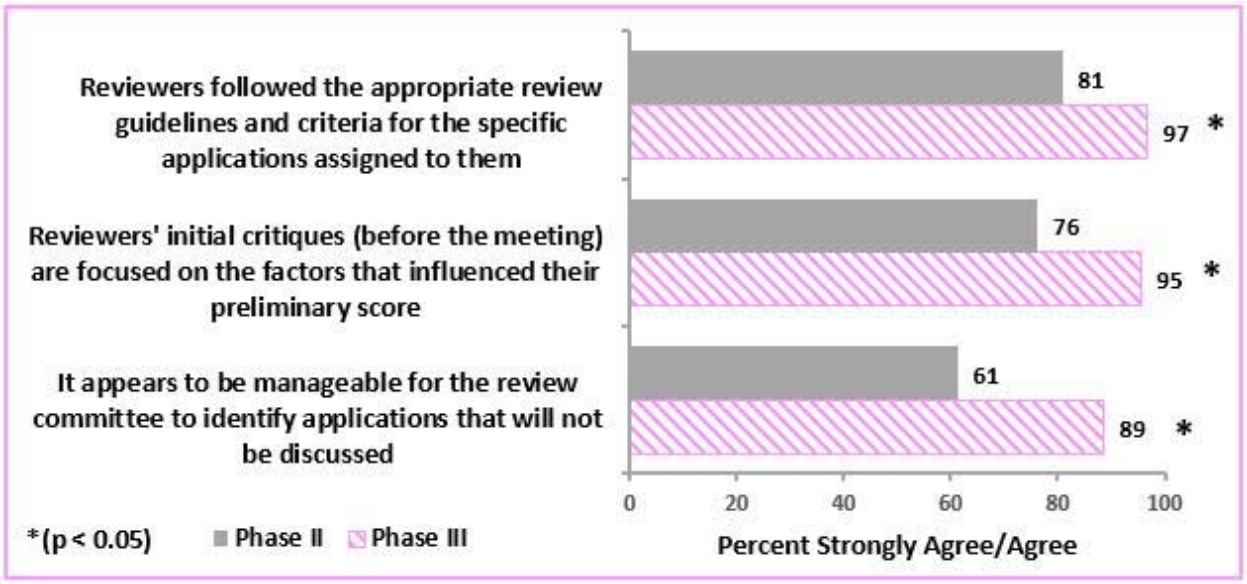
Figure 4: Significantly fewer reviewers agreed that the information contained in grant applications is adequate for them to identify potential conflicts of interest in their assigned applications in Phase III than Phase II.

SROs

SROs addressed a series of questions regarding reviewers' activities before and during study section meetings. SRO's rated reviewers as better focused on the most important, score-driving factors both before and during the study section meeting, and better able to achieve consensus.

- Significantly more SROs in Phase III (97%) than in Phase II (76%) agreed that during the past two rounds of review, the discussion of applications is focused on the strengths and weaknesses that affect the overall impact score of the applications (**Figure 5a**).
- SROs agreed significantly more often in Phase III that reviewers follow the appropriate review guidelines, focus on factors that influence their overall impact scores, and work effectively to identify applications that will not be discussed at the meeting (**Figure 5b**; opposite page).
- SROs agree significantly less often in Phase II that reviewers are willing to accept the applications assigned to them compared Phases I and III. This may be related to reviewers' transition to the shorter application format and increased reviewer workloads (**Figure 5c**; opposite page).





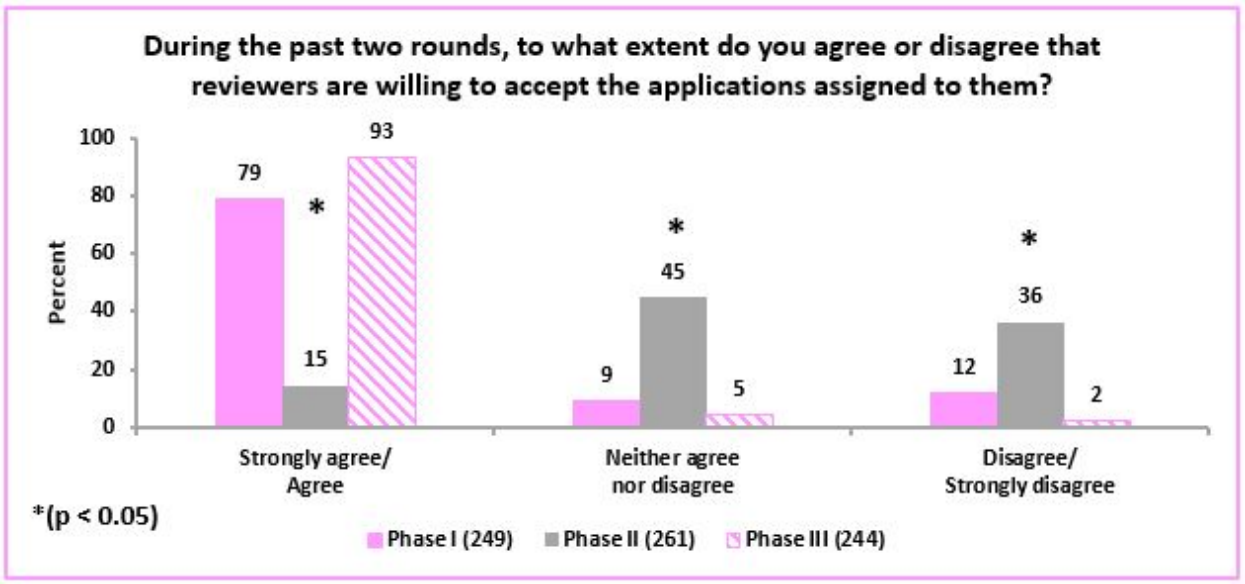


Figure 5a (opposite page): Significantly more SROs strongly agreed/agreed in Phase III than in Phase II that the discussion of applications has been focused on the strengths and weaknesses that affect the overall impact score of the applications during the past two rounds of review.
Figure 5b (top panel, above): SROs agreed significantly more often in Phase III in comparison to Phase II that reviewers follow the appropriate review guidelines, focus their critiques on the factors that most influence their scores, and work together as a committee to identify applications that will not be discussed during the meeting.
Figure 5c (bottom panel, above): SROs were significantly less likely in Phase II to strongly agree/agree that reviewers were willing to accept the applications assigned to them in comparison to Phases I and III. This is likely related to reviewers' transition to the shorter application format and increased reviewer workloads in Phase II.

POs

Questions about fairness for POs focused on their observations of the activities and work of reviewers, including the discussions at meetings and the content of critiques.

- Significantly more POs agreed in Phase III (65%) than in Phase II (49%) that during the two most recent rounds of review, reviewers followed the appropriate review guidelines and criteria for the specific applications assigned to them (**Figure 6**).
- Significantly more POs in Phase III (61%) than in Phase II (35%) agreed that during the past two rounds, there was clarity regarding the strengths and weaknesses that affect the overall impact score of the application (**Figure 6)**.
- Significantly, more POs agreed in Phase III (77%) than in Phase II (60%) that the discussions of applications are focused on the strengths and weaknesses that affect the overall impact score (**Figure 6**).

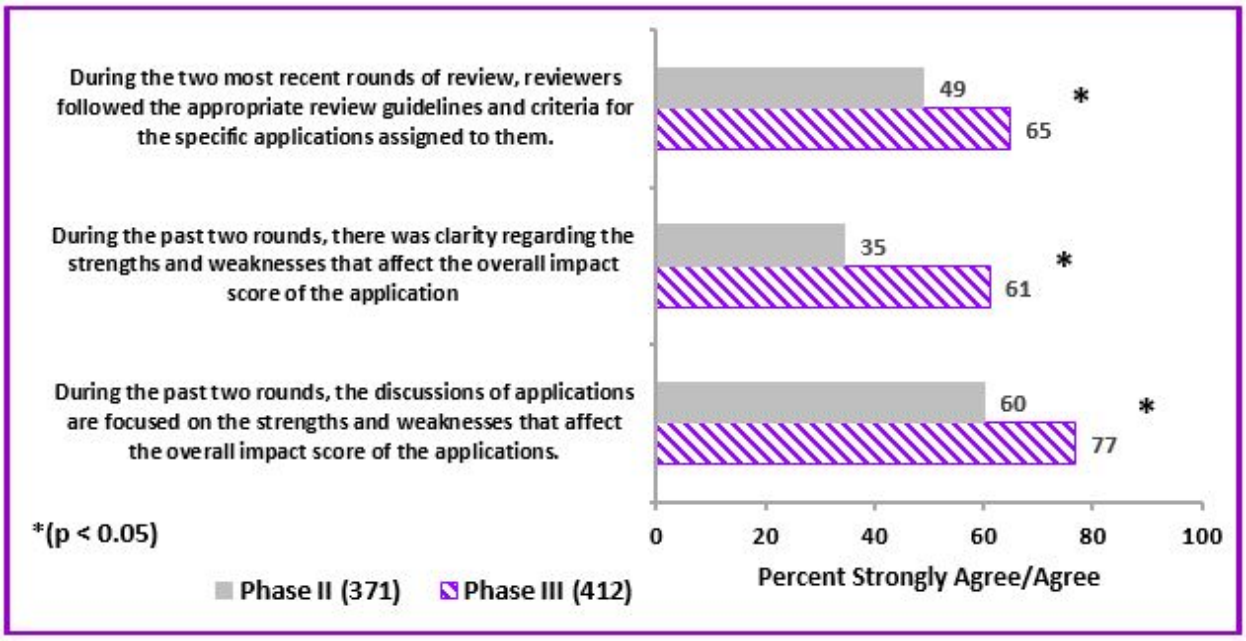
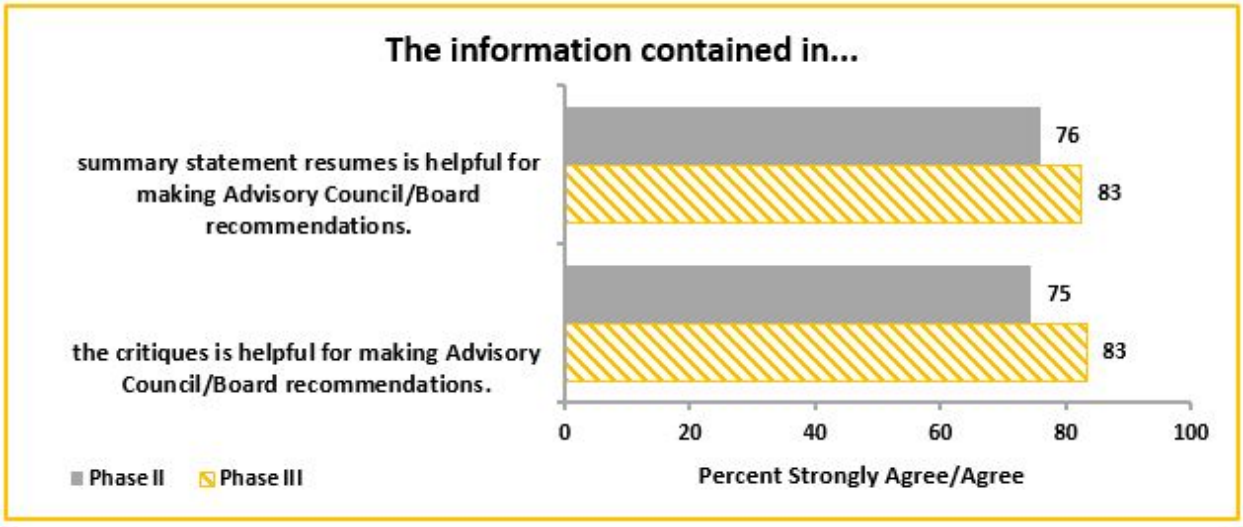


Figure 6: POs responded to questions about whether reviewers are using the appropriate review guidelines for the applications assigned to them, and whether reviewers' discussions and critiques are clear about the strengths and weaknesses that affect the overall impact score of the application. Reviewers strongly agreed/agreed significantly more often in Phase III in comparison to Phase II.

Advisory Council Members

Summary statements serve as the formal report summarizing the outcome of peer review for each application. Advisory council members use the information in summary statements to guide their funding recommendations to IC Directors.

- Most Advisory Council members strongly agreed/agreed in Phases II and III that information contained in summary statement resumes and critiques is helpful for making recommendations (**Figure 7a**).
- Advisory council members strongly agreed/agreed significantly more often in Phase III than in Phase II that the overall impact/priority score appears consistent with the information in the Resume and Summary of Discussion (**Figure 7b**).



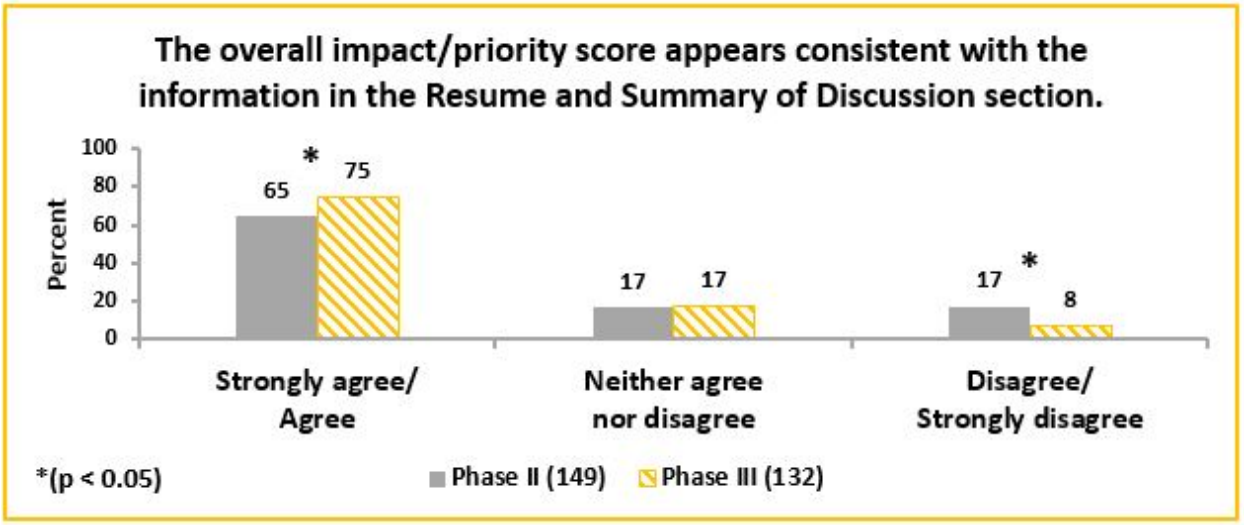


Figure 7a (top panel): Most Advisory council members agreed that the information in Phases II and III that summary statement resumes is helpful for making advisory council/board recommendations.
Figure 7b (bottom panel): Advisory council members strongly agreed/agreed significantly more often in Phase III in comparison to Phase II that the overall impact score appears consistent with the information in the Resume and Summary of Discussion.

Robust Scoring

Reviewers score each of five scored review criteria, and the overall impact of each application, on a nine-point rating scale that was introduced in 2009. Although reviewers who responded to this survey have consistently rated the new 9-point scoring system as adequate, analyses of overall impact score distributions have identified a tendency for reviewers to be overly generous in assigning scores, diminishing their ability to effectively distinguish the scientific impact of individual applications. Efforts were made in FY 2014 to encourage reviewers to use the full range of the 9-point scale in assigning scores. Those efforts included developing additional scoring guidance that encourages reviewers to use the entire range of scores to assess their assigned applications. In Phase III, the 9-point scoring system was assessed in the context of the new scoring guidance. Questions on the Enhancing Peer Review surveys assessed stakeholders' opinions on the adequacy of the nine-point rating scale, the range of overall impact scores, and whether tied scores posed a problem in making funding decisions.

Reviewers

- Seventy-nine percent of reviewers strongly agreed/agreed that the additional scoring guidance for research applications was useful for assigning overall impact scores. This rate of agreement is similar to reviewers' ratings of the previous scoring guidance assessed during Phases I (70%) and II (84%).
- Significantly more reviewers in Phase III (87%) than in Phase II (76%) strongly agreed/agreed that the 1 to 9 rating scale had sufficient range for them to communicate meaningful differences in the quality of applications (**Figure 8**).

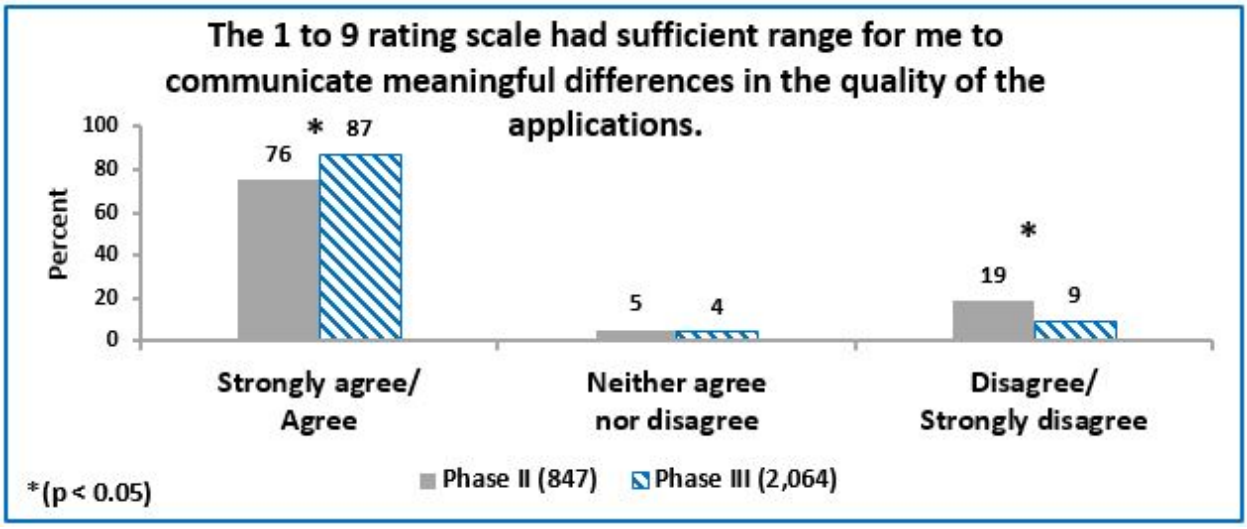


Figure 8: Significantly more reviewers agreed that the 1 to 9 rating scale had sufficient range to communicate meaningful differences in application quality in Phase III in comparison to Phase II.

SROs

SROs strongly agreed/agreed significantly more often in Phase III (65%) than in Phase II (43%) that reviewers used the full range of the 1 to 9 rating scoring scale for preliminary scores (**Figure 9**).

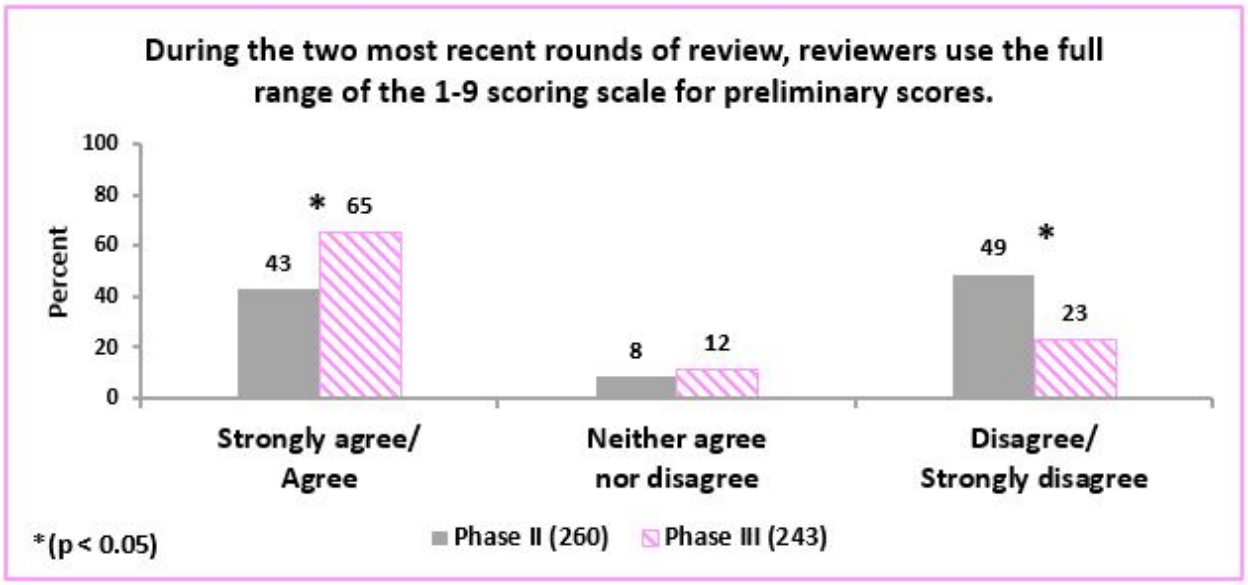
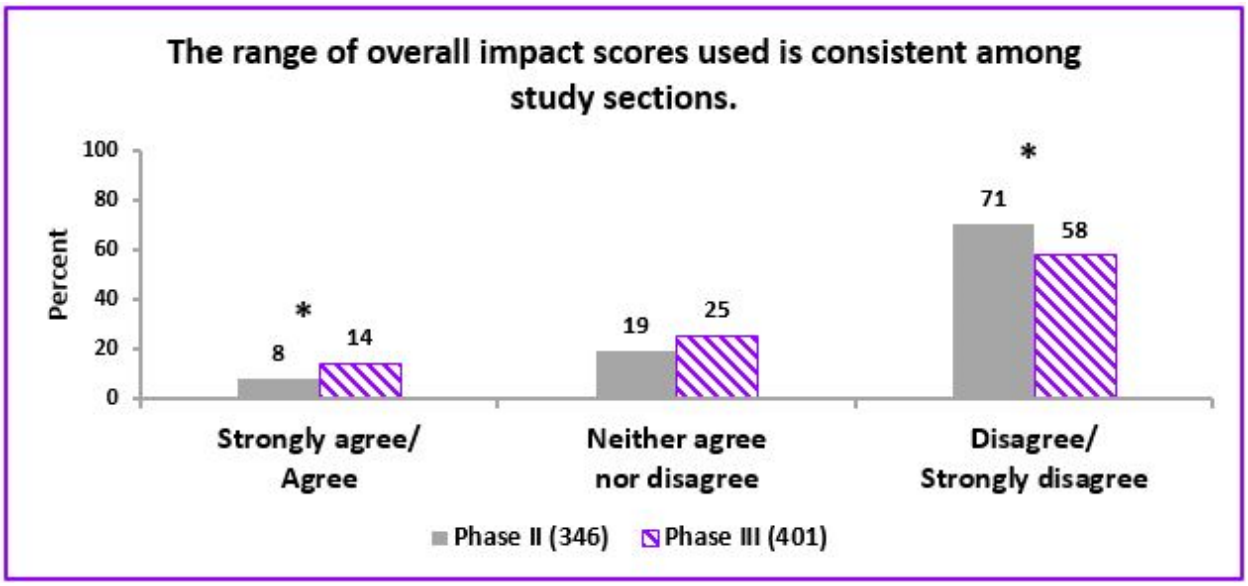


Figure 9: SROs agreed significantly more often Phase III in comparison to Phase II that reviewers use the full range of the 1 to 9 rating scoring scale for preliminary scores.

POs

Overall impact scores are percentiled to help normalize the rankings from hundreds of standing study sections. When some study sections assign scores more generously than others, the percentile scores can appear inconsistent across study sections, making it difficult to discern from scores alone which applications are most meritorious. The Enhancing Peer Review surveys queried Program Officers about the extent to which they observe consistent scoring practices across study sections and whether the summary statements are useful for reconciling tied scores.

- Although the rate of favorable responses was significantly higher in Phase III, most POs continued to disagree/strongly disagree (58%) that the range of overall impact scores used is consistent among study sections. (**Figure 10a; top panel**).
- Significantly more POs in Phase III (39%) than Phase II (25%) agreed that information contained in Summary Statements is useful for resolving tied scores (**Figure 10b; bottom panel**).



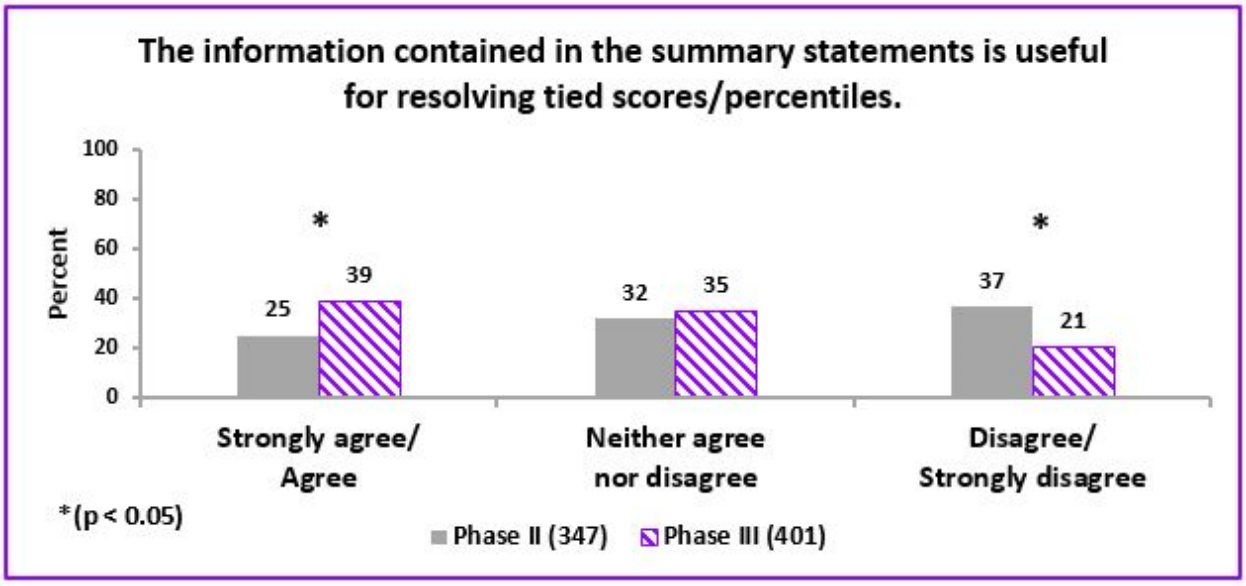


Figure 10a (opposite page): PO response to questions about whether the range of overall impact scores is consistent among study sections. Most POs continued to disagree that the range of overall impact scores used is consistent among study sections in Phase III in comparison to Phase II.
Figure 10b: Significantly more POs agreed that the information contained in Summary Statements is useful for resolving tied scores/percentile.

Advisory Council Members

Advisory council members in Phase III (62%) in comparison to Phase II (41%) agreed that during the most recent council rounds, the number of ties among the overall impact/priority scores and percentile rankings for applications has not been a problem in making advisory council/board recommendations (**Figure 11**).

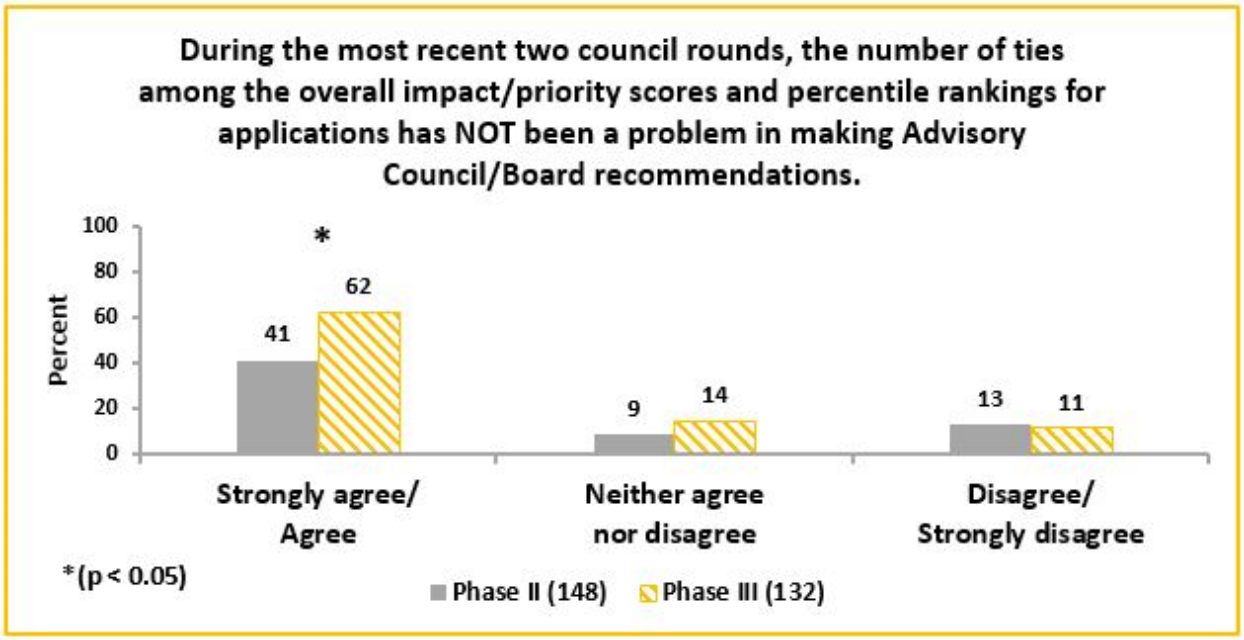


Figure 11: Advisory Council members strongly agreed/agreed significantly more often in Phase III than in Phase II that information contained in Summary Statements is useful for resolving tied scores/percentiles.

Transparency of Reviews

The NIH peer review process offers transparency to stakeholders by providing information about the factors that affected the review outcome in a Summary Statement. The Summary Statement provides a narrative critique and criterion scores from the assigned reviewers. If the application was discussed and scored during the review meeting, the summary statement includes the overall impact score, the average of scores from all reviewers who voted on the application, its corresponding percentile score, and a Resume and Summary of Discussion, prepared by the SRO. Questions in the Enhancing Peer Review surveys assessed stakeholders' opinions about how well the information contained in summary statements informs their next actions: recommending applications for award, advising applicants of their future options for their applications, and identifying problem areas in the proposed research that might be corrected through revisions.

Applicants

In the Phase III applicant survey, as in previous phases, there were significant differences in the responses of applicants with different review outcomes. Overall, 53% of applicants strongly agreed/agreed that the information in their summary statement helped them to focus on problem areas in the application. Favorable responses were more prevalent among funded applicants than applicants whose applications were not funded and those whose applications were not discussed (**Figure 12**).

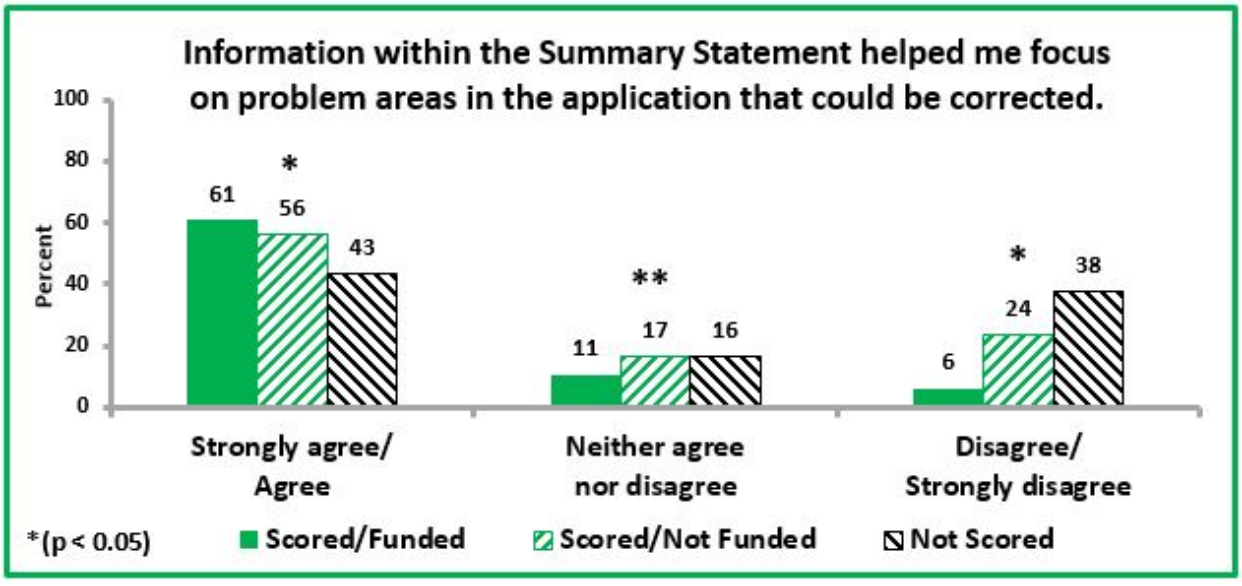


Figure 12: Applicant responses to a question about whether information within the Summary Statement helped them focus on problem areas in the application. *Applicants whose applications were funded strongly agreed/agreed significantly more often than those whose applications were scored but not funded, and those whose applications were not assigned a score. **Applicants whose applications were funded were significantly less likely to select the neutral response (neither agree nor disagree) than applicants whose applications were not funded.

POs and SROs

Program Officials are the designated contacts for applicants who seek advice about their possible next steps after learning their review outcome. The surveys queried POs' about the helpfulness of summary statements for advising applicants, and how well summary statements convey information about reviewers' assessments of scientific merit. Overall, there was improvement in POs ratings of the information contained in summary statements in comparison to their responses in Phase II.

- POs strongly agreed/agreed more often in Phase III (75%) than in Phase II (62%) that the information contained in the Summary Statements is helpful for advising applicants about appropriate next steps (**Figure 13**).
- POs were asked a series of questions about critique content. They strongly agreed/agreed more often in Phase III in comparison to Phase II that the strengths and weaknesses described in reviewers' written critiques reflect the discussions at the meetings (69% in Phase III vs. 57% in Phase II), individual criterion scores are consistent with the strengths and weaknesses described in the critiques (43% in Phase III vs. 25% in Phase II), and bulleted comments reflect complete, well composed thoughts (49% in Phase III vs. 23% in Phase II; **Figure 14**).
- Most SROs (75%) and POs (63%) agreed that the critiques are generally helpful for understanding how the five criteria that are assigned criterion scores contributed to the overall impact score (**Figure 15**). POs were significantly more likely to strongly agree/agree with this statement in Phase III than Phase II (41%).

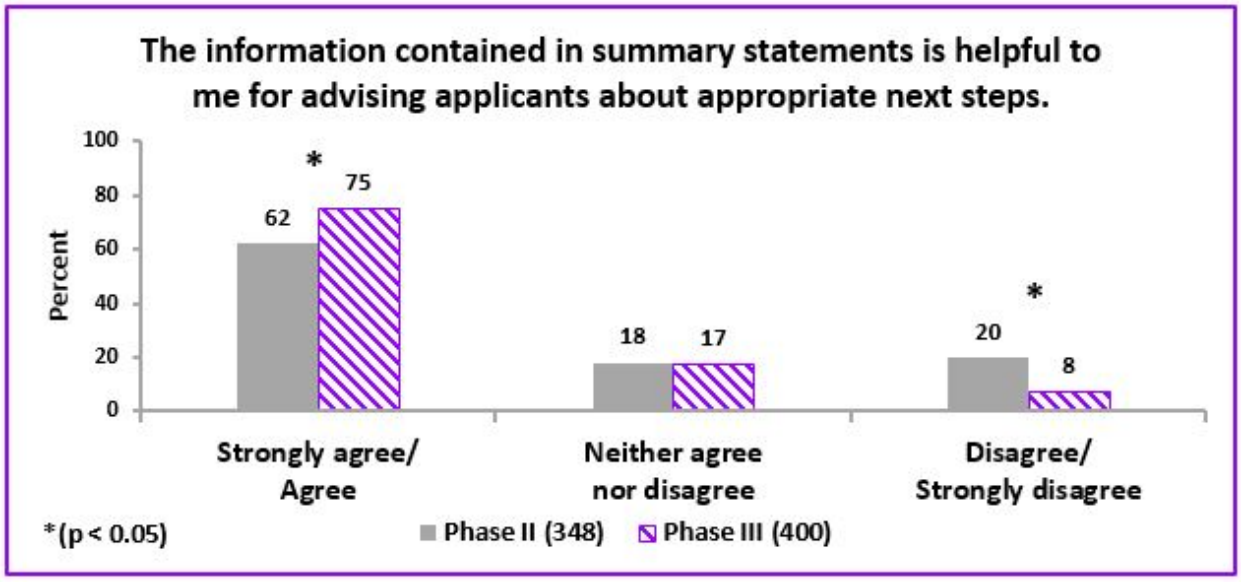


Figure 13. Program Officers' responses to a question about the helpfulness of summary statements for advising applicants. POs were more likely to strongly agree/agree that summary statements were helpful in Phase III than in Phase II.

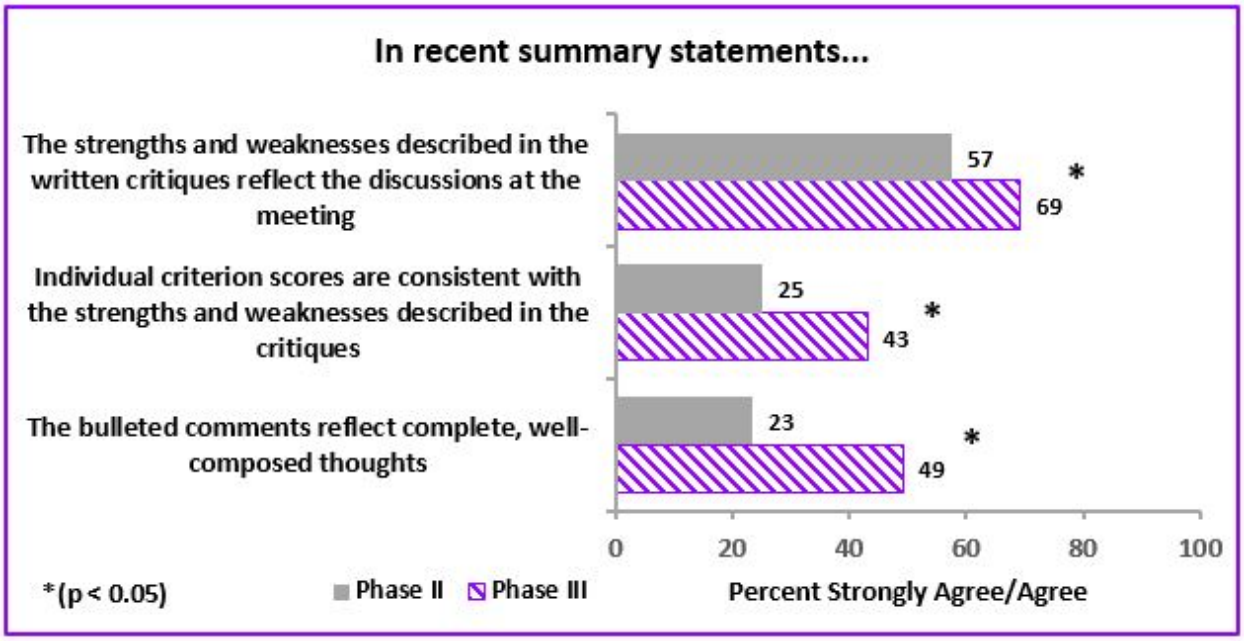


Figure 14: PO responses to questions about reviewers' written critiques provided in summary statements. POs were significantly more likely to strongly agree/agree that the written critiques reflect the discussions at the meeting, individual criterion scores are consistent with the strengths and weaknesses described in the critiques, and bulleted comments reflect complete, well-composed thoughts.

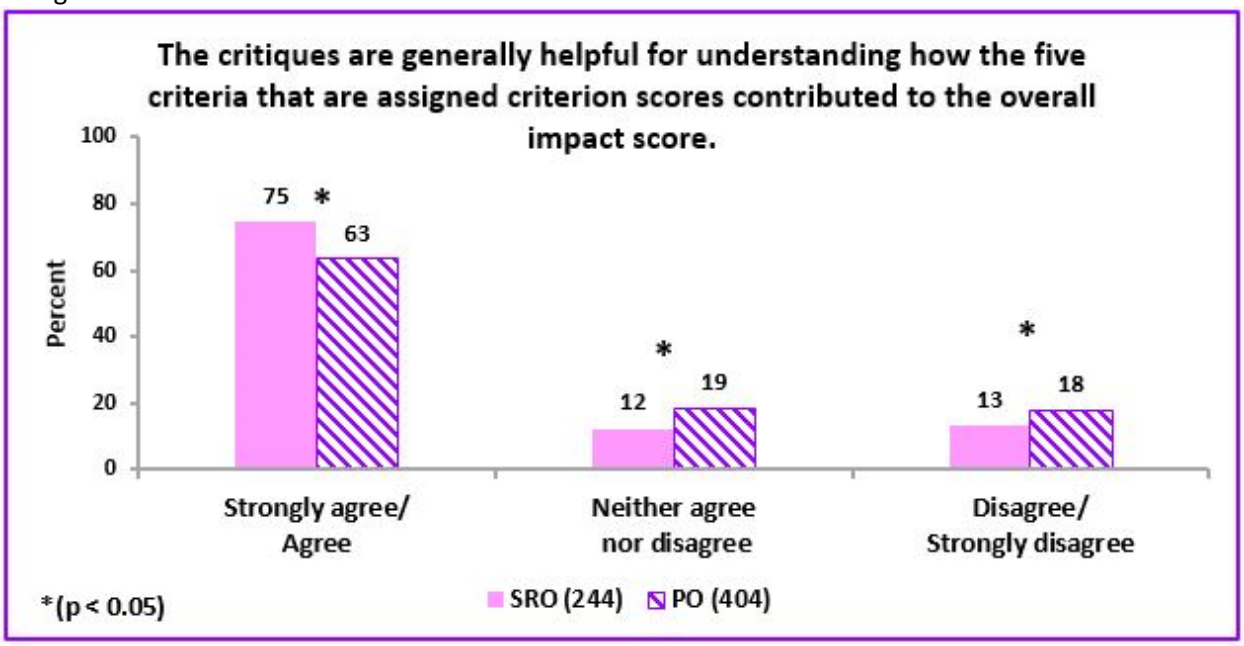


Figure 15: SRO and PO responses to a question about whether the critiques are generally helpful for understanding how the five criteria that are assigned criterion scores contributed to the overall impact score. Significantly more SROs than POs strongly agreed/agreed with the statement.

Advisory Council Members

Like POs, Advisory Council members responded to questions about reviewers' critiques with significantly more favorable ratings in Phase III.

Advisory Council members strongly agreed/agreed significantly more often that individual criterion scores are consistent with the strengths and weaknesses described in the critiques in Phase III (64%) in comparison to Phase II (48%), and that bulleted comments reflect complete, well composed thoughts (55% in Phase III vs 43% in Phase II; **Figure 16**).

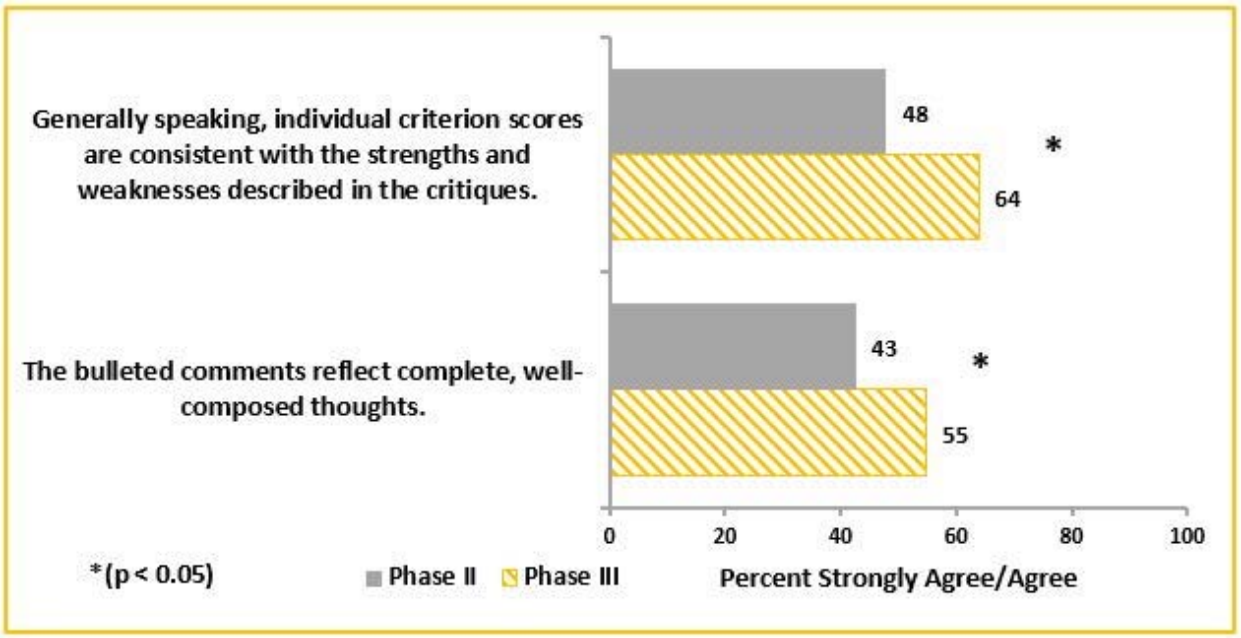


Figure 16. Advisory Council members' responses to questions about reviewer critiques. They were more likely to strongly agree/agree with the statements in Phase III compared to Phase II.

Burden of Review

The NIH continues to seek new approaches to reduce peer review burden while maintaining the core values of peer review. The NIH supports a fully automated web-based system for SROs to securely disseminate meeting materials to reviewers, and collect critiques, scores, confidentiality agreements, and other materials from reviewers. NIH develops extensive guidance materials for reviewers and applicants on peer review policies and procedures. Questions on the Enhancing Peer Review surveys queried SROs about their perceptions of reviewers' delivery on the products required from them. POs and reviewers were asked about the usefulness of the resources provided by NIH to support reviewers and applicants. Reviewers were also asked about whether the format and duration of the discussions at the study section meeting were sufficient for reviewers not assigned to an application to cast well-informed votes and whether an appropriate amount of time was spent on the potential impact of applications.

SROs and POs

- Significantly fewer SROs strongly agreed/agreed that draft summary statements are easy to correct for problems with critique content in Phase III (47%) in comparison to Phase II (61%) (**Figure 17**).
- SROs responded similarly in Phases II and III that draft summary statements are easy to correct for problems with critique format (**Figure 17**).
- Significantly fewer SROs strongly agreed/agreed in Phase III (49%) compared to Phase II (65%) that assigned reviewers completed all required elements in the critique template (**Figure 17**).
- SROs responded similarly Phase II (71%) and III (67%) that they contacted 1 - 24 percent of their reviewers after the meeting for updated critiques or criterion scores **(Figure 18).**
- Significantly more POs in Phase III (64%) than in Phase II (49%) strongly agreed/agreed that information resources for applicants on the OER website are useful as provided (**Figure 19**).

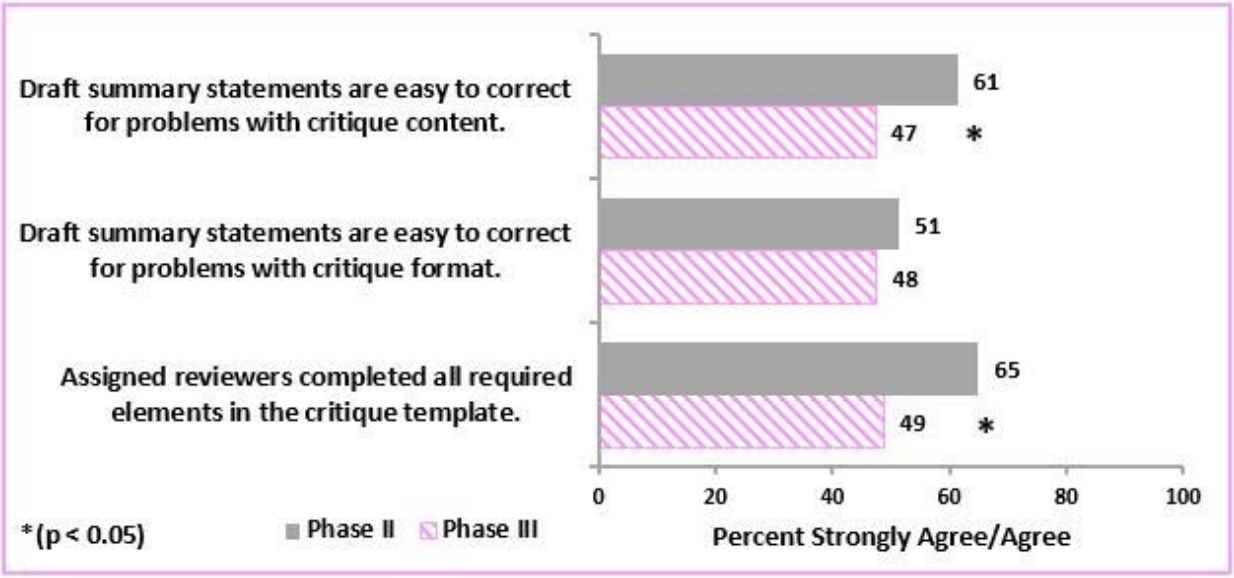


Figure 17 (bottom of previous page): Significantly fewer SROs strongly agreed/agreed in Phase III than in Phase II that draft summary statements are easy to correct for problems with critique content and assigned reviewers complete all required elements in the critique template. SRO responses were similar across phases to a question about whether draft summary statements are easy to correct for problems with critique format.

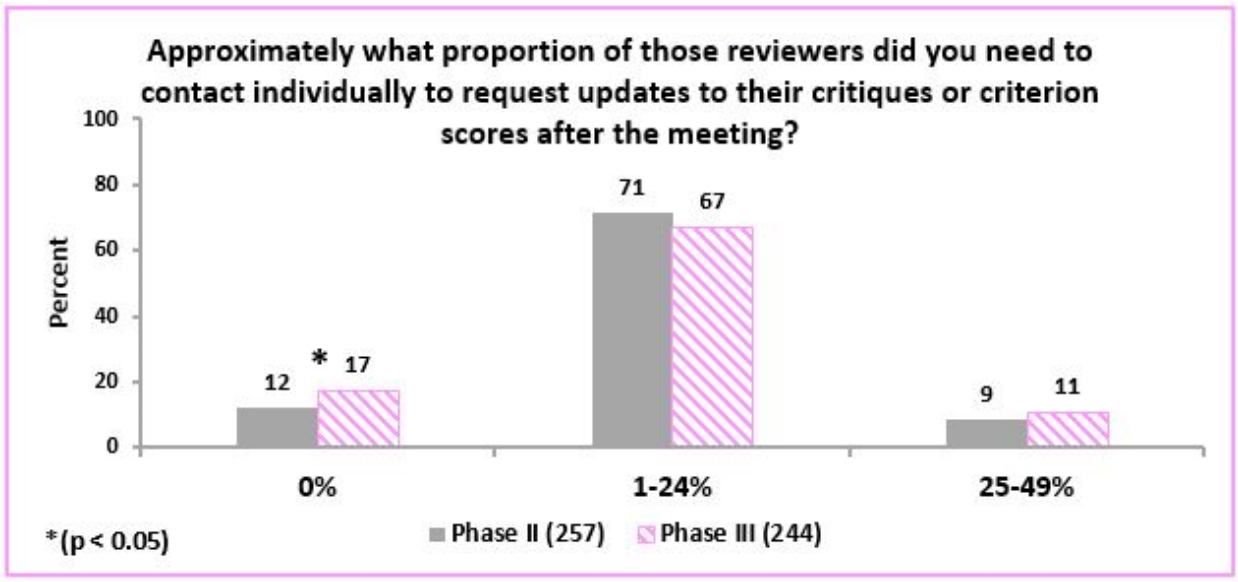


Figure 18: Most SROs responded that they needed to contact approximately 1 -24 % of their reviewers to request updated critiques or criterion scores after the meeting. However, significantly more SROs responded in Phase III than in Phase II that they needed to contact none of their reviewers for updates.

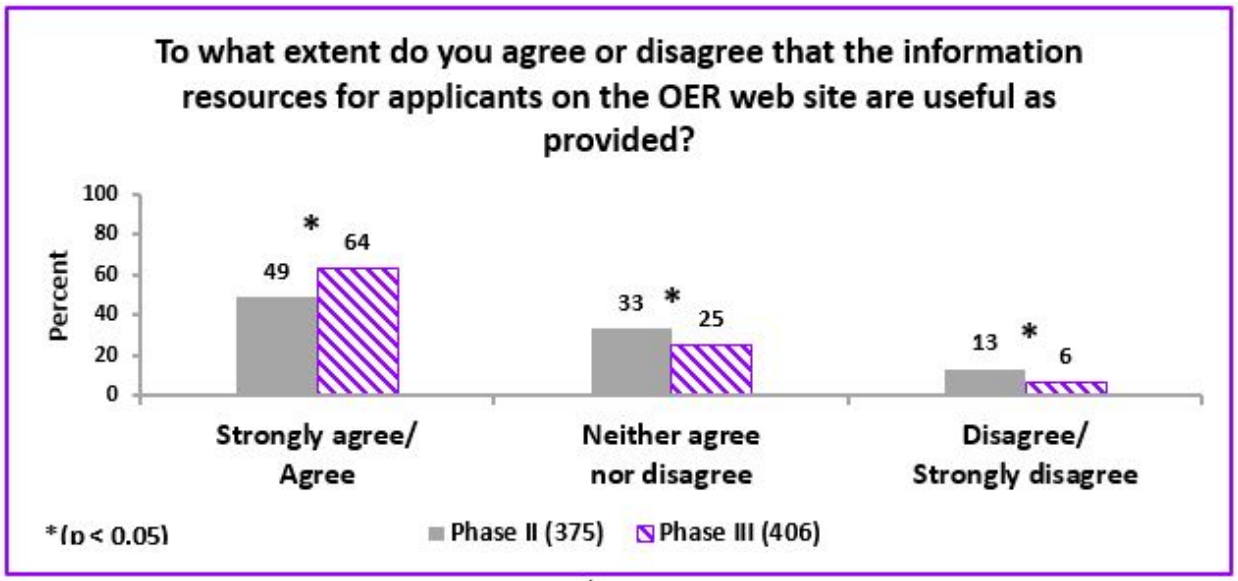


Figure 19: Significantly more POs strongly agreed/agreed in Phase III in comparison to Phase II (49%) that information resources for applicants on the OER website are useful as provided.

Reviewers

- Most reviewers responded that appropriate review guidelines, criteria and instructions were provided in Phases II and III (no significant differences across phases) (**Figure 20**).
- Most reviewers rated the format and duration of discussions at study section meetings as sufficient for reviewers not assigned to an application to cast well-informed votes (**Figure 20**).
- Most reviewers strongly agreed/agreed that the time spent discussing the potential impact of applications was appropriate (**Figure 20**).
- Most reviewers agreed in Phases II and III that the scoring guidance provided was useful for assigning scores in advance of and during review group meetings (**Figure 20**).

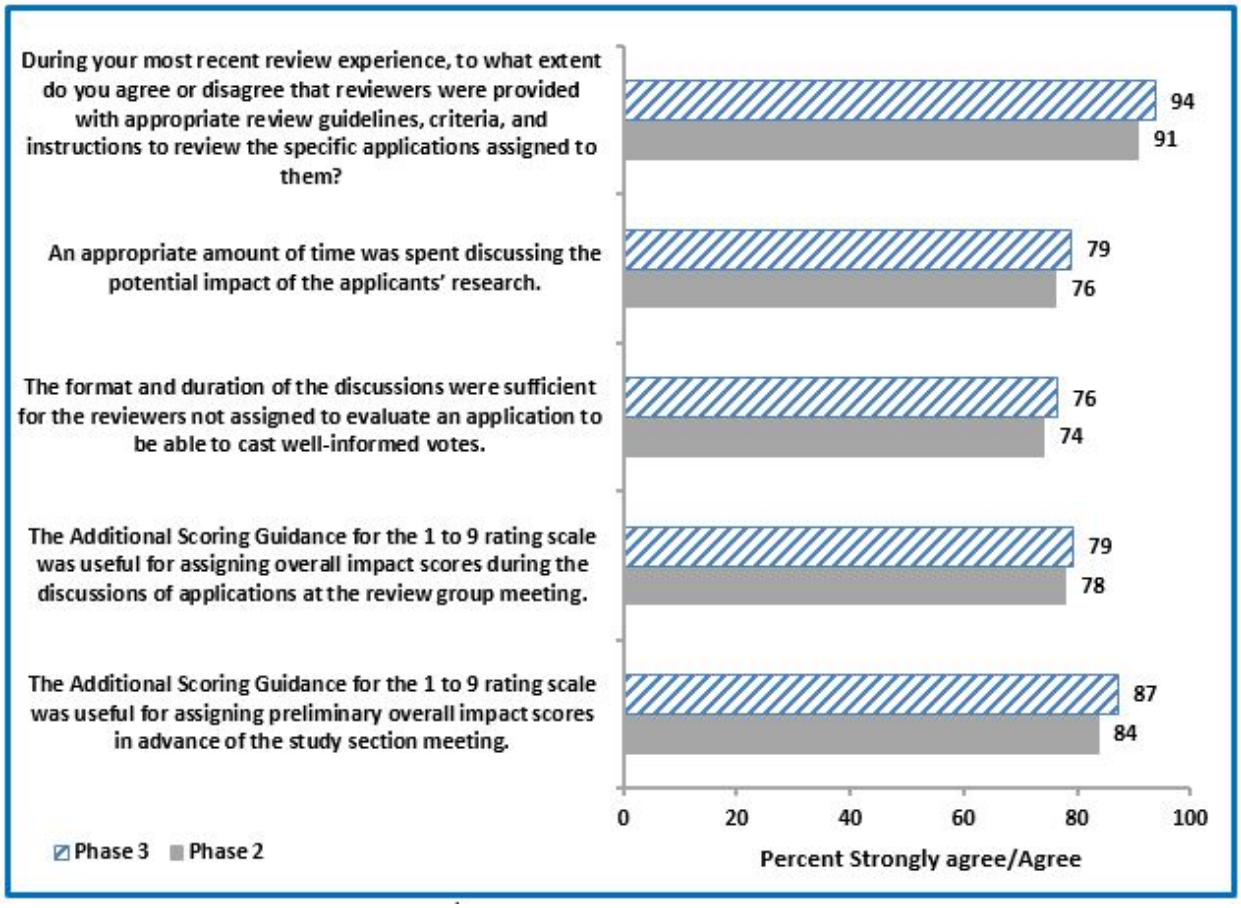


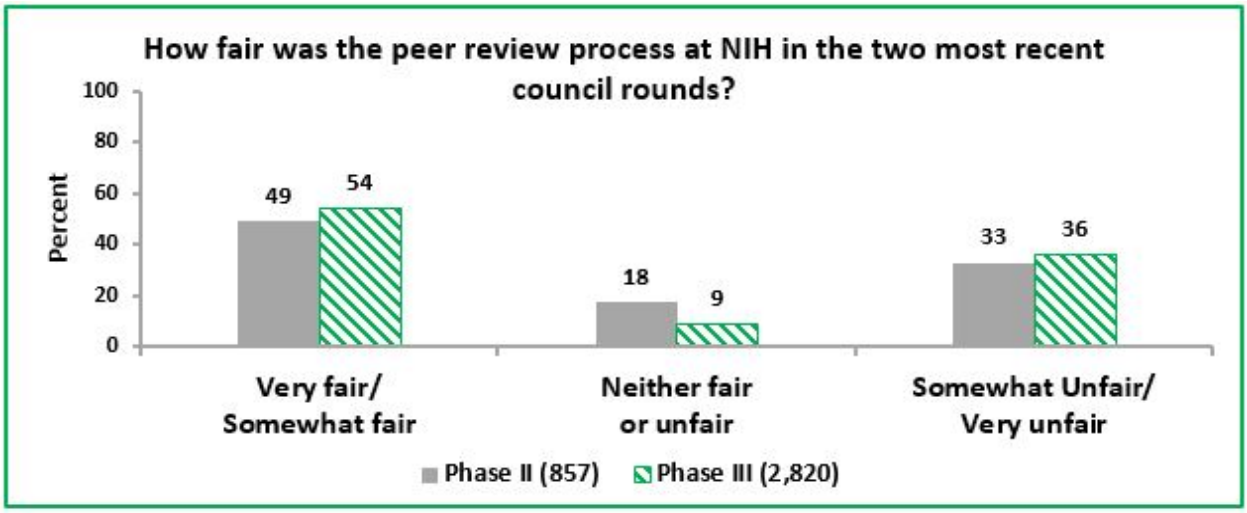
Figure 20: Reviewers' strongly agreed/agreed across Phases II and III that they were provided with adequate instructions, review guidelines and criteria to review the applications assigned to them. Most reviewers also agreed with statements about the appropriateness of the format, duration and composition of review group discussions. Finally, reviewers strongly agreed/agreed the table of scoring guidance provided in Phases II and III was useful for assigning scores before and during the meetings. There were no significant differences across Phase II and III in reviewers' responses to these questions.

Overall Satisfaction

All stakeholders were asked to respond to overall satisfaction questions:
1) During the past two rounds, how fair is the peer review process at NIH?
2) How satisfied are you with the peer review process at NIH during the past two rounds of review?

Applicants

- Applicants rated the peer review process at NIH as fair (54%) more often than they rated it as unfair (36%; **Figure 21, top panel**).
- Applicants were moderately more likely to rate themselves as satisfied (46%) than dissatisfied (44%) with the peer review process at NIH (**Figure 21, bottom panel**)



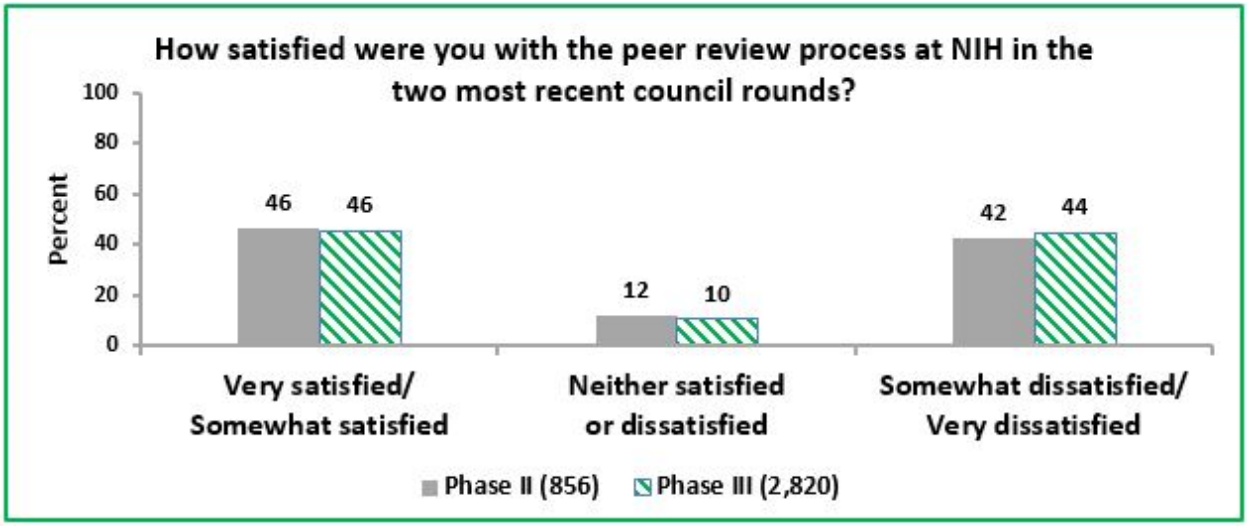
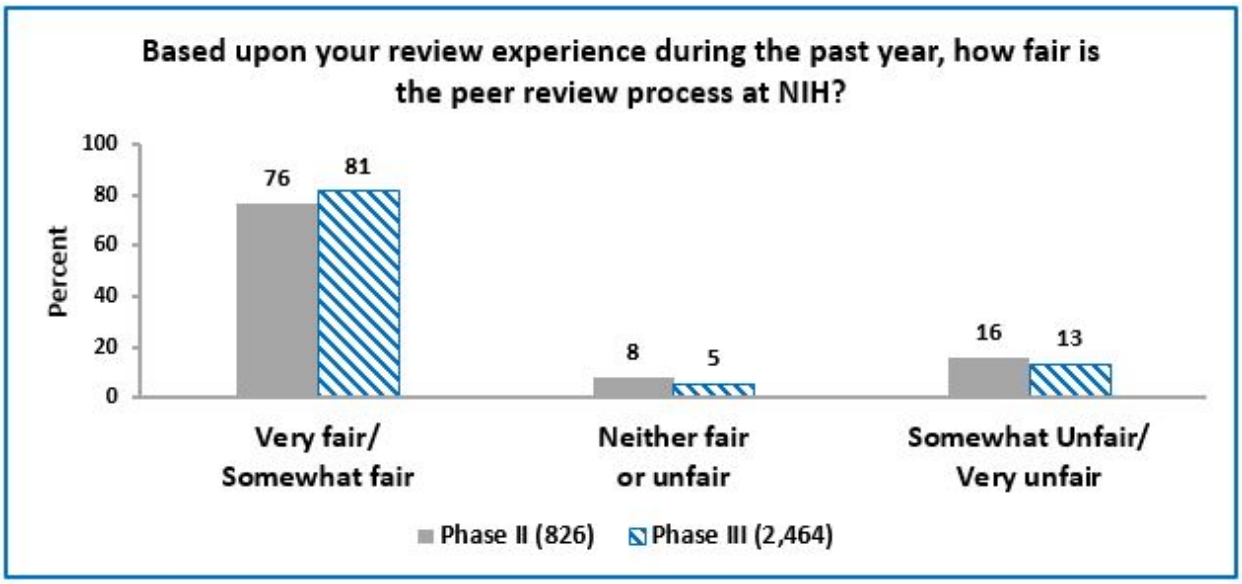


Figure 21. Applicants' responses to overall satisfaction questions. Most applicants rated the peer review process as very fair/somewhat fair in Phase III (upper panel), but fewer than half rated themselves as very satisfied/ somewhat satisfied (lower panel). There were no significant differences between Phases II and III.

Reviewers

- Most reviewers rated the NIH peer review process as fair in both phases, yet significantly more reviewers rated the system as fair in Phase III (81%) than in Phase II (76%; **Figure 22, top panel**)
- Most reviewers rated themselves as satisfied in both phases (72%; 77% for Phase II and III, respectively; **Figure 22, bottom panel**).



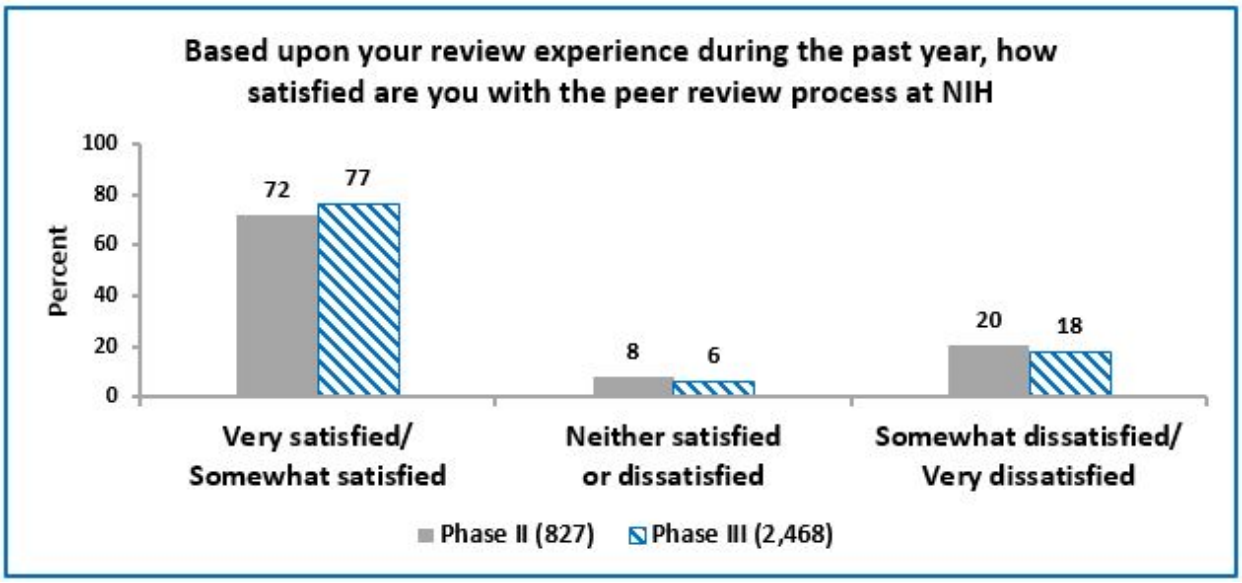


Figure 22: Reviewer responses to the two overall evaluation questions on the Phase II and Phase III surveys. Significantly more reviewers rated the peer review process as very fair/somewhat fair in Phase III than in Phase II (upper panel), and there was a modest increase in the percentage of reviewers who rated themselves as very satisfied/ somewhat satisfied (lower panel).

SROs

- Most SROs rated the peer review system as fair and rated themselves as satisfied in both phases (**Figure 23**)
- Significantly more SROs rated the system as fair in Phase III (96%) than in Phase II (79%; **Figure 23, top panel**)
- Significantly more SROs reported themselves as satisfied with the peer review system in Phase III (92%) than in Phase II (77%; **Figure 23, bottom panel**)
- Approximately a quarter (24%) of SROs' comments describing their satisfaction with the peer review process cited concerns regarding reviewer expertise.

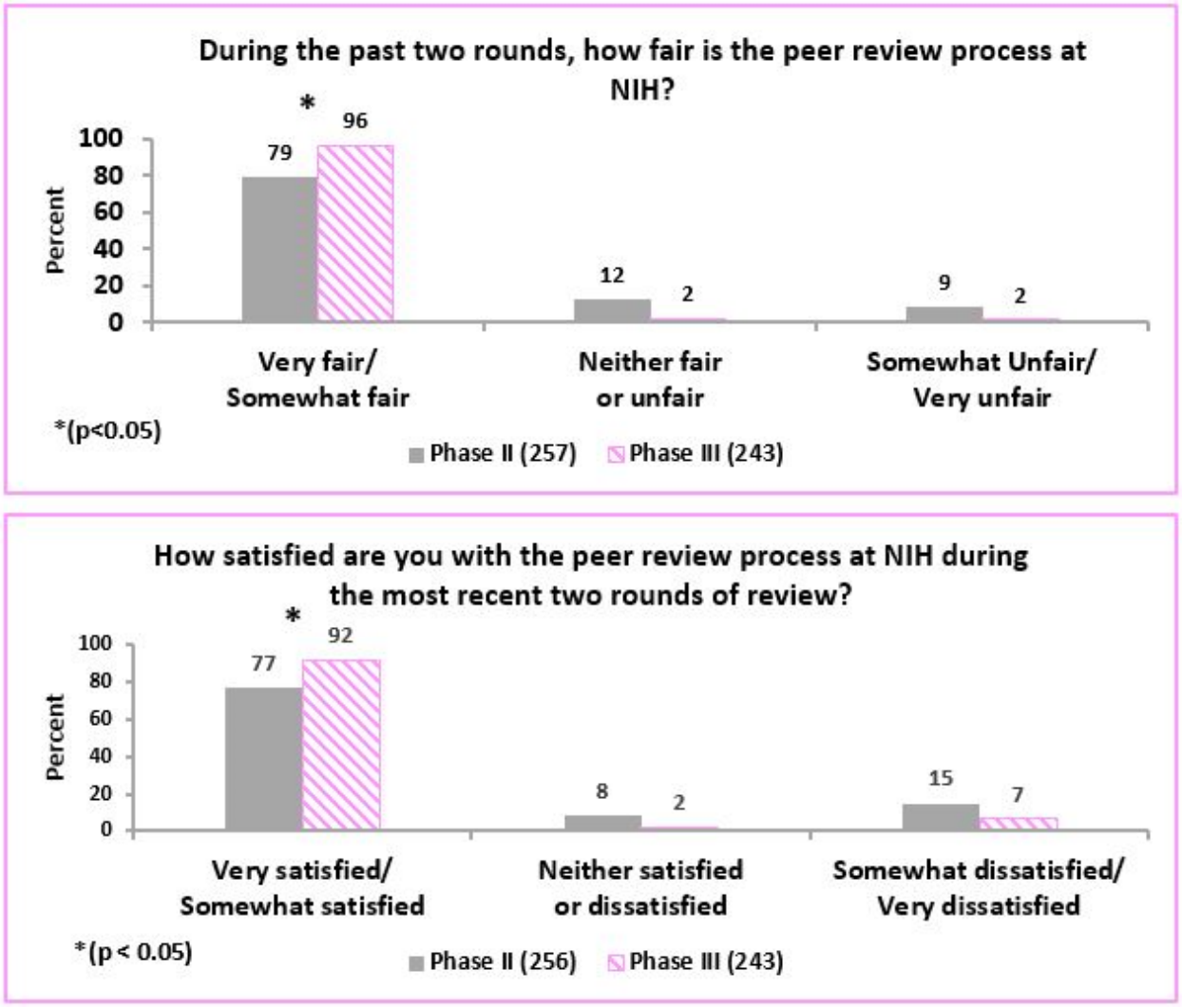
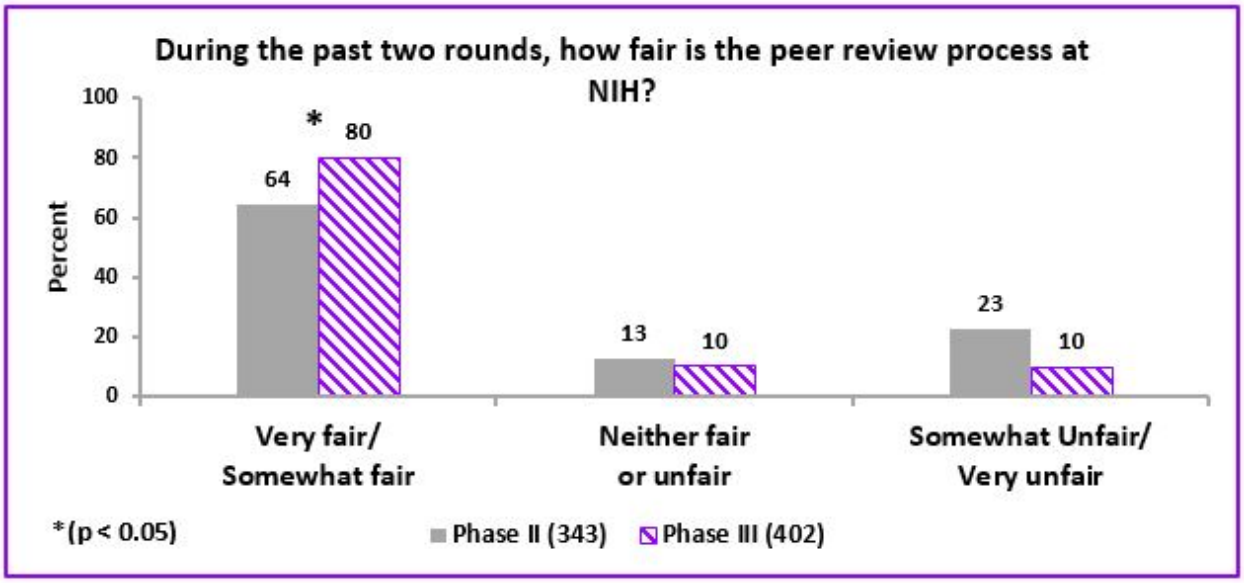


Figure 23: SRO responses to the overall evaluation questions on the Phase II and Phase III surveys. Significantly more SROs rated the peer review process as very fair/somewhat fair (upper panel) and rated themselves as very satisfied/somewhat satisfied (lower panel) in Phase III in comparison to Phase II.

POs

- Significantly more POs rated the system as fair in Phase III (80%) than in Phase II (64%; **Figure 24, top panel**)
- Significantly more POs reported themselves as satisfied with the peer review system in Phase III (73%) than in Phase II (51%; **Figure 24, bottom panel**)
- Under a third of POs' comments describing their satisfaction with the peer review process (29%) expressed concerns with reviewer expertise.



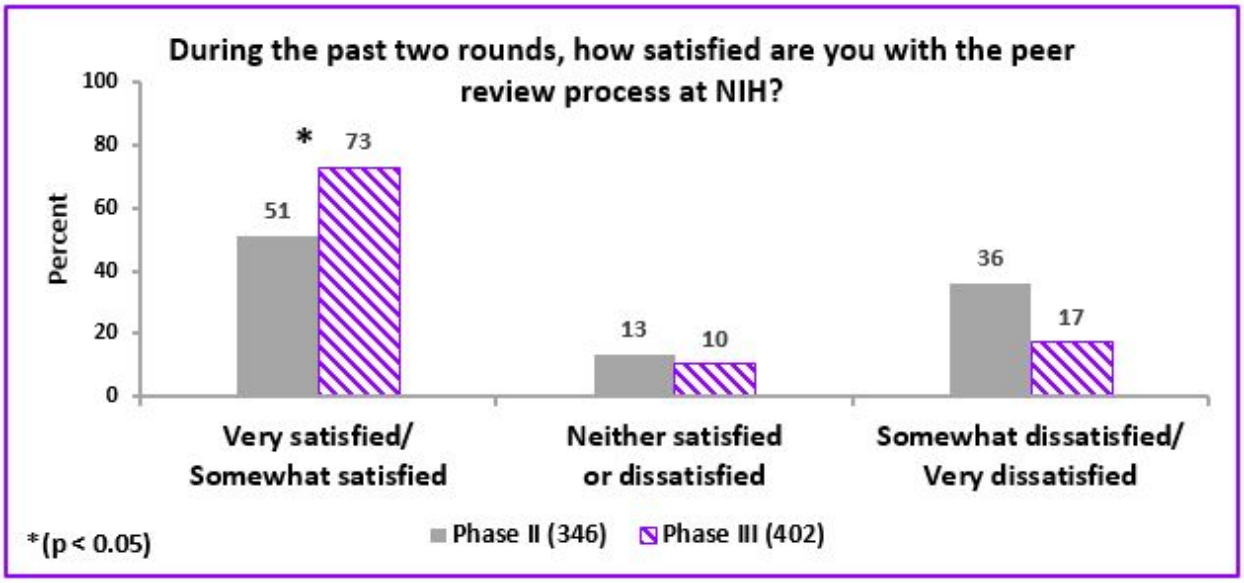


Figure 24: PO responses to the overall evaluation questions on the Phase II and Phase III surveys. Significantly more POs rated the peer review process as very fair/somewhat fair (upper panel) and rated themselves as very satisfied/ somewhat satisfied (lower panel) in Phase III than in Phase II.

Advisory Council Members

- Most advisory council members rated the peer review system as fair (79% and 81% for Phase II and III, respectively) and rated themselves as satisfied (72% and 77% for Phase II and III, respectively) in both phases (**Figure 25**).
- Significantly fewer advisory council members rated themselves dissatisfied with the peer review process in Phase III (12%) than in Phase II (21%)

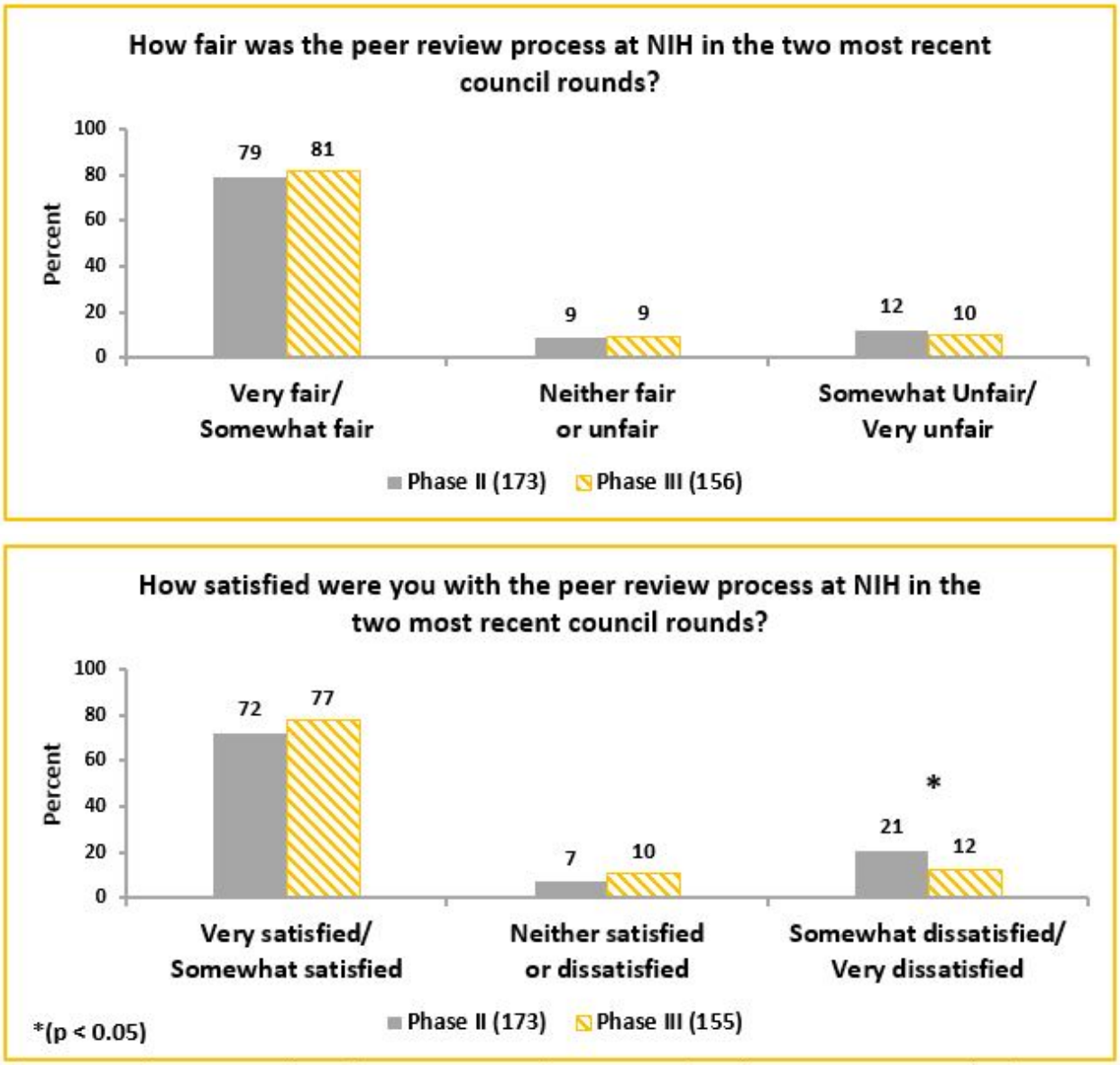


Figure 25: Advisory council member responses to the two overall evaluation questions on the Phase II and Phase III surveys. Most advisory council members rated the peer review process as very fair/somewhat fair (upper panel) and rated themselves as very satisfied/ somewhat satisfied (lower panel) in both phases.

Discussion

This report explores the continuing performance of the NIH peer review system, as indicated through Phase III survey responses from a variety of NIH stakeholders. Overall, the survey responses identified a number of changes implemented as part of Enhancing Peer Review that uphold the core values of NIH peer review. Of all the respondent groups, SROs responded most favorably to Phase III questions regarding the core values of peer review. In particular, most SROs responded favorably to questions about whether the NIH peer review process is fair and balanced.

Applicants expressed the least agreement that the current NIH peer review systems upholds the core values of NIH peer review. For example, when asked to rate whether reviewer expertise was appropriate during their most recent review or application experience, applicants were significantly less likely than reviewers to agree that reviewer expertise was appropriate; most reviewers agreed that reviewer expertise has been appropriate, while only half of applicants agreed.

Across all respondent groups, responses to questions about NIH peer review were largely more favorable in Phase III than in Phase II. However, there were instances in which Phase III survey responses were less favorable than that of Phase II. For example, SROs responded less favorably in Phase III than in Phase II to questions regarding peer review burden. In Phase III, fewer than half of SROs agreed that draft summary statements are easy to correct for problems with critique content and that assigned reviewers completed all required elements in the critique template. In phase III, over 60% of SROs agreed with these statements.

Roughly three-quarters of SROs, POs, reviewers, and advisory council members rated the NIH peer review process as fair and rated themselves satisfied with the peer review process at the NIH. In particular, SROs and POs responded favorably to questions about their overall satisfaction with the peer review process significantly more often in Phase III in comparison to Phase II. In contrast, over half of applicants rated the peer review process as fair and less than half rated themselves satisfied with the peer review process at NIH.

As discussed in the NIH-wide strategic plan, the NIH has established a framework to enhance its ability to achieve its mission while serving wisely as a steward of public resources. Some of the initiatives outlined in this plan may enhance NIH's ability to uphold the core values of peer review. The NIH stated that during FY 2016-2020, the agency would develop efforts to communicate its expectation that all NIH grantees serve on NIH peer-review study sections when asked, which could improve NIH's ability to involve reviewers with the appropriate expertise in the review of grant applications. Moreover, NIH described plans to centrally post information on IC award information on RePORT to provide the community with information on the use of the select pay option. Select pay refers to funds set aside to support grant applications that, based upon scores from peer review, do not fall within the payline, but that fill an important research gap and/or are of particular programmatic relevance to an ICO's scientific and health priorities. The continuous review of peer review, amid several policy developments such as these, strives to ensure that applications submitted to the NIH are reviewed under a peer review system that is fair, equitable, timely, and free of bias.

Appendix 1. Sampling of respondents and analysis of the Enhancing Peer Review Surveys

Identifying respondent populations

All SROs, POs and Advisory council members were identified from NIH's eRA database if they were assigned to administer applications in at least one of the two review or council rounds that occurred prior to the deployment of their respective surveys (June and October 2015 review rounds, or October 2015, January 2016 council rounds). Accordingly, 440 SROs, 1,147 POs and 250 Advisory Council members were invited to complete the surveys.

The applicant population comprises those individuals who submitted R01, R03, R21, U01 and R34 applications to NIH reviewed in any of the advisory councils/boards of NIH's constituent Institutes and Centers (ICs) in January or May 2015.These council rounds were chosen to permit sufficient time for applicants to be notified of their funding decisions and/or to decide whether to resubmit the application prior to completing the survey. A total of 28,647 eligible individuals were identified in the eRA Commons database as applicants and were included in the sampling frame.

The reviewer population was defined as individuals who served in NIH study sections that reviewed R01, R03, R21, U01 and R34 applications that were subsequently reviewed by the advisory councils/boards in January or May 2015. The target population of reviewers includes regular (appointed/permanent) and *ad hoc* (temporary) reviewers. A total of 13,987 individuals were identified in the eRA Commons database as reviewers and were included in the sampling frame.

Some individuals belonged to both the applicant population and the reviewer population. The sampling design was developed so that no one who was a member of both populations would be contacted for both the Applicant Survey and the Reviewer Survey.

A probability-based sampling design was created to ensure that statistical estimates would be unbiased as well as to ensure sufficient representation of various racial and ethnic groups. The number of individuals who could be contacted, not the number who were ultimately surveyed) was defined by burden limits under NIH's Office of Management and Budget (OMB) Generic Clearance No. 0925-0627. For the applicant and reviewer surveys, the total number of persons sampled under the burden limits established by the NIH guidance was 4,460. A sample of 2,322 applicants was invited to participate in the Applicant survey and a sample of 1,943 peer reviewers were invited to participate in the Reviewer survey. The 250 Advisory Council members were also counted in the OMB burden limit for this survey effort.

Survey Administration

Sampled Respondents were invited to participate in the surveys using the email addresses listed in the eRA database. Tracing was performed to identify correct email addresses for individuals whose invitations "bounced" on the first attempt. Each respondent was assigned their own electronic survey to complete, and they were sent periodic reminder emails until their survey was submitted. The surveys were available to SROs and POs in July and August of 2015. The surveys were available to applicants, and

reviewers in August, September, and October 2015. The surveys were available to advisory council members in November and December 2015.

Statistical Analysis

The following information collected from survey respondents was used to conduct statistical modeling of the survey responses to characterize underlying factors that may have contributed to survey responses:

SROs and POs:

- Number of years in their current position (as an SRO or PO)
- Number grants (POs) typically assigned to them
- SROs: Whether respondent works in an Institute or Center or in the Center for Scientific Review
- POs: How many study sections meetings they typically attend in a round by phone or in person

Applicants and Reviewers:

- Age
- Gender
- Race
- Ethnicity
- Type of Organization
- Job title
- Degree
- Applicants: New Investigator status
- Applicants: Whether the application proposed a clinical research project
- Applicants: Whether the application was a resubmission
- Applicants: Whether the application was assigned an Overall Impact score
- Applicants: Whether the application was funded
- Reviewers: Whether the reviewer has completed a term of chartered service