IDA | SCIENCE AND TECHNOLOGY POLICY INSTITUTE

April 12, 2022

MEMORANDUM

To: Ravi Basavappa, Becky Miller, and James Li
Office of Strategic Coordination and Center for Scientific Review
National Institutes of Health

From: Sally S. Tinkle, Xueying (Shirley) Han, Emma M. Thrift, Luke P. Newell, Gabriella G. Hazan, Nathan N. Dinh
Science and Technology Policy Institute (STPI)

Through: Kristen M. Kulinowski
Director, STPI

Subject: TRA Evaluation of the Anonymized Review of the FY2020 Transformative Research Award

---

The National Institutes of Health Office of Strategic Coordination (NIH/OSC) supports high-risk, high-reward research through targeted research programs, one of which is the NIH Director's Transformative Research Award (TRA). This award supports individuals or teams proposing transformative projects that are inherently risky and untested but have the potential to create or overturn fundamental scientific paradigms. For the 2021 TRA application cycle, the applicants were asked to anonymize their applications, and reviewers were not provided with the applicants' identity until the final phase of the review. NIH asked STPI to evaluate how anonymization has impacted the review process for reviewers, applicants, and NIH. The final report contains details of methods, results, conclusions, and recommendations. The report of these findings is attached to this memo.

IDA SCIENCE AND TECHNOLOGY POLICY INSTITUTE

Assessment of an NIH Anonymized Review Process

Sally S. Tinkle

Xueying “Shirley” Han

Emma M. Thrift

Luke P. Newell

Gabriella G. Hazan

Nathan N. L. Dinh

April 2022

Executive Summary

The National Institutes of Health (NIH) employs a competitive peer-review process that aims to identify the most promising research through a review process that is fair, independent, expert, and timely (NIH Center for Scientific Review 2021). A grant review process of this magnitude presents many challenges, and two advisory committees to the NIH Director made recommendations to enhance diversity of grantees, suggesting consideration of unconscious bias and an anonymized grant review. The NIH Office of Scientific Coordination which manages the Common Fund High Risk, High Reward (HRHR) research programs partnered with the Center for Scientific Review to develop an anonymized grant review process that would be piloted using the FY2021 HRHR Transformative Research Award (TRA) initiative.

The anonymized application required any identifiers for the Principal Investigator (PI), collaborators, laboratory, or institution be removed from the applications.[1] The anonymized review process had four components (Figure 1), and NIH constrained the anonymized information presented to the reviewers in Phase I (anonymized Specific Aims only), Phase II (anonymized Specific Aims and Research Strategy), and Phase IIIa (anonymized Specific Aims and Research Strategy, technical review). Phase IIIa reviewers produced preliminary scores for the applications. Phase IIIb reviewers had access to the complete, de-anonymized application for final scoring. Phase I and III reviews were conducted by an Editorial Board (EB) composed of senior scientists and Phase II technical reviews by subject matter experts.

Figure 1. Components of the Anonymized Review Process

NIH tasked the IDA Science and Technology Policy Institute (STPI) to evaluate whether anonymization impacted the review process for applicants, reviewers, and NIH administrative staff participating in the review, and whether anonymized review changed

[1] Section IV "names of individuals and institutions, honor and awards, hyperlinks, reference to any investigator attributes or accomplishments, citations that provide specific information about the source, and any other text from which the identity of any participating individual or institution can be reasonably inferred". See 2021 NIH Director's Transformative Research Awards FOA, available at https://grants.nih.gov/grants/guide/rfa-files/RFA-RM-20-013.html

the diversity of the applicants and awardees. Through consultation with NIH, STPI translated these overarching research questions into the following key study questions:

- Reviewers: Was anonymity maintained across the review process? Did reviewers have sufficient information to perform a rigorous review? Sufficient information to review the transformative potential of research proposal?
- Applicants: Did anonymization impact the decision to submit an application? Did anonymization impact the ability to prepare a competitive application? The ability to convey the transformative potential of the proposed research?
- NIH: Is the anonymized review process sustainable?

Critical to the initial concerns of the NIH and the advisory committees:

- Did anonymization change the diversity of applicants and awardees?[2]

To address these questions, a multi-modal study design was developed for data collection and analysis (Figure 2). Building out the initial review phases to correspond to the study questions, Pre-phase 1a was added to assess the applicant experience; and Pre-phase Ib and Pre-phase II, the NIH administrative staff experience. Surveys were conducted for each phase of review (red asterisks) and an analysis of demographic data was performed (narrow blue arrows). Although the survey populations vary from 25 to 176 individuals, the response rates for the six surveys performed in this evaluation are all

[2] "Diversity" includes demographic diversity (gender, race/ethnicity; See See 2021 NIH Director's Transformative Research Awards FOA, available at https://grants.nih.gov/grants/guide/rfa-files/RFA-RM-20-013.html) and institutional diversity.

greater than 40% and exceed 60% for the Pre-phase Ia and Phase II and 80% for Pre-phase 1b and Phase III.

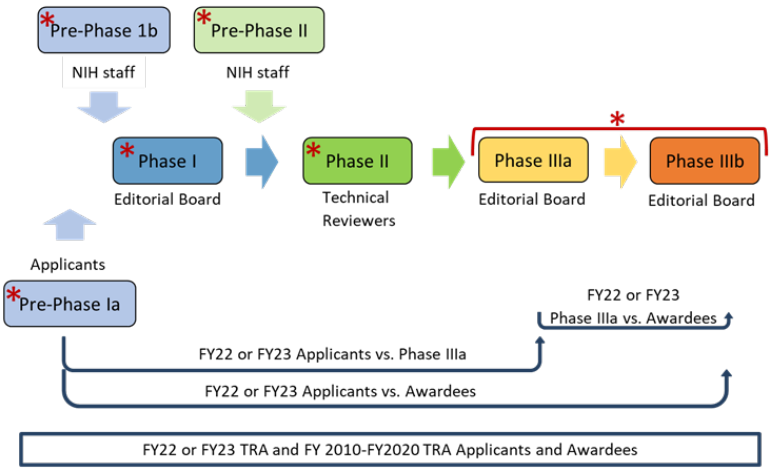


Figure 2. Anonymized Review Study Design

From these analyses, STPI developed a set of findings that address each study question.

Review

- **Was anonymity maintained across the review?**

No EB survey respondents reported they could identify an applicant; however, 20 (19%) technical review survey respondents reported identifying the applicant. STPI could determine that 17 of the technical review respondents flagged 15 unique applications among 54 applications by recognizing the work or noting information that was a personal identifier for a subject matter expert. These data indicate that with increasing information and subject matter expertise, anonymization is difficult to maintain at the technical review phase.

For a comprehensive accounting across all parts of the review, NIH provided data to STPI on their administrative review. Eleven of 176 applications were identified as noncompliant with anonymity instructions in Pre-phase I and 5 of 56 applications in Pre-phase II.

- **Did reviewers have sufficient information at each step of the process to perform a rigorous review? Assess transformative research potential?**

Most EB and technical reviewers considered the information available to them at each phase of the review sufficient to perform their review but expressed interest in having additional information (e.g., PI experience, resources, and supporting data) to further confirm their conclusions. Despite some respondents reporting an inability to assess transformative potential in Phases I and III, the majority of respondents selected *somewhat* or *very confident* to describe their ability to determine transformative potential.

Applicants

- **Did anonymization impact the decision to submit an application?**

While three-fourths of the applicant survey respondents reported that anonymized review was not a factor in their decision to submit an application, 25% reported they applied because anonymized review had the potential to be less biased; citing diversity factors and institutional prestige, among others; and the potential for more focus on science and not the scientist.

- **Did anonymization impact the ability to prepare a competitive application? Impact the ability to convey transformative research proposal?**

Almost all applicant survey respondents reported challenges in writing their applications, primarily, the omission of personal identifiers while trying to convey competitive, transformative research. They cited restrictions on the use of information that demonstrated expertise, feasibility, and unique capabilities and resources, and requested more examples of anonymized information (e.g., preliminary data, previous data, collaborators, unique technologies and methods).

NIH

- **Is the anonymized review process sustainable?**

The summary data in the previous sections show that the review process was completed as designed; however, concerns about anonymity at the technical review phase should be noted. These data also demonstrate the impact of anonymized review on the content and organization of the application and the sufficiency of information to conduct rigorous review.

Overall, EB respondents reported that anonymized review was *similar to* or *easier* than traditional NIH review, and technical and EB respondents expressed high levels of confidence in their reviews. Despite the high percentages of applicants and reviewers reporting sufficient information for their component of the process, reviewers requested additional information at all stages of the review. Most importantly, the majority of reviewers expressed interest in participating in future anonymized reviews.

Diversity

- **Did anonymization change the diversity of applicants and awardees?**

For FY2021, no significant differences were detected in the breakdown between applicants and awardees by gender, race, or ethnicity.

With respect to demographic diversity, STPI considered diversity to have increased for a demographic factor (i.e., gender, race, and ethnicity) in FY2021 if at least one non-dominant group within a demographic factor (e.g., female or Other for gender) increased in percentage relative to FY2010–2020 or the dominant group within a demographic factor

(i.e., males for gender, Whites for race, and Not Hispanic or Not Latino for ethnicity) decreased in percentage relative to FY2010–2020. Overall, applicant diversity increased for gender, race, and ethnicity in FY2021 relative to FY2010–2020. With respect to awardees, gender and ethnicity diversity did not increase in FY2021 relative to FY2010-2020 but diversity did increase for race.

With respect to geographic diversity, no new States were represented by TRA applicants in FY2021. With respect to institution diversity, the number of new institutions observed in FY2021 is comparable to what was observed in the previous 2 years. In addition, institutional diversity for FY2021 was within the range of what has been observed in previous years, suggesting that institutional diversity as measured by Simpson's Diversity Index did not increase in FY2021 relative to FY2010–2020. Among funded TRA applications in FY2021, two are from institutions that had never had a contact PI apply to the TRA program prior to FY2021.

Final Considerations

The goal of NIH review is to employ a competitive peer-review process that is fair, independent, expert, and timely to identify the most promising biomedical research. The goal of anonymized review is to increase the diversity of the NIH applicants and awardees by shielding from reviewers any information in the applications that would identify the PI, collaborators, laboratory, or institution.

As a first consideration, a competitive process should attract a robust number of applications. The FY 2021 TRA FOA received 176 applications, which is not significantly different from what has been observed in the previous four fiscal years (FY2016-FY2020). Given the novelty of the anonymized review process and its implications for drafting an application, a noncompliance withdrawal rate less than 10% is not unreasonable.

The EB and technical review phases of the anonymized process incorporated a second consideration: independent and expert review. Most EB and technical reviewers reported that they could conduct rigorous reviews of anonymized and constrained information at each step of the process and were confident of their efforts. Redundancy in the review process through the assignment of multiple reviewers to each application also provides evidence for rigorous review.

Anonymization was incorporated into this review to address the issue of fairness raised, in part, by publications demonstrating lack of diversity in the NIH awardee population. For FY2021, no significant differences were detected in the applicants and awardees for gender, race, or ethnicity, suggesting that an increase in applicant diversity could reasonably be expected to translate to the awardee population. STPI analysis identified more applicant diversity for gender, race, and ethnicity in FY2021 relative to

FY2010–2020. Neither geographic or institutional diversity, as measured in this study, increased in FY2021 or in FY2021 relative to FY2010–20.

This report documents an assessment of the first use of an anonymized review process for a Common Fund HRHR initiative. The data confirm that the anonymized process meets the NIH review criteria, and there is evidence for increased diversity for applicants. Because the number of awardees stratified by demographic, institutional, and geographic diversity is small, robust observations of diversity changes will require pooled data from multiple review cycles.

Contents

1. Introduction

The National Institutes of Health (NIH) aims to advance the health of the Nation by funding biomedical research (NIH 2020a). According to the website, NIH invests approximately $41.7B annually in biomedical research (NIH 2020a). Since 1946, NIH has employed a competitive peer-review process that aims to identify the most promising research through a review process that is fair, independent, expert, and timely. The process places an emphasis on scientific ideas, not applicants or institutions, and strives for transparency to the applicant by providing comments from the reviewers to the applicant (NIH Center for Scientific Review 2018). Review is conducted, for the most part, by the NIH Center for Scientific Review (CSR).

A grant review process of this magnitude presents many challenges, and numerous papers have examined racial and gender diversity among the NIH grant awardees and found disproportionate funding patterns (Mervis 2019; Reardon 2014; H.A. Valantine, Lund, and Gammie 2016; Wadman 2012). Although NIH leadership has been working to balance scientific workforce diversity for many years, most recently the NIH Director asked the Advisory Committee to the Director (ACD)[3] to examine the NIH review process and the demographic diversity of the applicants and awardees. In 2017 the ACD Working Group on Diversity (WGD)[4] developed 13 recommendations designed to accelerate diversity in the scientific workforce, including one to assess the effect of anonymization of peer review:

> NIH should design an experiment to determine the effects of anonymizing applications with respect to applicant identity as well as that of an applicant's institution. The WGDBRW understands that the nature of implicit bias cuts across processes, structures, organizations, and societal groups. The prospect of bias in the NIH peer review process is a serious matter that calls for deliberative action in a timely fashion (H. Valantine, Serrano, and The Advisory Committee to the Director Working Group on Diversity 2017).

In 2019, the ACD High-Risk, High-Reward Working Group (HRHR WG) analyzed the participation of women and other underrepresented groups in the applicant, finalist, and

---

[3] See NIH Advisory Committee to the Director, available at https://www.acd.od.nih.gov/

[4] The ACD Working Group on Diversity in the Biomedical Research Workforce was first convened in 2011. See NIH Scientific Workforce Diversity: About Us, available at https://diversity.nih.gov/about-us, and NIH Advisory Committee to the Director: ACD Working Group on Diversity in the Biomedical Research Workforce, available at https://acd.od.nih.gov/working-groups/dbr.html

awardee pools of HRHR programs to identify possible causes for their underrepresentation and make recommendations to enhance diversity (NIH Advisory Committee to the Director 2018). The ACD HRHR WG found low representation of females and underrepresented minorities in the High-Risk, High-Reward (HRHR) applicant pool, which was also reflected in the awardee pool. This working group recommended that NIH consider ways to mitigate the potential for unconscious bias in the grant review process. NIH published a Notice of NIH's Interest in Diversity in November 2019 that detailed NIH's definition of disadvantaged groups and reiterated NIH's commitment to "enhance the participation of individuals from groups that are underrepresented in the biomedical, clinical, behavioral and social sciences" (NIH 2019).

Consequent to these findings and actions, the NIH Office of Scientific Coordination (OSC), which manages the HRHR programs, partnered with CSR to develop an anonymized grant review process that would be piloted using the FY2021 NIH Director's Transformative Research Award (TRA) initiative. This initiative is a component of the NIH Common Fund's HRHR program that is focused on "projects that are inherently risky and untested but have the potential to create or overturn fundamental paradigms and may require very large budgets" (NIH 2021b). The purpose of anonymized review was to determine whether anonymization influenced the demographic and institutional diversity of the applicants and awardees.

NIH published the TRA Funding Opportunity Announcement (FOA) RFA-RM-20-013 (hereafter called the TRA FOA) in the NIH Guide on May 21, 2021, with a due date of September 30, 2020. The submission requirements and review criteria can be found in Appendix A and Appendix B (NIH 2020b). In summary, applicants were instructed to remove from their application identifying information such as *names of individuals and institutions, honor and awards, hyperlinks, reference to any investigator attributes or accomplishments, citations that provide specific information about the source, and any other text from which the identity of any participating individual or institution can be reasonably inferred*. The five traditional NIH review criteria were maintained (significance, investigator, innovation, approach, environment); however, technical review only assessed significance, innovation, approach; and investigator and environment were reviewed in the last step—Editorial Board (EB) discussion and final scoring. As an additional process change, the information provided to reviewers at each step of review was constrained and anonymized until the last at which time the EB reviewed the entire application.

a. Phases of the Anonymized Review Process

The NIH developed a three-phase anonymized review process in 2009 and adapted it to the anonymized review process (Figure 1). In Phase I, EB members review the anonymized Specific Aims excerpted from applications and assign each Specific Aims a

score. Each reviewer reviews about 40–50 Specific Aims and each Specific Aims is assigned to at least three reviewers. A subset of applications moves to Phase II and technical subject matter expert (SME) review. The SMEs conduct a technical review of the anonymized Specific Aims and Research Strategy using three NIH review criteria (significance, innovation, and approach) and provide written comments. Phase III has two parts: Phase IIIa and Phase IIIb. In Phase IIIa, EB members receive the technical reviewers' assessments and the anonymized Specific Aims and Research Strategy. From this information, they assign preliminary scores. In Phase IIIb the EB members discuss the complete application using all five NIH review criteria. This adds assessment of the investigator and environment to the previous three review criteria (significance, innovation, approach). Following discussion, the EB members assign a final score to each application.

NIH also developed an administrative review to assess anonymization of the Specific Aims prior to Phase I review and the Research Strategy prior to Phase II review. These reviews were conducted by NIH staff, and those Specific Aims and research strategies that did not follow the rules of anonymity detailed in the TRA FOA were administratively withdrawn from the review.

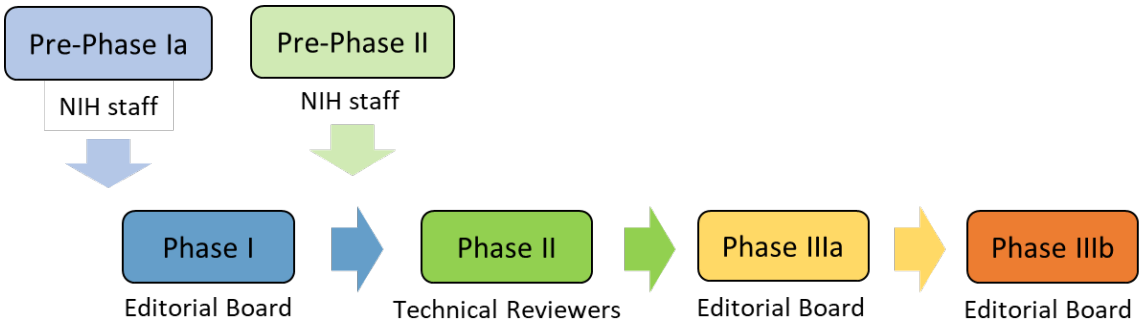


Figure 1. Phases of the Anonymized TRA Review Process

b. Scope of the Evaluation

Recognizing the complexity of an anonymized review process and the importance of an independent evaluation, NIH tasked the IDA Science and Technology Policy Institute (STPI) to assess each phase of the anonymized review process, evaluate whether the process is sustainable, and make recommendations to improve the process. NIH also tasked STPI to determine whether anonymization impacted the diversity of applicants and awardees. STPI understood demographic diversity to mean gender, race, and ethnicity, (NIH 2021a) and geographic diversity to mean institution and State. The results of this assessment were used by NIH to inform the Council of Councils at their May 2021 meeting and the FY2022 TRA FOA and application and review process.

The key question in the Statement of Objectives asks how anonymization impacted the review process. Through discussion with NIH, the key question was parsed to study questions for:

- Reviewers:
  - Was anonymity maintained across the review?
  - Did reviewers have sufficient information to perform a rigorous review? Review of transformative potential of research proposal?
- Applicants:
  - Did anonymization impact the decision to submit an application?
  - Did anonymization impact the ability to prepare a competitive application? A transformative research proposal?
- NIH
  - Are the mechanics of the review process sustainable?

To assess whether the anonymized review process will increase the diversity of TRA awardees, STPI will examine the demographic and geographic diversity of

- applications submitted and awards received
- changes in diversity across the phases of the FY2021 TRA review cycle
- TRA FY2021 applicants and awardees compared to TRA FY2010 to FY2020 cohorts

c. Study Design

To assess the multi-faceted questions outlined above, STPI developed a multi-modal study design that consists of surveys and statistical analyses of diversity data. Surveys will provide data on impact of anonymization on application development and each review step. Surveys will query the effectiveness and sustainability of the process (was anonymity maintained and was the information at each phase sufficient for the review step). The diversity analysis will evaluate changes across the FY2021 TRA review process, and compare the TRA applicants and awardees to a FY2010–2021 R01 cohort.

STPI revised the graphic of the review process (Figure 1) to include applicants (Phase Ib) and the demographic and geographic diversity analysis (Figure 2). The persons participating in each of the Pre-phase and Phase I–III components of the review were invited to participate in a survey and the applicant and awardee data were obtained from the NIH Query, View, Report (QVR) system. Detailed methods and content for the surveys

are found in Part I and surveys and data for the diversity analysis in Part II Diversity Analysis.

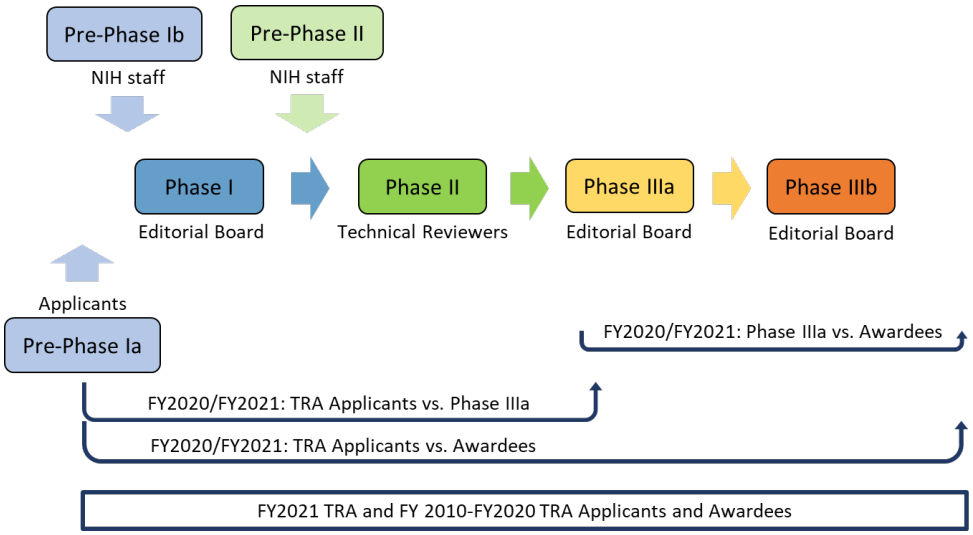


Figure 2. Study Design for FY2021 TRA Anonymized Review Assessment

2. Surveys

A. Methodology

This study assesses applicants who applied to the FY2021 TRA FOA: RFA-RM-20-013.[5]

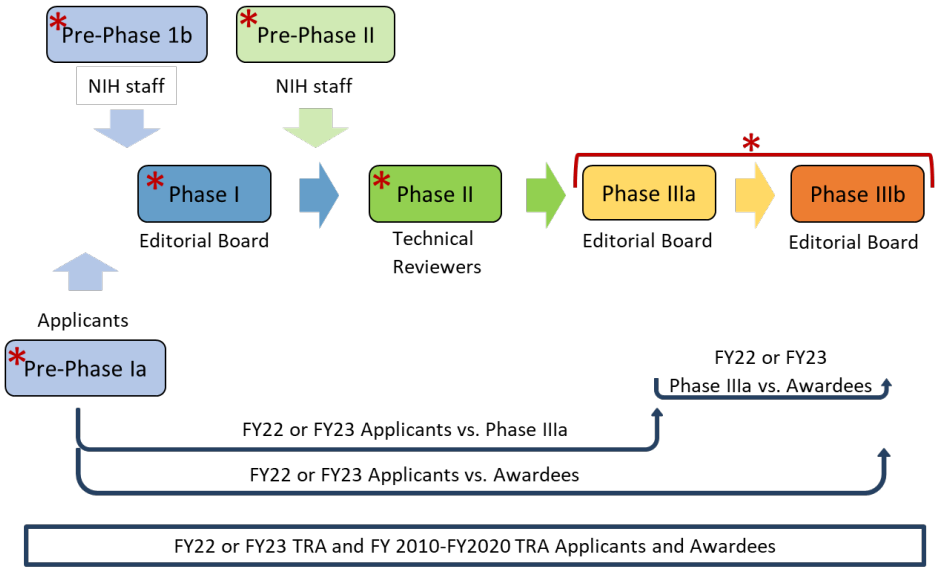
1. Survey Development and Administration

STPI designed six surveys to examine applicants' abilities to prepare an anonymized application, and reviewers' abilities to review the applications and interest in participating in future anonymized reviews. Each survey was iterated with NIH, and content and format were tested through STPI focus groups. The content for each survey is described in the introduction to each survey section below, and the complete survey is provided in Appendices D, F, H, J, L, and N.

STPI developed and administered the surveys in Alchemer, a web-based survey platform.[6] Surveys were administered to four groups: TRA applicants, administrative reviewers, EB members, and technical reviewers (Figure 3). Each applicant and reviewer received either a personalized or general link to the survey through email. Administrative reviewers and EB members participated in two phases of the review and received separate surveys for each phase. EB reviewers received one survey for Phases IIIa and IIIb. Technical reviewers reviewed one or two applicants and were asked to complete a survey for each applicant. Responses were kept confidential, and only aggregate results are provided to NIH.

[5] Applications responsive to the Notice of Special Interest (NOSI) Common Fund TRA FOA for ALS-related research: NOT-RM-20-019, or the Emergency TRA FOA for SARS-CoV-2-related research: RFA-RM-20-020 were not included in this analysis.

[6] See Alchemer web page, available at https://app.alchemer.com/



Surveys are designated by red asterisk and diversity analyses by blue arrows at the bottom of the figure

Figure 3. Six Surveys Developed for the Anonymized Review Process

2. Survey Analysis

Each survey included yes or no, select all, numeric, and free response questions. All analyses were performed in R.[7] Some yes or no questions were accompanied by free response questions allowing the applicant to expand on their previous answer. Free response questions were analyzed using inductive coding. Generally, when more than four applicants responded with a similarly coded answer, these answers were considered their own category. The term *respondents* is used to delineate analyses and discussions that pertain solely to those who completed the survey. Categories with four or fewer respondents were grouped into a category called *other*. For free response questions, the respondents within each group vary by response. For example, if 10 respondents are grouped for *other* response A and 10 respondents for *other* response B, the 10 respondents in each group are not the same.

Only completed surveys were analyzed. Any response for which the respondent reached the end of the survey was considered complete. Because survey questions were optional, the number of responses to each survey question varied. Descriptive statistics (i.e., the number of responses and the percentage of survey respondents who selected each answer choice) are provided for each question. Qualitative information was analyzed using

[7] See 2020 R: A language and environment for statistical computing, available at https://www.R-project.org/

coded qualitative categories. Free responses that did not answer the question asked were eliminated from the analysis.

Exact binomial tests were used to assess whether the percentage of respondents who selected *yes* is significantly different from the percentage of respondents who selected *no* for two answer multiple choice survey questions. For *select all that apply* survey questions, a Cochran's Q-test was used to determine whether there was a significant difference in the percentage of respondents who selected each answer choice. In cases where Cochran's Q-test was significant, pairwise comparisons between all answer choices were performed to determine how the percentage of respondents who selected each answer choice differed from one another. All statistically significant findings in this analysis are significant at $p < 0.05$.

B. Applicant Survey

The anonymized TRA award process begins with the preparation of an anonymized application (Figure 4). The applicant survey queried the applicants on their experience preparing and submitting an anonymized application according to TRA FOA instructions (NIH 2020b). The survey questions are located in Appendix C and the data analysis in Appendix D.

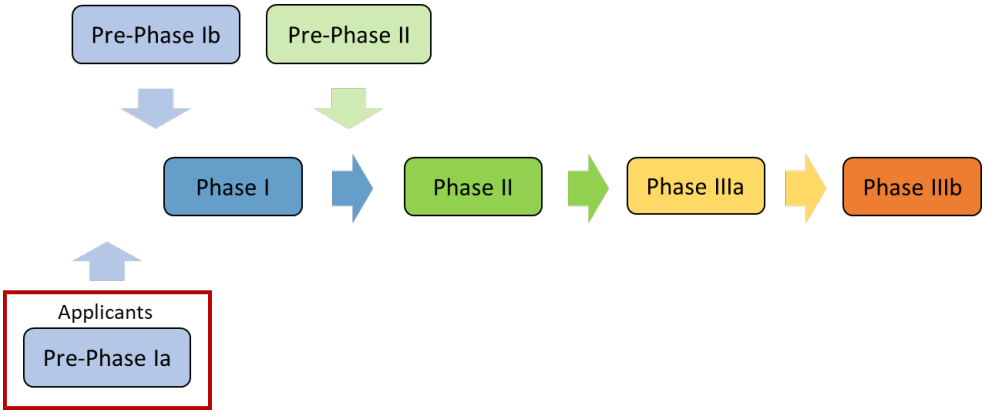


Figure 4. Pre-phase 1a Focuses on the Applicant

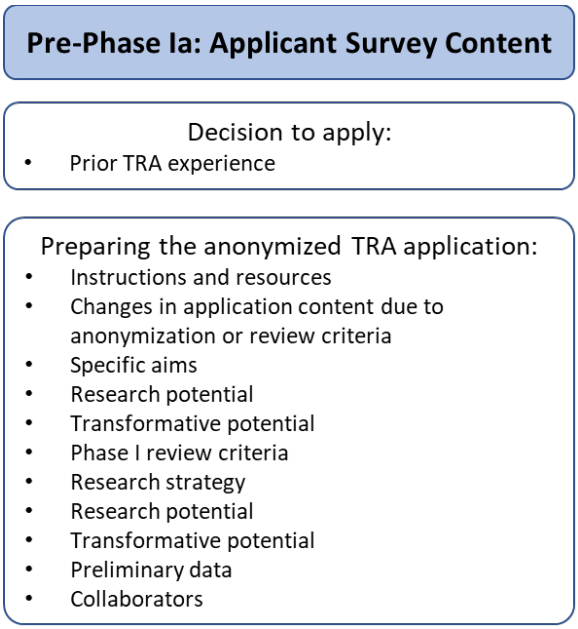


Figure 5. Applicant Survey Content

1. Survey Administration

An email containing personalized links was sent on October 14, 2020 to the 176 PIs who applied to TRA FOA. The first reminder was sent on October 21 and the final reminder on October 26 for a survey close date of November 4. In an effort to survey all those relevant in the application process, the same survey was sent to 119 non-contact PIs[8] applicants 1 week later on October 21 with a reminder sent on November 2 for a survey close date of November 4. The survey data were downloaded for analysis on November 12. Two additional responses were received after the applicant data were downloaded; they were not included in the analysis.

2. Results

a. Response Rate

One hundred and five of 176 applicants listed as the PI (62%) and 22 of the 119 co-PI applicants (18%) completed the survey.[9] STPI received responses from both the contact PI and 1 co-PI for 11 applications and from the contact PI and 2 co-PIs for 1 application. Nine co-PIs responded to a survey for an application that had no response from a contact

---

[8] The TRA application can have multiple PIs. For the purposes of this evaluation, STPI refers to the *contact PI* for the application and all other PIs as *non-contact PIs*.

[9] There were twenty-one partial (i.e., incomplete) survey responses from TRA applicants. These twenty-one partial survey responses were not included in any of the survey analyses.

PI. Due to the low number of responses from non-contact PIs and the potential for multiple survey responses for a single application, non-contact PI responses were excluded from all analyses.

b. Survey Data

Questions were optional so the total number of respondents for each question may vary. Responses for each survey question are provided in Appendix D and followed by the statistical analysis in table format.

Question: Please indicate if use of an anonymized review impacted your decision to submit your application.

Seventy-nine of the 105 respondents (75%) reported that the use of the anonymized review process had no effect on their decision to submit their applications ($p < 0.001$). STPI performed a secondary analysis of these data and determined that the percentage of applicants reporting that anonymized review impacted their decision to submit an application did not differ significantly by gender, race, or ethnicity.

Twenty-six of the 105 respondents (25%) indicated that anonymized review impacted their decision to apply, and 24 described how the review impacted their decision.[10] Seventeen of 24 respondents noted that they were encouraged to apply because the anonymized process could be less biased; more focused on the proposed science, not the scientist; and more equitable (Figure 6). Responses falling outside of these categories cited awards being made based on reputation of the investigator; anonymized review being better suited for interdisciplinary work, and less conflict of interest.

In addition to these comments, respondents mentioned:

- Size, location, or perceived prestige of their institution (7 respondents)
- Younger applicants with less experience or name recognition (6 respondents)
- Gender, ethnic and racial bias; and bias against those in non-academic fields

[10] Three responses were not included in the analysis because the respondent did not answer the question asked.

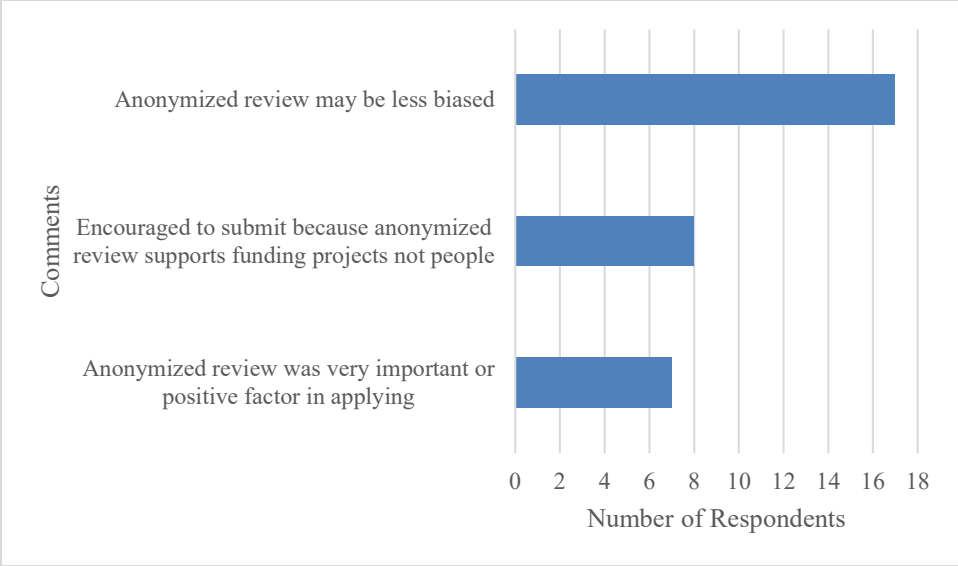


Figure 6. Impact of Anonymized Review on the Respondents' Decision to Submit

Question: Please indicate if the FOA instructions you received from NIH were sufficient to prepare your application.

The TRA FOA provides the primary set of instructions for applicants. Eighty-five of 104 survey respondents (81.7%) said the FOA instructions were sufficient for them to prepare their application ($p < 0.001$). Respondents who were administratively withdrawn at Pre-phase I or Pre-phase II were less likely than those not administratively withdrawn to report that the FOA instructions were sufficient.

Sixty-four of the 85 respondents (75%) described the TRA FOA elements that were sufficient to prepare an application (Figure 7). Twenty-six of the 64 respondents (40%) said the instructions and guidance on the format of the application were clear, detailed, and helpful.

Seventeen respondents indicated that the TRA FOAs guidance helped them to anonymize personal or laboratory identity to maintain anonymity. Respondents specifically noted:

- Examples of prohibited and accepted information (10 respondents)
- Specific formatting and anonymization instructions (8 respondents)
- TRA FOA reviewer requirements (6 respondents)
- Webinar provided useful material supplementary to the TRA FOA instructions (5 respondents)
- Guidance given to them regarding their personal references, and references to previous work was helpful (5 respondents)

Ten respondents provided responses that summed to less than five similar responses per topic. Placed in the *other* category, these responses included comments about the TRA FOA Section IV. Application and Submission Information and Section V. Application Review Information, preliminary data instructions, the use of multiple warnings within the instructions regarding compromising texts, and other technical aspects of the TRA FOA. Other responses cited sample grants, external links, and guidance on how to reference institutions.

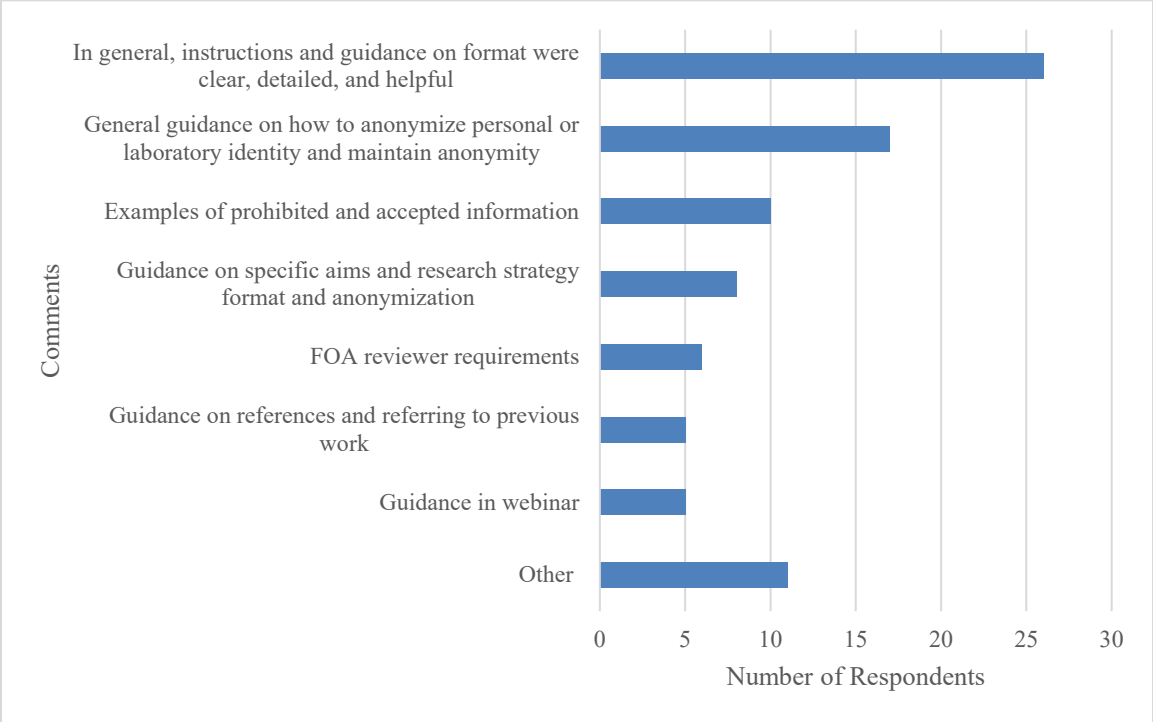


Figure 7. Elements in the TRA FOA Helpful to Developing an Anonymized Application

Seventeen of 19 respondents (89%), who indicated that the TRA FOA instructions were not sufficient to prepare an application, added feedback. Of these seventeen, fifteen respondents mentioned some aspect of the anonymization process they would like clarified:

- clearer instructions on how to anonymize
  - institution-specific resources and technologies
  - letters of support, or if letters of support were allowed
  - figures from published work
  - preliminary data
  - location of collaborators
- additional examples

  - anonymized grants
  - accepted or prohibited language
- summary of pros and cons for including preliminary data

Five respondents provided responses that summed to less than five similar responses per topic. Placed in the *other* category, these responses cited more clarity in TRA FOA instructions, whether Specific Aims were necessary, instructions on how to submit an application involving multiple institutions, how likely they were to receive the funding amount they requested, how to use page numbers, the necessary parts of the application and their components, instructions on page limits, details about the review panel expertise and structure; and better communication mechanisms with the program officers.

Question: What additional information would have been useful to help you prepare an anonymized application?

NIH hosted a Q&A webinar on June 29, 2020 to answer applicants' questions regarding the TRA FOA.[11] At the webinar, and in general, applicants were encouraged to reach out directly to NIH program officers with questions about the application process if needed. Thirty-nine of 103 (37.9%) respondents reported consulting NIH program officers, 54 (52.4%) reported using the webinar, and 27 (26.2%) reported referring to other NIH resources for help when preparing their applications (*Cochran's* $Q = 11.68, p = 0.003$; Figure 8).

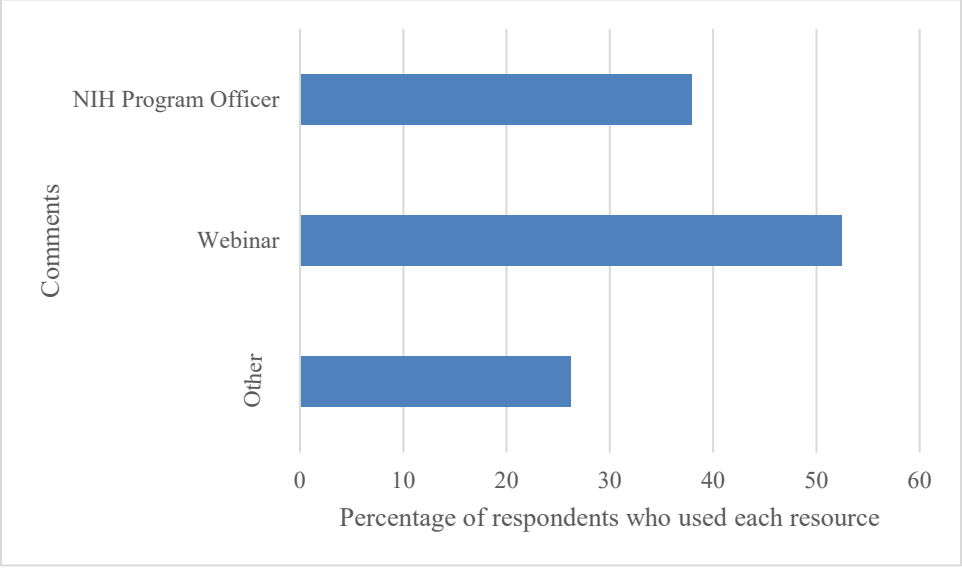


Figure 8. Other Resources Consulted by Applicants

[11] Webinar is available at: https://www.youtube.com/watch?v=JzNJubfp2_8

Fifty-six of the 99 survey respondents (56.6%) used an additional resource to prepare their application, selecting among the survey options: webinar, NIH program officer, and other. Twenty-one of these 56 respondents reported receiving "general assistance," 14 respondents received helpful information on application preparation and the submission or review process. Eight respondents used the NIH TRA website, and fewer than 5 respondents indicated they used proposal examples, consulted an NIH PI not involved with the TRA program, and received assistance from collaborators.

Response categories that summed to less than five similar responses were grouped as *other* and identified the webinar and Q&A session as clarifying the rationale and scope of the anonymized review process and providing guidance on application preparation and anonymization. Specifically, they noted guidance on how to emphasize transformative potential while remaining anonymous; applicability of their research topic for the TRA award; budget, deadlines, and timelines; institution-specific resources and technologies; and letters of support. Survey respondents also referenced other awards that used similar mechanisms as guidance—and guidance from other NIH PIs and program officers. Additionally, they reviewed a task force report to gain insight on the impetus of the anonymized review, and received reassurance from NIH program officers that their background would be acceptable for them to apply and be competitive within the TRA process.

Question: Specific Aims: Anonymization - Please indicate if you made changes in your Specific Aims to comply with the anonymization instructions.

The TRA FOA contained a list of specific identifying information that applicants should not include in their Specific Aims and Research Strategy sections of their applications (NIH 2020b). Ninety-three of 105 (88.6%) respondents reported making changes to their Specific Aims in order to comply with the anonymization instructions ($p < 0.001$). The percentage of respondents who reported making changes did not differ whether or not respondents had previously submitted a TRA application.

Eighty-five of the 93 (91%) survey respondents described these changes (Figure 9). To comply with anonymization instructions, the respondents reported they:

- excluded, or limited, references to their (as PIs) or their lab's previous publications, data, and findings, [12] or preliminary data[13] (55 respondents)
- removed all language that could possibly lead to the identification of the PI(s), collaborators, institutions, or communities involved with the research, and anonymized their Specific Aims in general (21 respondents)

[12] Including citing other publications instead of just one's own.

[13] Including removing descriptions of "we have shown."

- limited significantly or did not mention the PI, laboratory, or research team's expertise or accomplishments (16 respondents)
  - excluded information that demonstrated the ability to successfully complete the proposed research
  - limited mention of unique collaborations necessary to complete the proposed research
- interpreted the TRA FOA instructions as excluding Specific Aims (7 respondents)
- excluded or modified the details about methods, experiments, or technologies used in their proposed research (5 respondents)

Response categories that summed to less than five similar responses were grouped as *other*. Within the *other* category, seven respondents excluded reference to other researchers who are doing similar research, emphasized the transformative potential and scientific impact of their research, were concerned about publishing results from their research during review, wrote their Specific Aims for a more general audience, used only numerical

references, and made the Specific Aims more comprehensive since it was the main component in the first review.

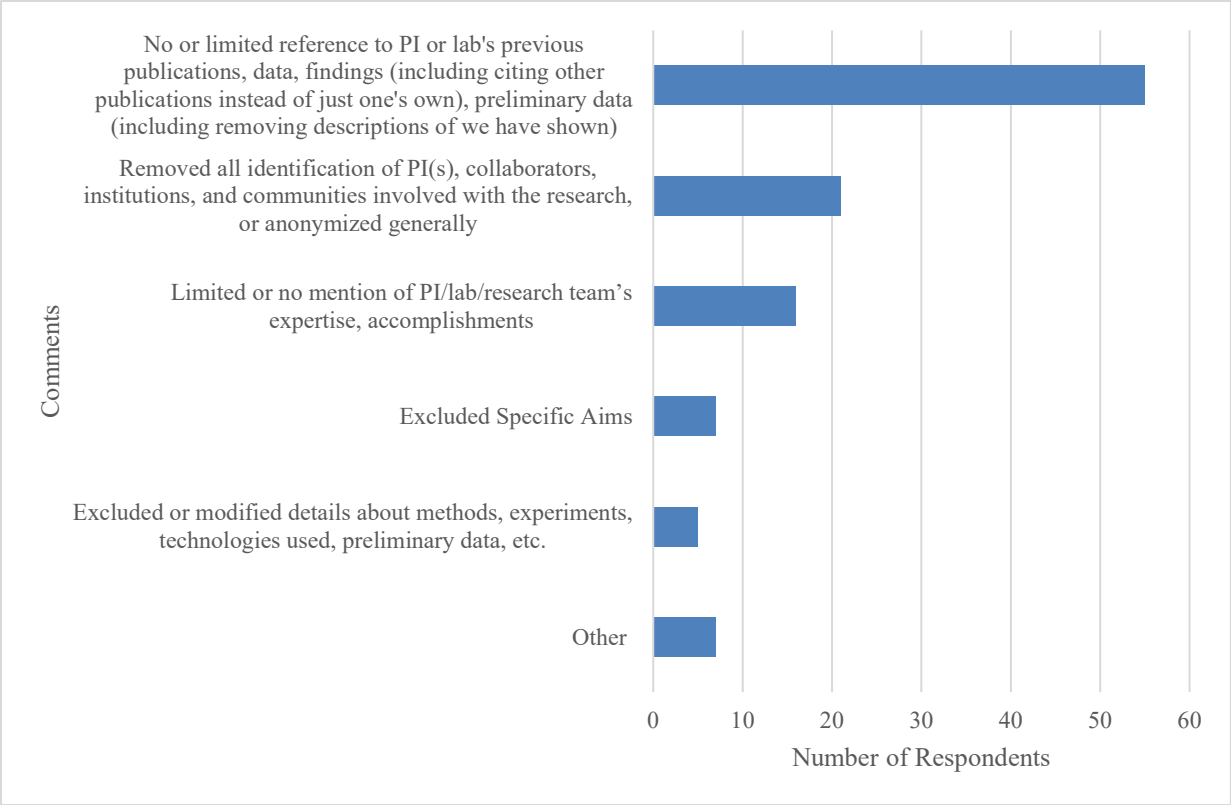


Figure 9. Changes Applicants Made to Their Specific Aims to Comply with the Anonymization Instructions

Question: Specific Aims: Transformative Potential - Please describe how the anonymization instructions affected your ability to convey the transformative potential of your research in the Specific Aims.

The TRA initiative is unique in its request that applicants specifically discuss the transformative potential of their proposed research. Thirty-two of 105 (30.5%) respondents reported that the anonymization instructions impacted their ability to convey the transformative potential in their Specific Aims ($p < 0.001$). Secondary analysis determined that the percentage of respondents who reported this affect did not differ significantly by race, gender, or ethnicity.

Twenty-nine of the 32 (91%) respondents who reported that they had challenges in ability to convey transformative potential described these factors. Twenty of the 29 respondents reporting this challenge noted that anonymization did not permit them to reference past work, publications, accomplishments, and expertise. Response categories that summed to less than five similar responses were grouped as *other*. Within the *other*

category, respondents challenged to describe transformative potential reported positive and negative effects of anonymization. Specifically, they reported that anonymization prevented them from referencing technological discoveries important to the research they proposed and referencing unique collaborations. In contrast, respondents reported an emphasis on the transformative potential of their research that enabled them to more easily convey the transformative nature of their research and focus on the potential overall impact of the proposed research on science.

Question: Specific Aims: Review - Please describe the changes you made to the Specific Aims knowing that only the Specific Aims were used in step one of the review.

The Specific Aims were the only information considered by the EB in Phase I of the review process. Half (52 of 104; 50%) of respondents indicated that this information regarding Phase I review influenced how they wrote their Specific Aims ($p = 1$).

Of the 52 respondents who indicated that knowing the Specific Aims was the sole component of the application reviewed during Phase I EB review, 45 described changes they made when writing their Specific Aims (Figure 10). Twenty of the 45 respondents reported that they emphasized the transformative potential of their research, and 13 respondents made the Specific Aims more compelling and comprehensive. In contrast, eight of the 45 respondents generalized the Specific Aims and made them *higher level* so they would be well received by scientists from a broad range of disciplines and backgrounds.

Response categories that summed to less than five similar responses were grouped as *other*. Eleven respondents reported emphasizing the rationale for their proposed research, making the Specific Aims more concise and less focused on experimental details, or completely excluding conventional Specific Aims. Some respondents bolded key points within their research, emphasized the feasibility of their research, or removed figures from their Specific Aims.

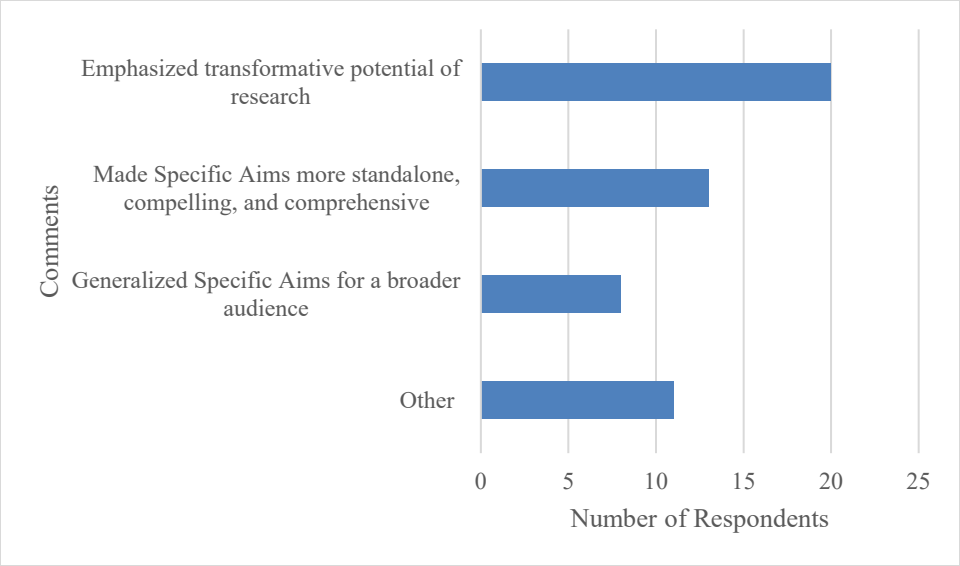


Figure 10. Changes the Applicant Made to the Specific Aims Because Only the Specific Aims Were Used in Phase I Editorial Board Review

Question: Research Strategy: Anonymization - Please indicate if you made changes in your Research Strategy to comply with the anonymization instructions.

A majority of respondents (96 of 105; 91.4%) reported that they made changes to their Research Strategy section to comply with the anonymization directions ($p < 0.001$).. The percentage of respondents who reported making changes to their Research Strategy to comply with the anonymization instructions did not differ based on whether applicants had previously submitted a TRA application.

Seventy-five of 96 (78%) survey respondents described these changes (Figure 11). The respondents:

- completely excluded, or limited references to the PI or lab's previous publications, preliminary data, or findings (59 respondents)
- removed all language that could be used to identify the PI(s), collaborators, institutions, and communities involved in the research. (29 respondents)
- limited or no mention of PI/lab/research team's expertise, accomplishments (26 respondents)
- omitted details about methods, experimental protocols, or technologies used (8 respondents)

Response categories that summed to less than five similar responses were grouped as *other*. Within this *other* category, four respondents reported they avoided using the names of their products or equipment to ensure their identity would not be revealed, used future tense within the Research Strategy and did not refer to any past activities, and included

disclaimers in relevant areas of the Research Strategy to ensure reviewers understood that any *product* name included did not imply an association with the research team.

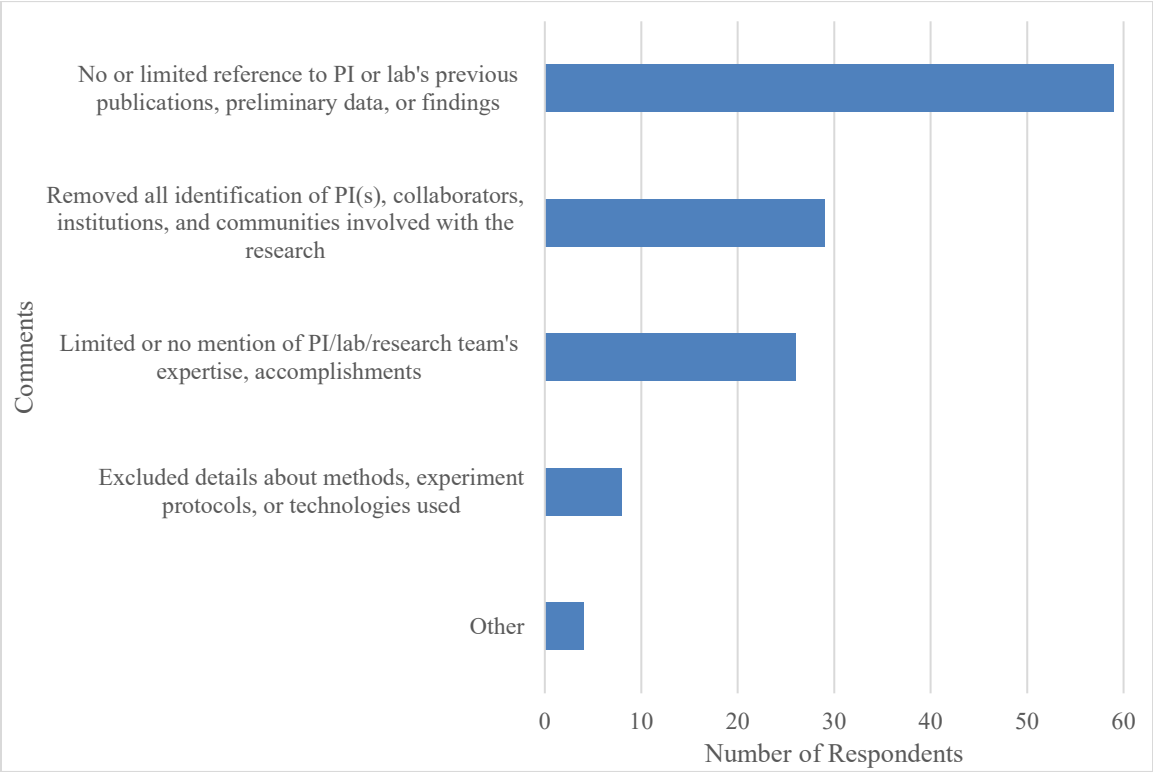


Figure 11. Changes Made to the Research Strategy to Comply with the Anonymization Instructions

Question: Research Strategy: Transformative Potential - Please indicate if the anonymization instructions affected your ability to convey the transformative potential of your research in the Research Strategy.

The TRA is unique in its request that applicants specifically discuss how their proposed research has transformative potential. Thirty-nine of 104 (37.5%) survey respondents reported an effect on their Research Strategy ($p = 0.014$). Secondary analysis demonstrated that the percentage of applicants who reported this effect for their Research Strategy did not differ significantly by gender, race, or ethnicity.

Thirty-four of the 39 respondents (87%) that indicated anonymization instructions affected their ability to convey the transformative potential of their research in the Research Strategy described those changes (Figure 12). Thirty-one respondents said that it was difficult to show transformative potential and feasibility without being able to show that the research team holds specialized expertise, has previous accomplishments, or has unique methods and knowledge to allow them to successfully complete the proposed research.

Response categories that summed to less than five similar responses were grouped as *other*. Within this *other* category, 11 respondents provided positive and negative comments. Respondents reported more difficulty conveying transformative potential or paradigm shift, especially without previous work history, and that anonymization is a heavier burden for accomplished researchers. Conversely, others reported that anonymization requirements necessitated more focus on the transformative nature of their proposed research and suggested that, because transformative ideas can be widely sourced, review of all high impact awards should be anonymized. Other respondents included previous work regardless of anonymization instructions, because they perceived it to be a strength in the application.

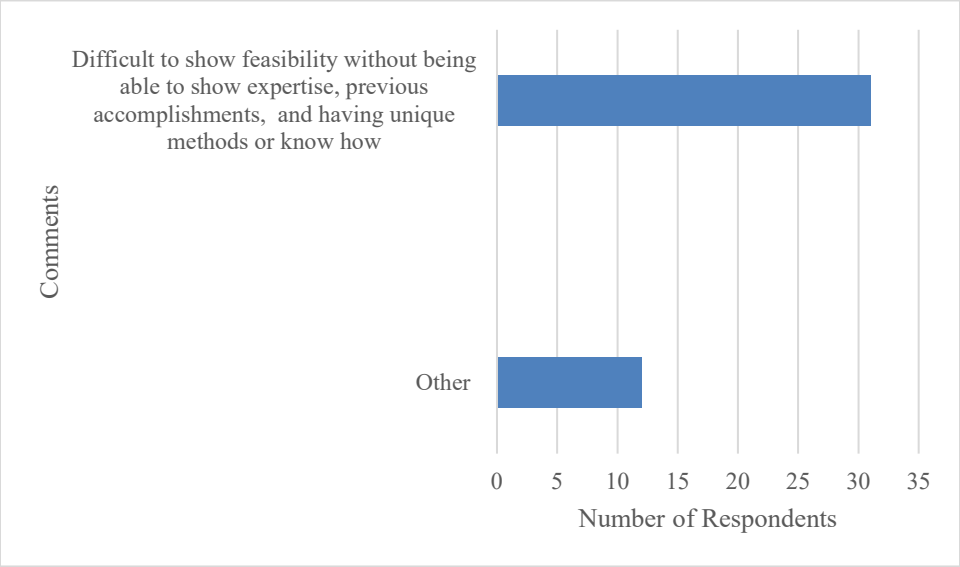


Figure 12. How Anonymization Instructions Affected Respondents' Ability to Convey the Transformative Potential in the Research Strategy

Question: Please indicate the number of times that you have submitted or won a TRA award prior to 2020.

Seventy-four of 105 respondents (70.5%) had not submitted an application to a prior TRA FOA (Figure 13). Twenty respondents reported having submitted 1 prior application, 8 had submitted 2 prior applications, 1 had submitted 3 prior applications, and 2 had submitted 5 or more prior applications. Respondents whose applications were administratively withdrawn in Pre-phase I or Pre-phase II were no more likely than respondents whose applications were not administratively withdrawn to report having submitted a previous application to a TRA award.

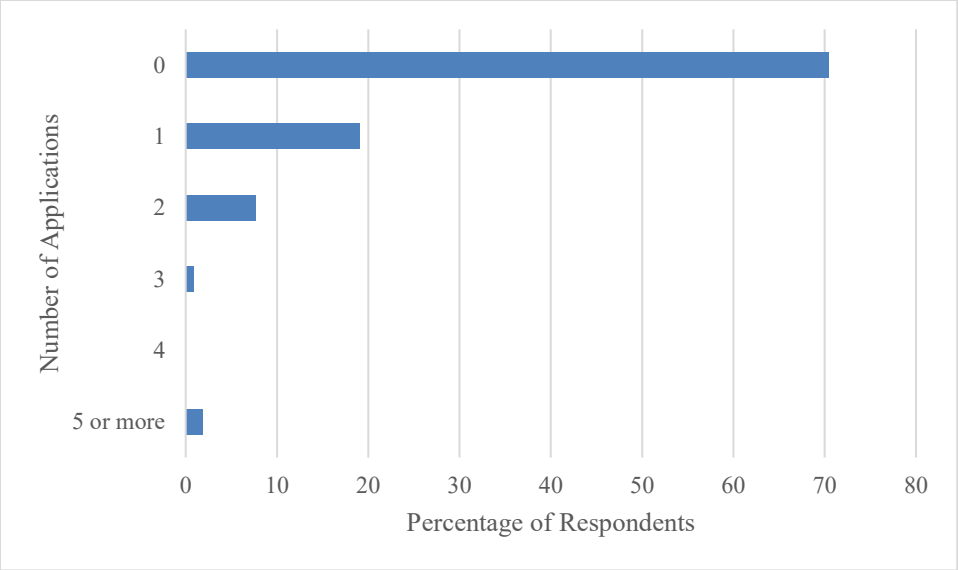


Figure 13. Number of Previous TRA Applications Submitted by Current TRA Applicants

Ninety-five of 101 respondents (94%) reported not receiving a prior TRA award 4 of 101 respondents had received 1 prior TRA award, and 2 respondents had received 2 prior TRA awards (Figure 14).

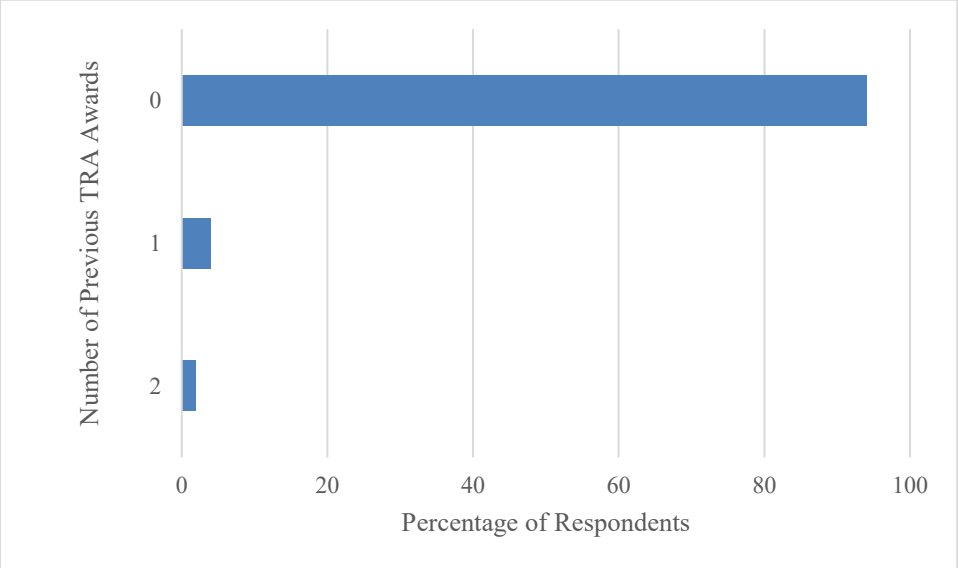


Figure 14. Number of Previous TRA Awards Submitted by Current TRA Applicants

Question: Please indicate if your institution provided any guidance or resources to assist you in writing your anonymized TRA application.

Twenty-one of 105 of respondents (20%) reported receiving guidance or resources from their institution ($p < 0.001$). Secondary analysis demonstrated that the percentage

of applicants who reported receiving guidance or resources from their institution did not differ significantly by race, gender, or ethnicity.

Respondents reported different types of assistance:

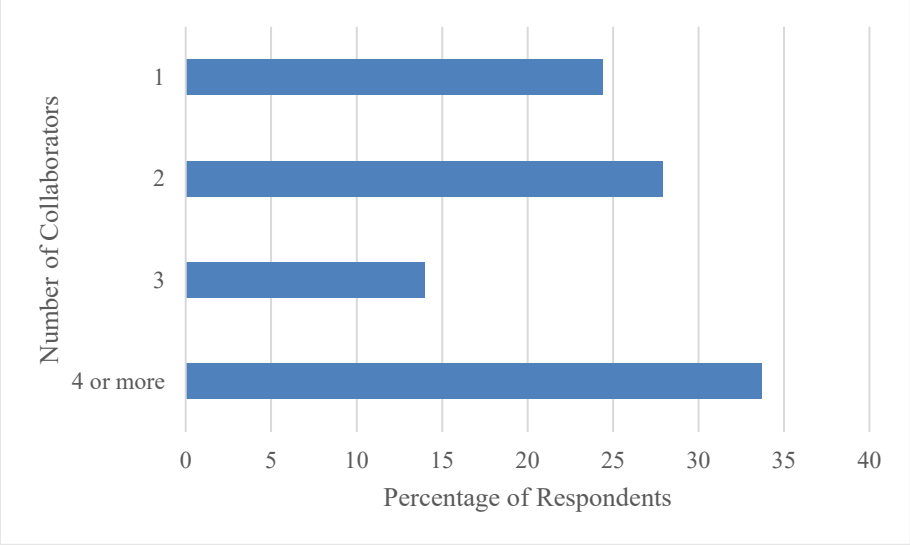
- writing an anonymized application (21 respondents)
- resources (18 respondents)
- review, editing, feedback, or check of their applications against the TRA FOA instructions (13 respondents)

Response categories that summed to less than five similar responses were grouped as *other*. Within this *other* category, seven respondents included:

- review of budget proposal
- general institutional review of the application
- administrative assistance (i.e., internal approvals)
- financial compensation for the time it took them to write their proposal
- active institutional role collecting the materials from them or coordinating with co-investigator

Question: Please indicate if your application had one or more collaborator(s).

As with many NIH grants, many respondents had collaborators in preparing their application: 85 of 105 respondents (81.0%) indicated they had at least 1 collaborator on their application ($X^2_1 = 78.02, p < 0.001$). Twenty of those 85 respondents reported 1 collaborator, 24 reported 2, 12 reported 3, and 29 reported 4 or more (Figure 15).



Respondents included only include those who reported at least one collaborator.

Figure 15. Number of Collaborators on Each Application

Question: Please indicate if anonymizing the identity of your collaborator(s) was challenging.

In the Specific Aims

Forty-seven of 85 respondents (55%) who reported having at least 1 collaborator indicated that anonymizing their collaborator(s) was challenging ($p = 0.386$).

Thirty-six of the 47 survey respondents indicated that anonymizing the identity of their collaborator(s) was challenging (Figure 16). They reported that:

- anonymized review reduced their ability to show necessary team expertise, knowledge, or tools to complete the proposed research (26 respondents)
- well-known collaborators precluded their inclusion in the application (12 respondents)
- applicant was unable demonstrate the uniqueness or importance of their research (12 respondents)

Response categories that summed to less than five similar responses were grouped as *other*. Within this *other* category, four respondents avoided mentioning collaborators and reported that anonymizing personal identifiers in preliminary data was difficult.

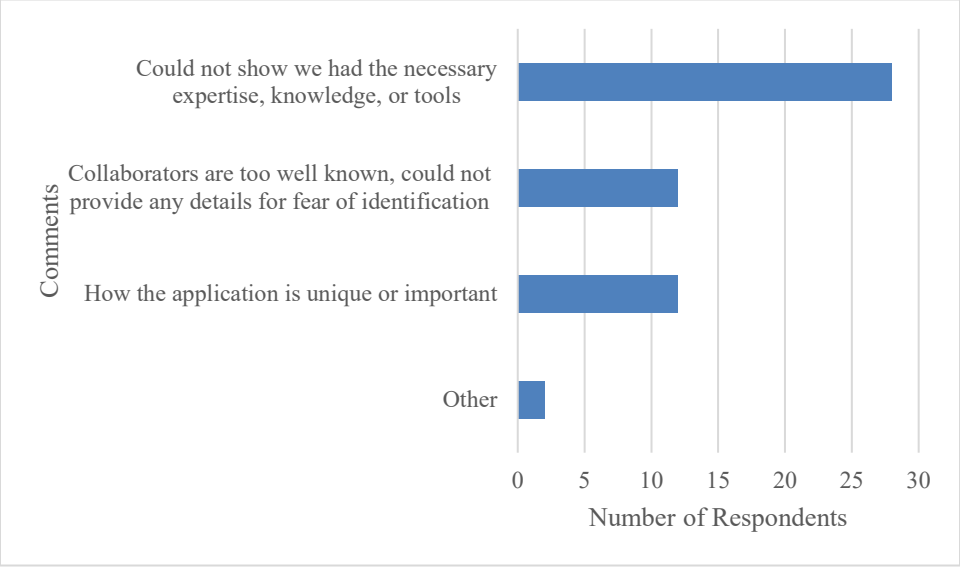


Figure 16. Information That Was Challenging to Anonymize in Specific Aims Because of Collaborators

In the Research Strategy

Thirty of 47 respondents (64%) described challenges to anonymizing the Research Strategy because they had collaborators (Figure 17). Respondents cited their inability to:

- include information (e.g., background and expertise) that showed their team was qualified to perform the proposed research (16 respondents)
- demonstrate the feasibility of their proposed research (8 respondents)
- describe approaches, methods, or protocols unique to collaborators and the research team (7 respondents)

Response categories that summed to less than five similar responses were grouped as *other*. Within this *other* category, eight respondents cited preliminary data would have revealed their collaborator's identity, references to previous research, and infrastructure that would enabled their proposed research.

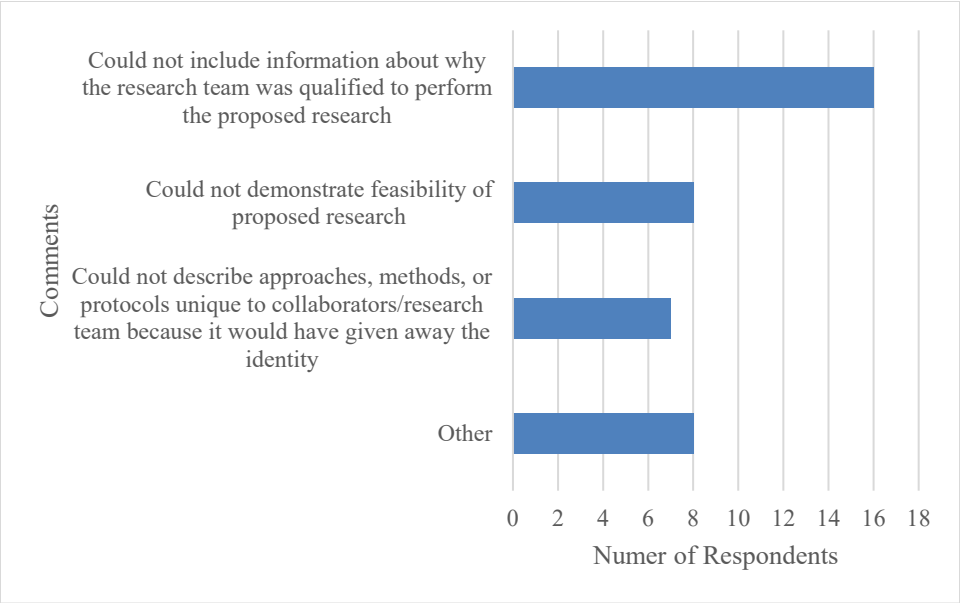


Figure 17. Information That Was Challenging to Anonymize in Research Strategy Because of Collaborators

Question: Please indicate if you included any preliminary data in your application.

Ninety-two of 103 respondents (88.5%) included some amount of preliminary data in their application ($p < 0.001$).

Question: Please indicate if anonymized review influenced your decision to include preliminary data in your application.

Twenty-eight of 92 respondents (30.4%) included some amount of preliminary data in their application because the review process was anonymized ($p < 0.001$).

Question: Please indicate if anonymized review influenced the amount or type of preliminary data included in your application.

Thirty of 92 respondents (32.6%) who included preliminary data reported that the anonymization process influenced the amount or type of preliminary data that they included ($p < 0.001$). Twenty-five of these 30 respondents described these factors (Figure 18). The impacts cited included:

- removing or decreasing the amount of preliminary data to avoid identification (13 respondents)
- providing preliminary data that had not been published nor presented (7 respondents)
- including more preliminary data than normal to convey expertise and feasibility (6 respondents)

Only two responses were in the *other* category, and they are not reported here.

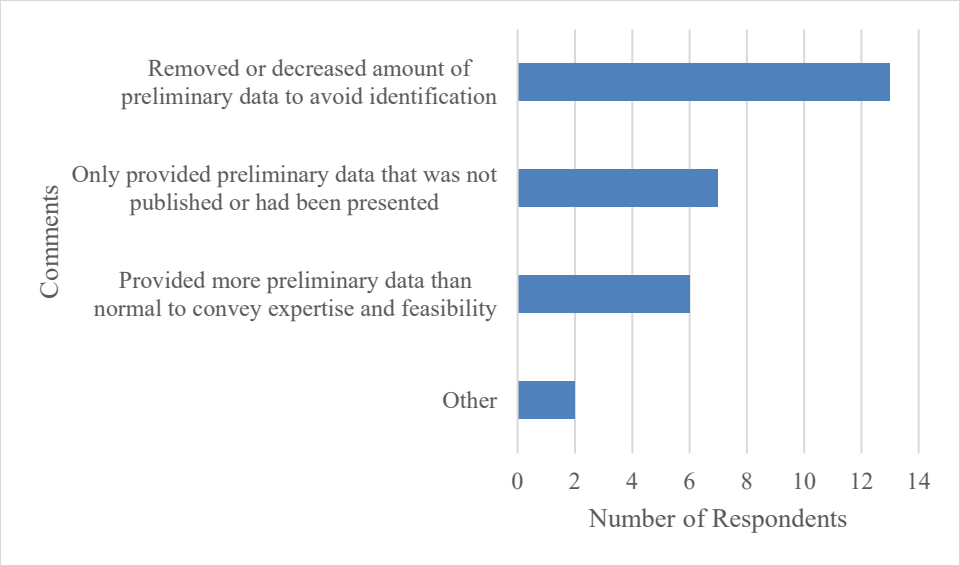


Figure 18. How Anonymized Review Influenced the Amount or Type of Preliminary Data Included in One's Application

Question: Please indicate if anonymized review influenced your decision to exclude preliminary data from your application.

Twelve respondents did not include preliminary data in their applications, 5 of the 12 (41.7%) indicated that the anonymized review process influenced their decision to exclude preliminary data ($p = 0.774$).

Response categories that summed to less than five similar responses were grouped as *other*. Within this *other* category, five of the respondents reported they felt that the TRA FOA discouraged preliminary data and only referenced published papers or they did not know how to provide preliminary data anonymously.

Question: Are there any additional comments you would like to provide to NIH about the anonymized TRA application process?

Sixty-three survey respondents provided additional comments about the anonymized TRA application process (Figure 19). These answers varied greatly, with 33 responses conveying positive sentiments about the anonymization process, including comment that the anonymization process was more fair and equitable than traditional review, and that NIH should have more anonymized review processes.

Fourteen respondents felt negatively towards anonymization, with many noting they felt that the anonymous application process is a waste of time, will not reduce bias, or that the requirements add an administrative burden to the application process.

Eight respondents were concerned about showing feasibility of the proposed research, and seven respondents felt that it was difficult to write strong, effective research proposals under anonymous application conditions. Five respondents liked the idea behind anonymized review, but were not convinced that the process would reduce bias.

Four respondents expressed interest in hearing the outcome of how anonymized review impacts diversity. Ten respondents' comments were grouped into the *other* category. Within this category some respondents felt that the inclusion of preliminary data would have led to applicant identification, or questioned if the problem with perceived bias of the traditional NIH review system lie instead with the traditional NIH reviewers.

Some respondents expressed interest in being notified of their rejection and reason for rejection if they do not pass a level of review. Others noted that preliminary data are necessary and they should be allowed to include it as long as the rest of the application was anonymized to the best of their abilities. Another respondent opined there was neither harm nor benefit from an anonymized review.

Other respondents felt that anonymized review may discourage certain applicants to apply, and that applicants should not be penalized if they have different writing styles.

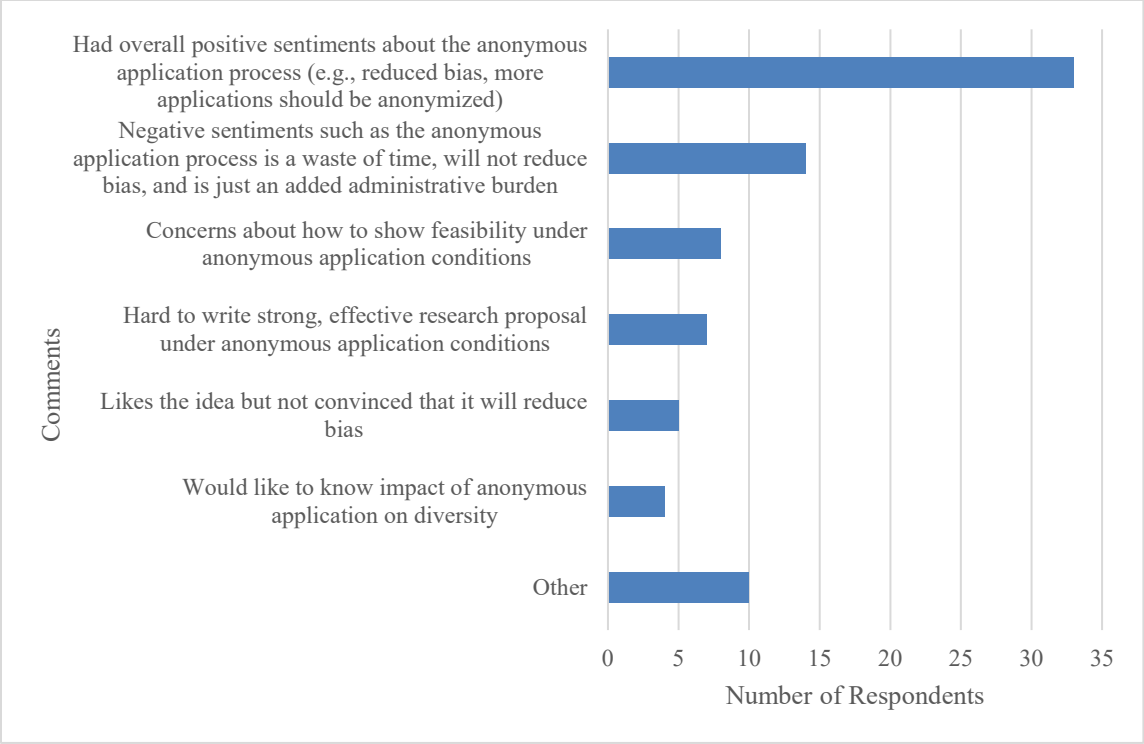


Figure 19. Additional Comments about the Anonymized TRA Application Process

C. Administrative Review Survey: Specific Aims

In Pre-phase Ib, the NIH staff review the Specific Aims in each application for compliance with the anonymity instructions (Figure 20). The administrative review survey queried staff about the applicants' compliance, their review instructions, and the process through which they conducted their review (Figure 21) (NIH 2020b). The survey is provided in Appendix E and the descriptive statistics in Appendix F.

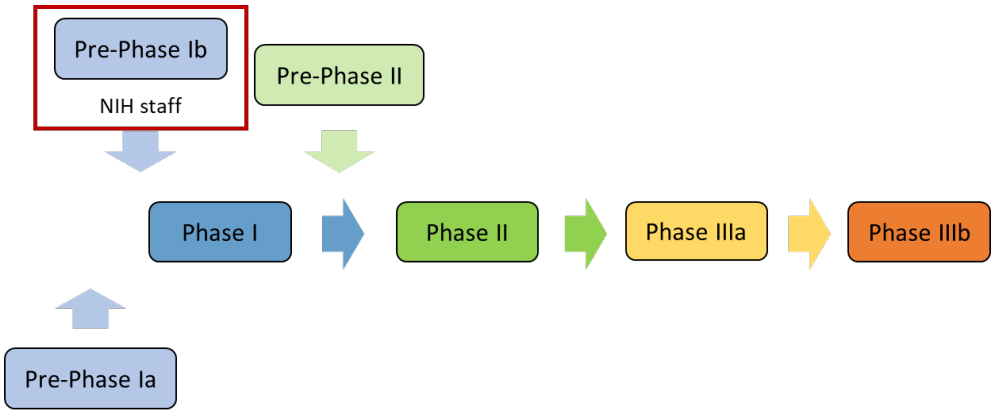


Figure 20. Pre-phase 1b Focuses on the NIH Staff Review of Specific Aims

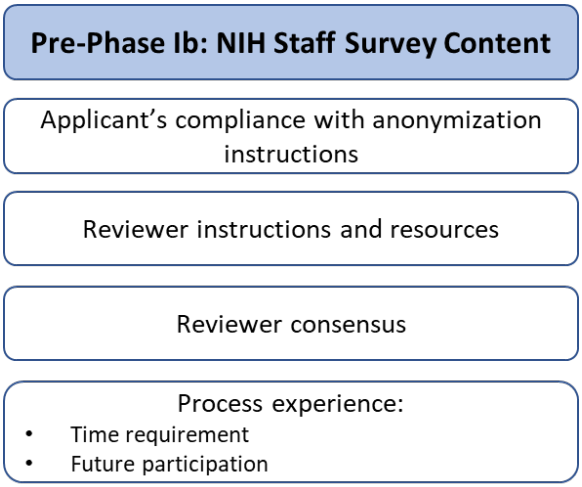


Figure 21. Administrative Review Survey: Specific Aims Survey Content

1. Survey Administration

An email invitation to participate in the survey was sent to 50 NIH Pre-phase Ib reviewers on December 1, 2020. Each email contained a personalized survey link. A reminder was sent to all reviewers on December 9 and a final reminder sent on December

11. The survey was closed on December 15, and the survey response received on December 17 was not included in the analysis.

2. Results

a. Response Rate

Fifty reviewers completed 43 surveys. The response rate is 86%.

b. Survey Data

Question: Please indicate if the instructions you received were sufficient for you to assess if the Specific Aims were anonymized.

Thirty-seven of 43 respondents (86.0%) indicated that the instructions they received were sufficient for them to assess if the Specific Aims were anonymized ($X_1^2 = 41.86, p < 0.001$).

Six respondents provided additional comments, with all six indicating that they would have appreciated clearer anonymization instructions (Figure 22), while a subset of these respondents was unclear how to determine if applications that contained preliminary data, unpublished data, or publicly available data would disclose the applicant's identity. Other respondents felt that the TRA FOA instructions were too vague, subjective, and ambiguous in general, and that the final review criteria, "any other text from which the identity of any participating individual or institution can be reasonably inferred," was difficult to interpret. Respondents also indicated uncertainty regarding whether references to investigator attributes or accomplishments would reveal identity and cited the need for more examples of acceptable and unacceptable language within the Specific Aims.

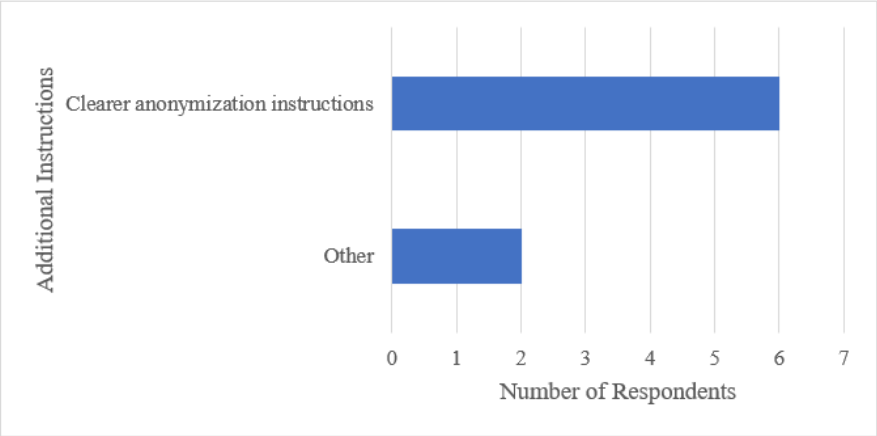


Figure 22. Additional Instructions That Would Have Been Helpful for Reviewers to Determine Whether Anonymity of the Specific Aims Was Compromised

Question: For the anonymized Specific Aims that you believe contained "other" identifying information, please indicate whether you made this determination because you possess related knowledge and expertise.

Ten respondents noted portions of the Specific Aims that contained identifying information *other than* the specific categories listed within TRA FOA. Six of 10 respondents indicated they were able to make that determination because they possessed knowledge and expertise related to the application ($X_1^2 = 0.2, p = 0.655$).

Question: Please describe the process through which you and your fellow reviewer reached consensus.

Two NIH staff reviewed each application. Twenty-one of 43 (49%) respondents described the process through which they reached consensus when conclusions differed regarding anonymity (Figure 23). Eighteen of the respondents cited discussion via email or video meeting platforms as the method used to reach consensus. Five respondents reported that they involved the program officer in their deliberation. The program officer sometimes participated as a *tie breaker* if reviewers could not reach consensus, or helped to provide general clarification to the reviewers during the review process.

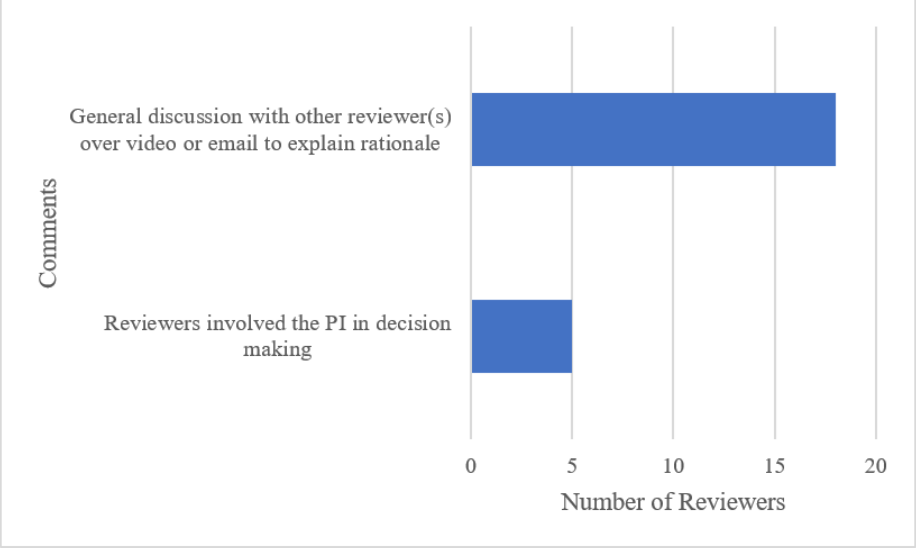


Figure 23. Factors That Helped Reviewers Reach Consensus with Other Reviewers

Question: Please describe what prevented you and the other reviewer from reaching consensus.

This question received mixed responses from the survey respondents. Ten of 14 respondents who did not reach initial consensus provided free response answers in explanation; however, several were confused by the survey question, indicating that they eventually reached consensus and the question did not apply to them (Figure 24). Seven described the initial disagreement with the reviewers, 5 of 7 respondents felt that they had

a different interpretation of the TRA FOA instructions than their counterpart. Two respondents disagreed about the reviewer's role in conducting an online search to determine if they would identify the applicant based on the information given.

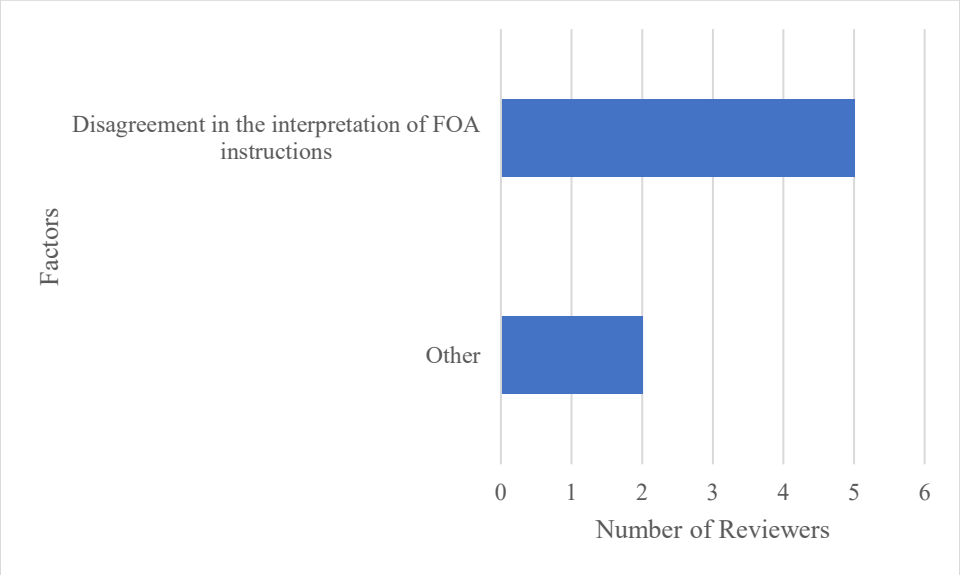


Figure 24. Factors That Prevented Reviewers from Reaching Consensus

Question: Our records indicate that you reviewed [number specified for each reviewer] Specific Aims. Please estimate, on average, how long (in minutes) it took you to review the Specific Aims for each application.

On average, applicants spent 15 minutes reviewing each Specific Aims section.

Question: If anonymized review is expanded at NIH, please indicate your level of interest in participating in future exercises.

Thirty-five of 43 respondents (81.4%) reported they were *somewhat interested* or *very interested* in participating in future, expanded exercises of anonymized review at NIH, while 5 indicated they were not very or not at all interested in participating ($X^2_1 = 39.31, p < 0.001$) (Figure 25).

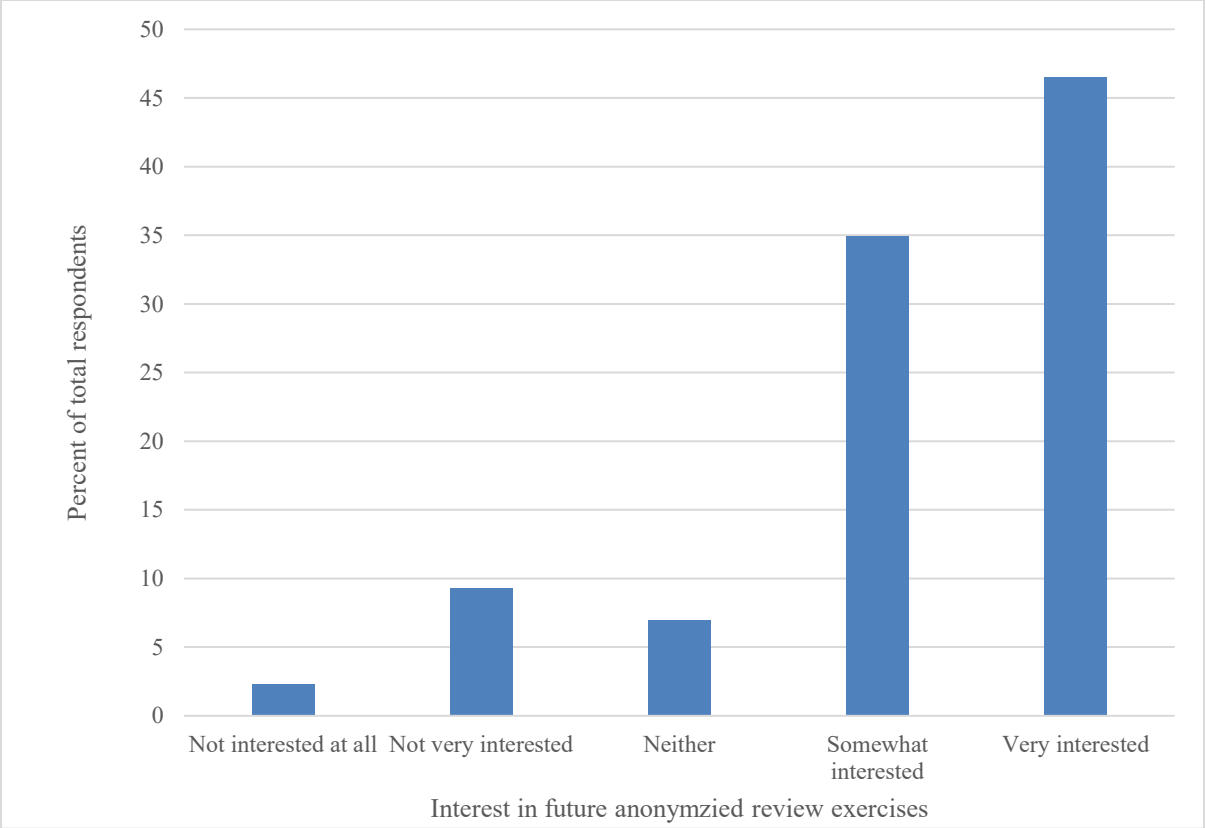


Figure 25. Reviewer Interest in Participating in Future, Expanded Efforts at Anonymized Review at NIH

Question: Please provide any additional information that you would like NIH to consider about administrative review of the anonymized Specific Aims.

Twenty-five of 43 respondents (58%) reported a diverse range of responses (Figure 26). Ten responses mentioned clarifying the PI and reviewer instructions, with some respondents indicating they would like increased general clarity in the instructions, more examples of acceptable and unacceptable phrases within instructions, and more guidance on how to use variants of the phrase, "we have shown." Other respondents would have liked the instructions to include more guidance on how to include published data, unpublished data, preliminary data within the grant, or preliminary data that have been presented in meetings within the application. And others indicated their confusion about review of applications that included unique reagents and techniques.

Five respondents commented on the program officer's involvement in the review process, with several giving positive feedback. Other respondents felt that there may be a conflict of interest with the program officers who had previously engaged with the

applicants, possibly affecting the review process, or that more program officers should participate in the review process.

Thirteen responses coded into the *other* category with respondents reporting that the anonymized TRA process was a valuable exercise, the process would improve over time, or that they were interested in seeing the results of the TRA process. Other respondents felt that the two-step process is more time consuming than necessary and anonymized review may not matter given that the identity of the applicant is revealed in the final round of the review. Other respondents sought clarification about the reviewer's role in externally investigating the applicant based on information included in the application, had engaged in a helpful group conversation with other reviewers, or felt that engaging in a group discussion with multiple reviewers would have helped them to clarify their thinking.

Other respondents felt that much of the review process to ensure anonymization could be completed through ML or other computer methods. There was also confusion regarding whether reviewers were expected or allowed to search for the applicant's published research using key words from the application, and if finding revealing information online was technically a breach of applicant anonymity.

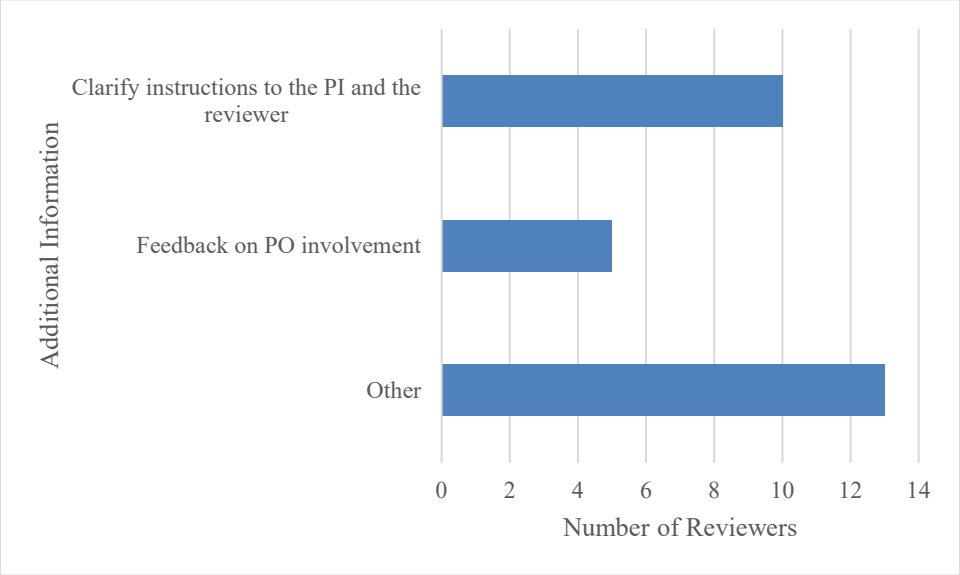


Figure 26. Additional Information from Pre-phase Ib Administrative Review Survey Respondents

D. Editorial Board Review: Phase I

In Phase I, Editorial Board members review a mix of Specific Aims from applications submitted to two TRA FOAs and one TRA Notice of Special Interest that were published in parallel (Figure 27).[14] The Editorial Board member survey queried the reviewers about their ability to identify the applicant, assess transformative research potential using only the Specific Aims, and their thoughts on the anonymized review process (Figure 28) (NIH 2020b). The survey questions are located in Appendix G and the data analysis in Appendix H.

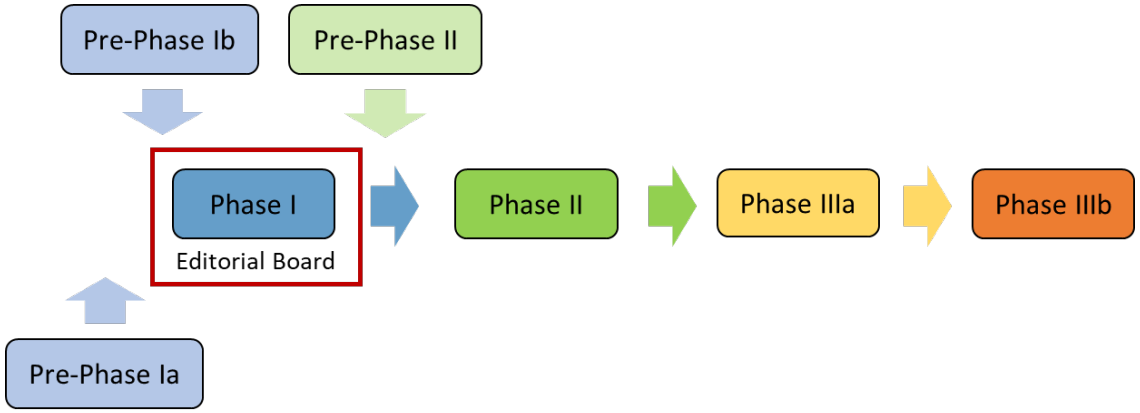


Figure 27. Phase I Focuses on Editorial Board Review

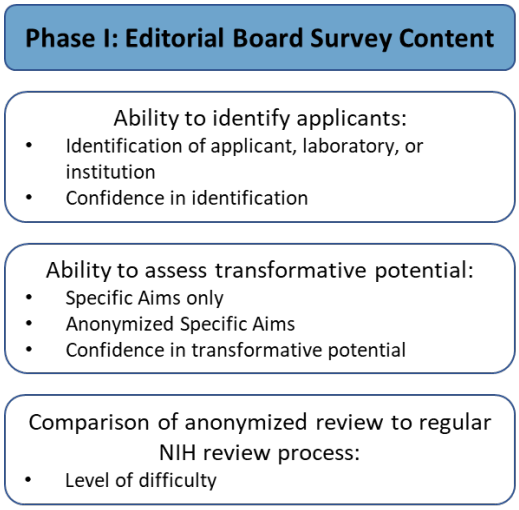


Figure 28. Editorial Board Review Survey Content

[14] General TRA: RFA-RM-20-013, Notice of Special Interest (NOSI) Common Fund TRA FOA for ALS-related research: NOT-RM-20-019, and the Emergency TRA FOA for SARS-CoV-2-related research: RFA-RM-20-020

1. Survey Administration

STPI sent an email invitation to participate in the survey to 27 EB members on December 10, 2020 using a generalized link to the survey on the Alchemer platform. Two reminders were sent prior to the December 18 close date. The data were downloaded from Alchemer on December 21 and included one survey received on December 19.

2. Results

a. Response Rate

Due to a validation error in Alchemer, three EB members could only complete the first two questions, and their surveys were categorized as *partially complete* by Alchemer. The validation error was corrected and no further issues with completing the surveys were recorded.

Twelve of 27 EB members completed the survey. The response rate is 44%.

b. Survey Data

For this portion of the analysis, the EB members were asked to recall their experience with a set of TRA Specific Aims that were derived from three TRA opportunities (see footnote 14).

Question: Please indicate the number of anonymized Specific Aims you reviewed.

The 12 respondents evaluated an average (± SE) of 50 (± 0.92) Specific Aims sections each.

Question: Of the [number assigned per reviewer] Specific Aims you reviewed, please indicate if there were any for which you believe you could identify the applicant or collaborator(s), lab group, or institution.

None of the 12 respondents who responded reported being able to identify the applicant, collaborators, lab group, or institution for any of the Specific Aims they reviewed ($p = 0.001$).

Because no respondents indicated they were able to identify the applicant, collaborators, lab group, or institution, no respondents answered the question of how many Specific Aims they were able to identify the applicant, collaborators, lab group, or institution for any application.

Question: Please indicate if there were any applications for which you could not assess transformative potential because you only had access to the Specific Aims.

Five of 12 respondents (41.7%) indicated that they were unable to assess the transformative potential of some applications because they only had access to the Specific

Aims ($p = 0.774$). The 5 respondents indicated that they could not assess the transformative potential for 13 applications.

Question: Please indicate if there were any applications for which you could not assess transformative potential because they were anonymized.

Three of 12 respondents (25.0%) indicated that they were unable to assess the transformative potential of the applications because they were anonymized ($p = 0.14$).

Question: Please indicate, on average, how confident you were in your ability to determine if the research in the anonymized Specific Aims you reviewed was transformative.

Eleven of 12 respondents (91.7%) indicated that they were *somewhat* or *very confident* in their ability to determine if the research in the anonymized Specific Aims sections they reviewed was transformative, and one respondent indicated they were *not very confident* (Figure 29).

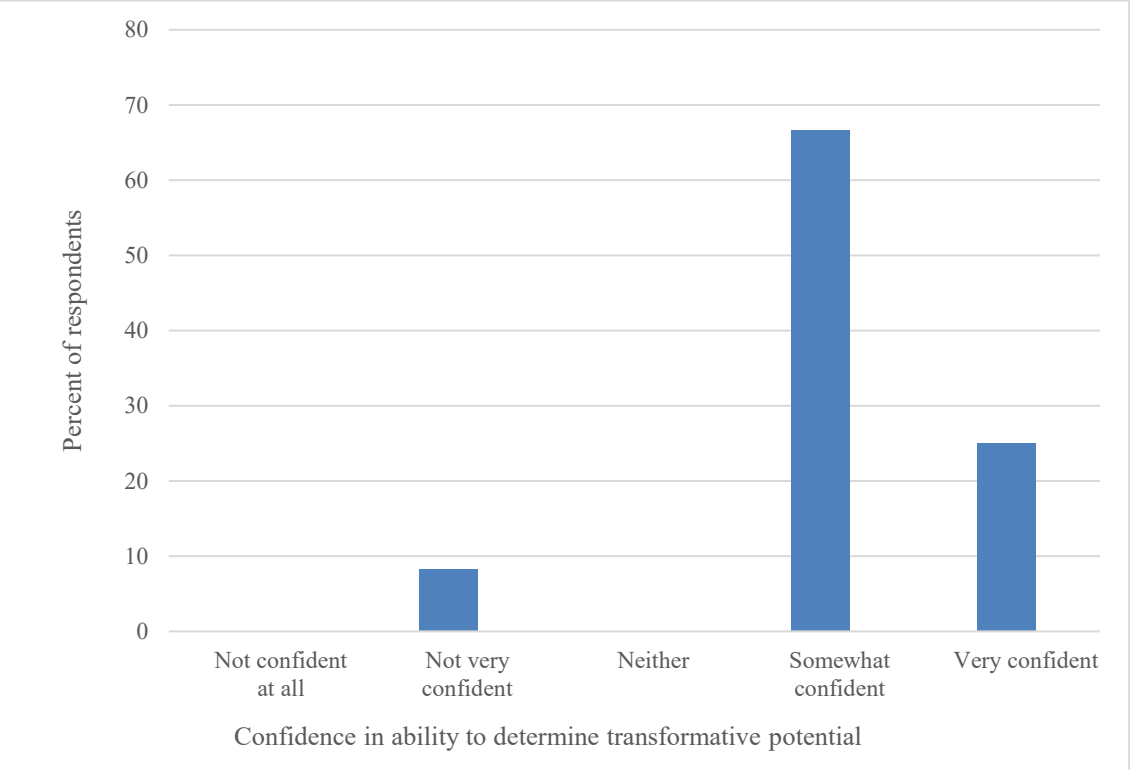


Figure 29. Reviewer Confidence in Ability to Assess Anonymized Specific Aims for Transformative Potential

Question: Please indicate how difficult it was to score anonymized Specific Aims compared to the traditional NIH review process.

Three of 12 respondents (25.0%) indicated that it was easier or much easier to score anonymized Specific Aims compared to the traditional NIH review process, while 4 of 12 indicated that it was more difficult, and 5 selected about the same as traditional NIH review (Figure 30).

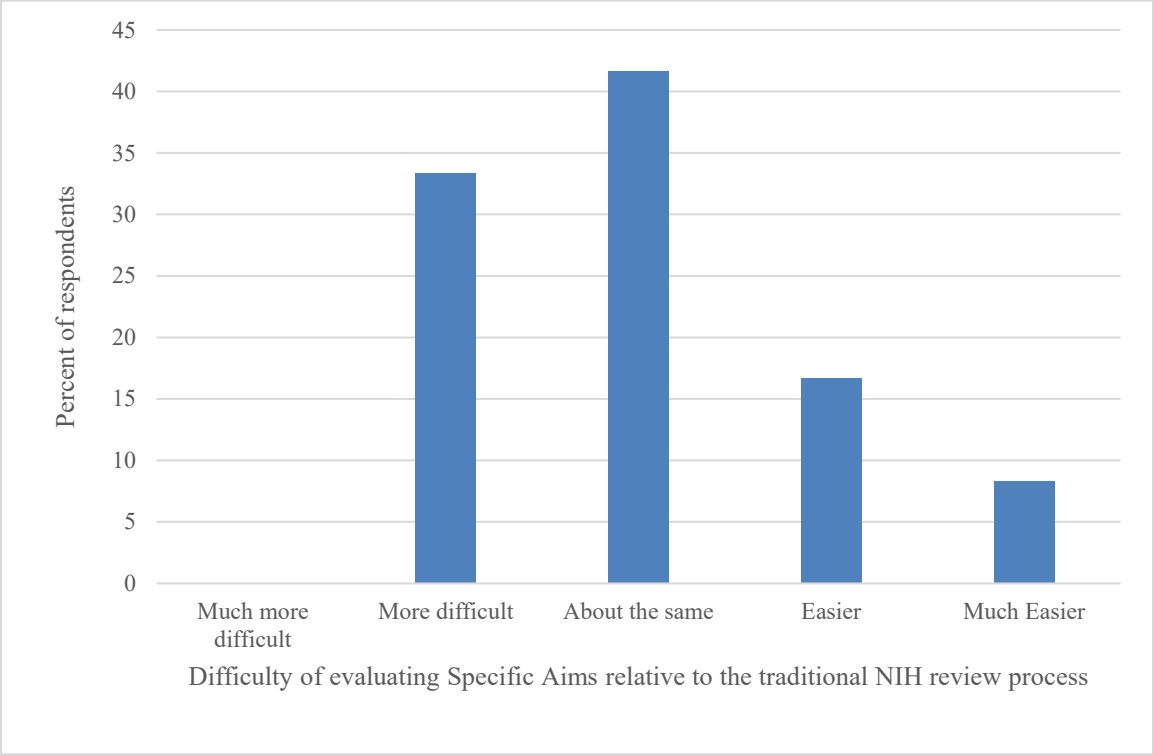


Figure 30. Reviewer Indication of Difficulty in Assessing the Transformative Potential of Anonymized Specific Aims Compared to the Traditional NIH Review Process

Seven of 12 respondents described the differences between the anonymized review processes for the anonymized Specific Aims versus the traditional NIH review.

More difficult to score:

- qualifications or experience of the research team behind the proposed research (4 respondents)
- feasibility of the research due to lack of detail, or because the proposals were outside of the reviewer's expertise (4 respondents)
- access to the applicant's publication record or supporting data (4 respondents)

Easier to score:

- Specific Aims were shorter and more focused on the science of the research proposed (3 respondents)

Two respondents felt less confident in their score because the Specific Aims are less detailed than the research proposal.

Question: Are there any additional comments you would like to provide to NIH about the use of anonymized Specific Aims for this step of the review process?

Five respondents felt positive about this step in the review process. Two respondents noted that they thought the applicants should have been given more instructions about the application process, including more instruction on formatting, and more guidance on the structure and content of the Specific Aims.

Three respondents grouped into the *other* category requested more reviewers with a wide background on each application, expressed concern that the inability to assess the investigator's background may reduce the reviewer's ability to identify significant research, and recommended that the anonymization review concept be expanded and utilized broadly, with the revision to score the Research Strategy before the Specific Aims.

E. Administrative Review: Research Strategy Survey

In Pre-phase II, the NIH staff review the Research Strategy in each application for compliance with the anonymity instructions (Figure 31). The administrative review survey queried staff about the applicants' compliance, their review instructions, and the process through which they conducted their review (Figure 32) (NIH 2020b). Survey questions are located in Appendix I and survey data in Appendix J.

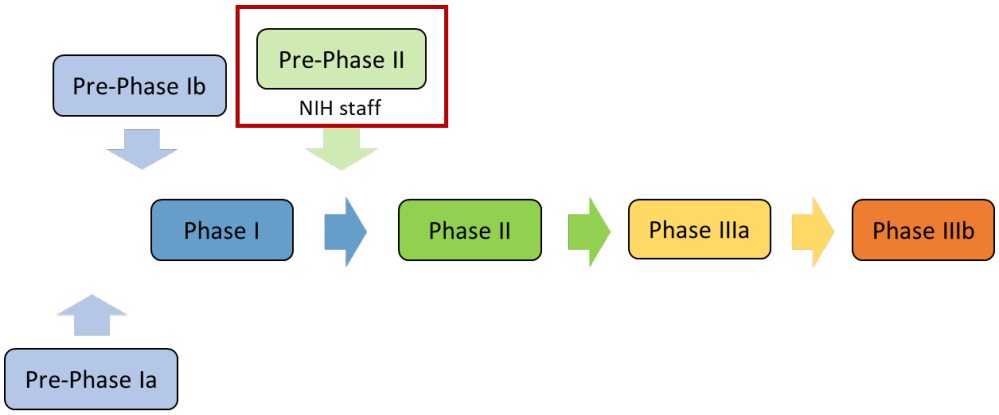


Figure 31. Pre-phase II Focuses on the NIH Staff Review of the Research Strategy

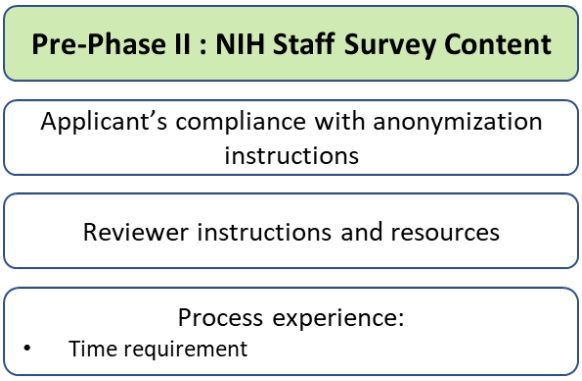


Figure 32. Administrative Review: Research Strategy Survey Content

1. Survey Administration

STPI sent an email invitation to participate in the survey to 46 NIH reviewers on February 2, 2021. The email contained a personalized survey link. Reminders to complete

the survey were sent on February 8 and the survey closed on February 12. The survey data was downloaded from Alchemer on February 19.

2. Results

a. Response Rate

Twenty-three of the 46 reviewers completed the survey. The response rate was 50%.

b. Survey Data

Question: Please indicate if the instructions you received were sufficient for you to assess if the Specific Aims were anonymized.

Twenty-one of 23 (91%) felt the instructions they received were sufficient to assess if the research strategies were anonymized, and 2 (9%) felt that the instructions were insufficient. The two respondents cited lack of clarity for the definition of anonymity, the need for additional materials, and instructions that apply to all possible applicant scenarios so disagreement between reviewers assigned to the same application can be avoided.

Question: Our records indicate that you reviewed [application specified]. Research Strategies from the general TRA FOA. Please estimate, on average, how long (in minutes) it took you to review the Research Strategy for each application in the general TRA FOA.

Nine of 23 respondents (39%) reported spending 16 to 30 minutes reviewing a Research Strategy, with 6 respondents spending 31 to 60 minutes, 3 respondents 11 to 15 minutes, 2 respondents spending 5 or fewer minutes, 2 spending more than 60 minutes, and 1 respondent spending from 6 to 10 minutes per Research Strategy (Figure 33).

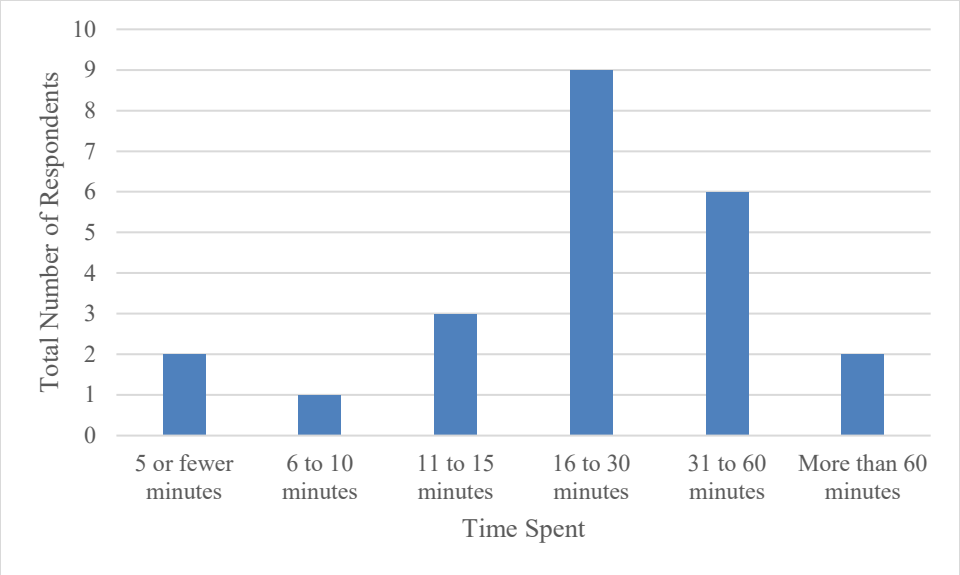


Figure 33. Average Estimated Time Each Reviewer Spent Reviewing Research Strategies

Question: If anonymized review is expanded at NIH, please indicate your level of interest in participating in future exercises.

Eleven of 23 respondents selected *very interested* as the gauge of their willingness to participate in anonymized review again, 5 respondents were *somewhat interested*, 3 each were *not very interested* or were *neither interested nor disinterested*, and 1 reviewer was *not interested at all* (Figure 34).

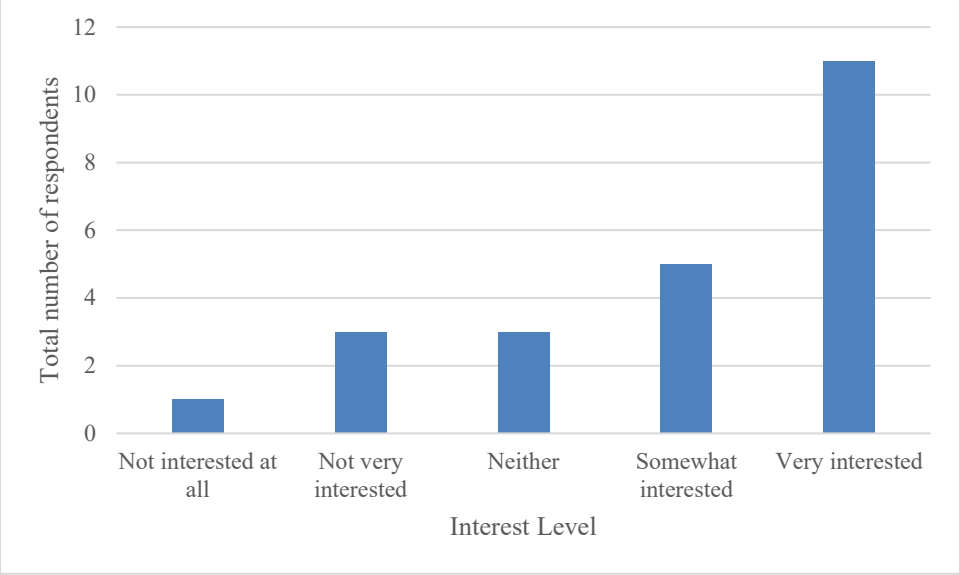


Figure 34. Reviewer Interest in Participating in Future Exercises

Question: Please provide any additional information that you would like NIH to consider about administrative review of the anonymized Specific Aims.

Eight survey respondents provided additional comments to NIH about Pre-phase II review. Four respondents felt that the application instructions were unclear, which led to different interpretations of the instructions and review process. Of these four, two respondents specifically mentioned confusion about the use of citations in general and citations for methods versus data generated by the principal investigator (PI) specifically.

Individual responses included the note that the process was clear and straightforward, the TRA FOA should require the applicant to write in passive tense to place more emphasis on the project and not the individual, and a mechanism to allow the applicant to fix minor anonymization infractions would be helpful. Three respondents expressed positive sentiment about the process, praising the program officers for their efforts managing the review activities and noting that the review efforts were appropriate to the potential outcome for anonymized peer review. Other respondents felt that anonymity is useful for certain types of grant mechanisms and had different opinions on the utility of artificial intelligence (AI) or machine learning (ML) in the review process. Conversely, the other respondent felt that ML may not be effective given the subjective nature of the review exercise.

F. Phase II: Technical Review Survey

In Phase II, subject matter experts perform a technical review of the anonymized Specific Aims and Research Strategy using three NIH review criteria (significance, innovation, and approach; Figure 35). The technical survey queried the reviewers about their ability to identify the applicant and the sufficiency of information to perform a technical review (Figure 36) (NIH 2020b). The survey questions are located in Appendix K and the survey data in Appendix L.

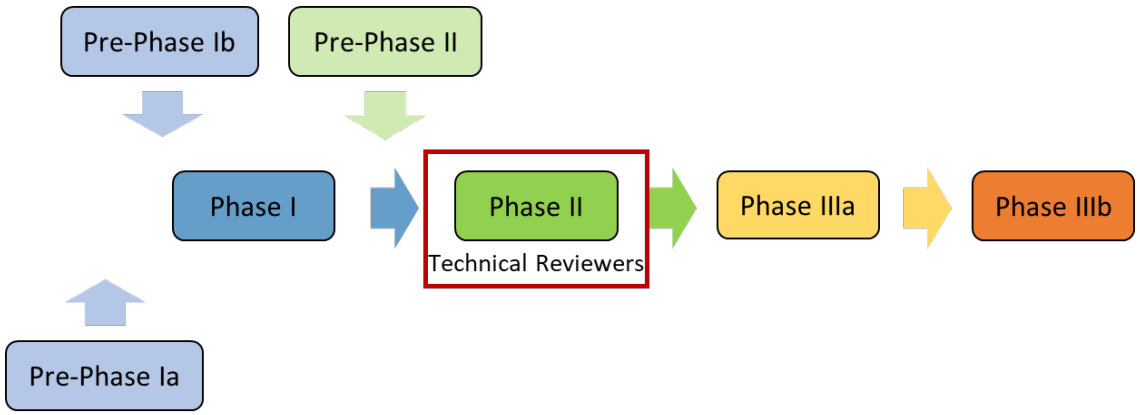


Figure 35. Phase II Focuses on the Technical Review

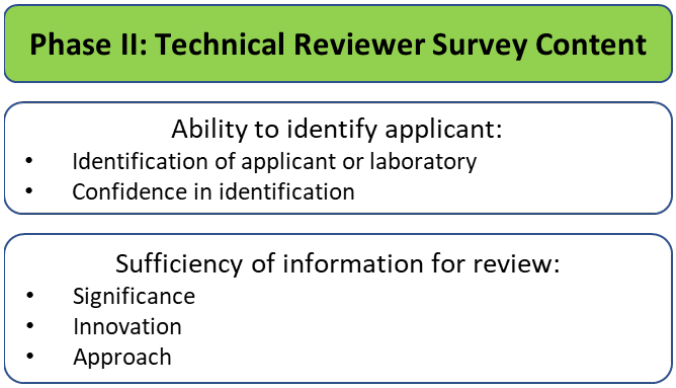


Figure 36. Phase II Technical Review Survey Content

1. Survey Administration

In contrast to the email invitations issued for the previous surveys, the technical survey link was directly attached to the electronic review critique form used to document each review. Instructions were included within the survey that directed each reviewer to only fill out the survey. Because there was a second TRA FOA for COVID applications and a TRA Notice of Special Interest for amyotropic lateral sclerosis, STPI reviewed all

electronic forms to ensure only surveys for applications responsive to FY2021 TRA FOA: RFA-RM-20-013 were included in the analysis.

2. Results

a. Response Rate

Applications are reviewed by 3 technical experts. Fifty-four applications were assessed by 149 reviewers for a total of 162 review critiques. One hundred of 149 reviewers responded to the survey. The reviewer response rate is 67%.

Because four respondents were assigned more than one application, 104 review critiques were received from the 100 respondents. The response rate for the 104 of the 162 possible review critiques is 64%.

b. Survey Data

Question: Please indicate if you could identify the applicant, lab group, institution, or the collaborator.

Eighty-four of 104 respondents (81%) reported they could not identify the applicant, lab group, institution, or collaborator, while 20 indicated they were able to identify groups ($p < 0.001$).

Eight of the 20 respondents who reported that they could identify an applicant, lab group, institution, or collaborator commented in the survey that the application contained unique technologies, methodologies, or resources. Five of the 20 respondents indicated that the applications contained specialized research topics or goals that aided them in deciphering the applicant's identity.

Thirteen of the 20 respondents who reported that they could identify an applicant, lab group, institution, or collaborator provided comments that summed to less than 5 similar responses per topic. Placed in the *other* category, these responses cited recognition of published figures, papers, or previous work; recognizable preliminary data; and inclusion of the PI's previous work or name of the lab group. Also cited as identifying information were the applicant's style of writing and an understanding of the scientific landscape within the field. One respondent conducted an internet search in an attempt to identify the applicant.

Question: Please indicate how confident you were in your ability to identify the applicant, lab group, institution, or the collaborator.

Twenty of 20 respondents (100%) who indicated they could identify the applicant, lab group, institution, or collaborator responded that they were either *somewhat confident* (N = 8) or *very confident* (N = 12) in their ability to identify the group (Figure 37).

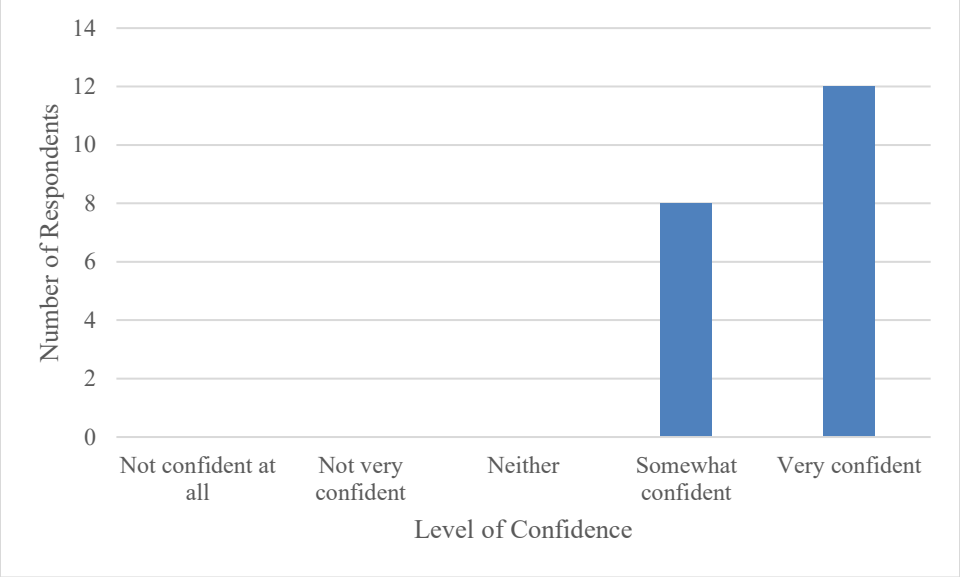


Figure 37. Respondent Level of Confidence in Their Ability to Identify the Applicant, Lab Group, Institution, or the Collaborator

Of the 20 respondents who indicated that they were *somewhat* or *very confident* in their ability to identify the applicant, lab group, institution, or collaborator, 8 respondents (40%) reported that unique technologies, methodologies, or resources in the Specific Aims or Research Strategy sections allowed them to identify the applicant or their instution. Six of the 20 respondents (30%) reported that the specialized research topic or goal allowed them to identify the applicant information. The respondents also reported other information or strategies that allowed them to identify the applicant, lab group, institution, or collaborator:

- Information in the application that the respondents used to conduct internet searches to narrow down the applicant
- Distinctive preliminary data
- Inclusion of already-published figures or data
- References to prior work
- Distinct styles of writing
- References to a lab group

To explore the accuracy of the 20 respondents in identifying the applicant, STPI sent a follow-up email asking the respondents to provide the name of the applicant. STPI could verify the contact emails of 17 of the 20 respondents using publicly available sources. These individuals received personalized links via email to a single question follow-up survey asking for the name of the individual whose anonymized application they reviewed. STPI could then verify this information through NIH QVR information.

STPI determined that the 17 individuals reviewed 15 unique applications, and 8 of the 17 individuals (47%) completed the follow-up survey. STPI confirmed that the 8 respondents reviewed unique applications, and 7 of 8 follow-on survey respondents were able to correctly identify the PI or institution.

Of the eight respondents who correctly identified the PI or institution, four of eight respondents (50%) cited specialized research topics or goals that allowed them to identify the applicant or institution. Other information in the application that allowed the eight respondents to correctly identify the PI or institution include information on unique technologies, methodologies, or resources; already-published figures, distinctive preliminary data; general knowledge of the field; reference to prior work; or use of internet searches to narrow down the applicant's identity.

Question: Please indicate if you felt you had sufficient information in the anonymized Specific Aims and Research Strategy to evaluate if the proposed research was significant, innovative, logical and compelling, and feasible.

Respondents reported sufficient information to assess:

- significance: 92 of 104 respondents (88.5%) ($p < 0.001$)
- innovation: 89 of 103 respondents (86.4%) ($p < 0.001$)
- logical and compelling: 85 of 104 respondents (81.7%) (, $p < 0.001$)
- feasible: 64 of 104 respondents (61.5%) ($p = 0.024$)

Respondents provided 48 responses describing what additional information would have been helpful in the review, citing:

- more details about the PI's track-record and experience, including details like publications, institutional resources, and research infrastructure (21 respondents)
- more preliminary data (5 respondents)
- more experimental or methodological details (12 respondents)
- more information in general to evaluate feasibility (19 respondents)

Twenty-eight responses expressed concern that the anonymized review was less rigorous and that it was more difficult to conduct an objective merit review without the inclusion of references.

Although some comments are similar to those listed above, 20 respondents provided comments that summed to less than 5 similar responses per topic. Placed in the *other* category, respondents indicated an interest in more details in general to evaluate approach, significance, environment, and innovation—and more details to avoid making decisions based on the applicant's lofty and unsupported statements, giving the applicant the "benefit

of the doubt" because the absence of details made it difficult to tell if the applicant could complete proposed research. In contrast, one respondent felt that fewer details in the Specific Aims allowed the reviewer to focus more on the transformative potential of the application.

Additional information cited as potentially helpful includes: a clear list of Specific Aims, budget, biosketches, and metrics for success. Other comments noted that the anonymized review process was a poor use of both NIH and reviewer resources, a disservice to ambitious proposals, and that the applicant should be informed more effectively about the anonymized application requirements.

Question: Are there any additional comments you would like to provide to NIH about the technical review step in the anonymized TRA review process?

Sixty-eight respondents of 104 survey responses (65%) described additional comments about the technical review step (Figure 38). Because of the large number of responses to this question, *other* responses that summed to less than five similar responses per topic are reported, as appropriate, under the major categories listed in Figure 38.

Twenty-three of the 68 respondents included comments about wanting the applications to include references. Within this *References* category, 21 of the 23 respondents felt that it is challenging to review applications rigorously without reference lists or bibliographies. Comments grouped into the *References (other)* category cited:

- inability to assess applicant validity without references
- review as a more valuable process for the reviewer when references are included, and without references there was less interest in serving as a reviewer
- time saved by not having to perform large literature searches that could identify the applicant; however, they reported not learning as much about the application
- more time and effort spent trying to review the literature but potentially missing key references
- bibliographies allow for better assessment of scientific rationale
- leaving out references as an ineffective method to protect the applicant's identity, thus adding more unnecessary burden to the reviewer

Nine of the 68 respondents included comments about the applicant's identity and were grouped in a main category for *Identity*. Within this category, seven people felt that they could discern the applicant's identity. Comments grouped into the *Identity (other)* category reported:

- knowing the identity of the applicant does not affect the way they score and they should be trusted to make an impartial decision

- potential for an uneven playing field because some PIs are more easily identified than others
- being outside the applicant's field of research not knowing the applicant even if they had the name
- another felt that their own knowledge of the landscape of the field led them to guess the applicant's identity

Seventeen of the 68 respondents included positive sentiments towards the *anonymized* review. Within the *Positive* category, 13 of 17 respondents expressed support for the goal of the anonymized review process, with 5 considering the review easier than the R01 process due to the shorter and more focused structure. Five respondents provided responses that were grouped into a *Positive (other)* category: the reviewer preferred anonymized review to non-anonymized review; the reviewer had no issues with review process; and the review should be expanded or applied to other R01 reviews.

The next category was the *Negative* sentiment category in which 15 of 68 respondents provided negative comments about anonymized review. Of these respondents, nine of 15 expressed understanding of the value of anonymized review but felt the approach had flaws. Five respondents felt it reduced their effectiveness to not have more details in the application. Respondents grouped into the *Negative (other)* category indicated a dislike for anonymized review in general while some noted that it was easier to review. Other respondents commented that anonymized review is a waste of money and time, and is unscientific.

Seventeen responses from the 68 respondents did not align with the previous *other* categories. Fewer than five respondents

- expressed general curiosity about the TRA review process and the application, remarked that the process would be cumbersome if there were multiple applications to review, and noted that reviewing a single grant made it difficult to put it into perspective with others
- acknowledged that the application would not be anonymized if the applicant's track record was included, but that this information is needed to complete the review
- relied on their own knowledge in the field to review application
- had difficulty evaluating rigor or safety of experimental design (e.g., sample size, human subjects, children)
- would benefit from an example of an exceptional anonymized application to understand how they should score their own applications
- wanted to know

- the size of the lab to determine if the applicant had the resources to complete the proposed work
- if the lab group already had "too much money" and therefore did not need the research grant
- previous work to determine if the PI had the foundation for the proposed research

Respondents also commented that reviewers should be trusted to be impartial and the review process should not be anonymized. Several respondents noted that, for traditional NIH review, technical reviewers might participate in discussions for final scoring and they noted in their survey comments that they missed interacting with other reviewers to debate conclusions. Respondents indicated that the anonymized format may not work well for traditional R01 applications but works well for TRA applications.

Figure 38. Additional Information That Would Have Been Helpful in the Review

G. Phase III: Editorial Board Survey

Phase III has two parts. In Phase IIIa, the EB reviewed anonymized Specific Aims, Research Strategies, and technical review critiques for each application and provided a preliminary score (Figure 39). In Phase IIIb, EB members reviewed the complete

application, discussed each application as a group, and provided a final score. The Phase III survey queried EB members about their ability to identify the applicants; assess transformative potential; the level of difficulty of anonymized review versus traditional review; and how access to the complete application may have influenced final scoring (Figure 40). The survey questions are located in Appendix M and the data analysis in Appendix N.

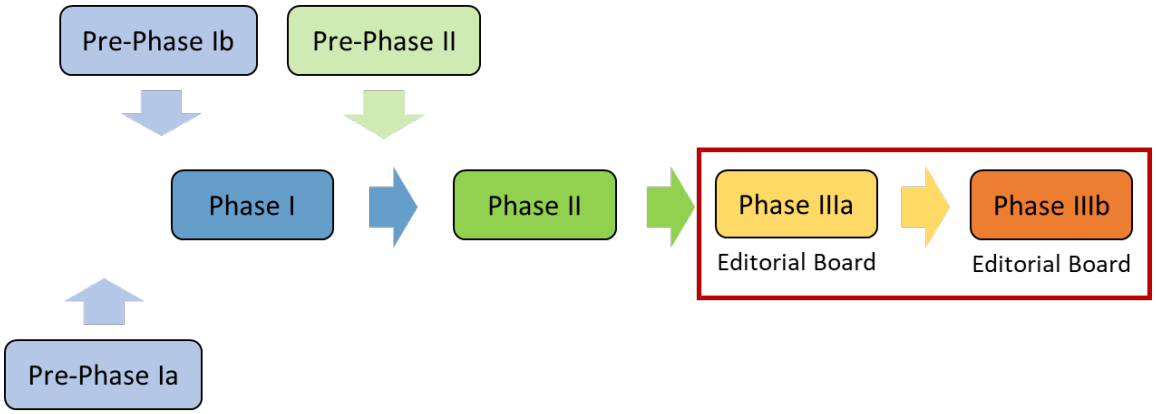


Figure 39. Phase III: Editorial Board Survey Flow Diagram

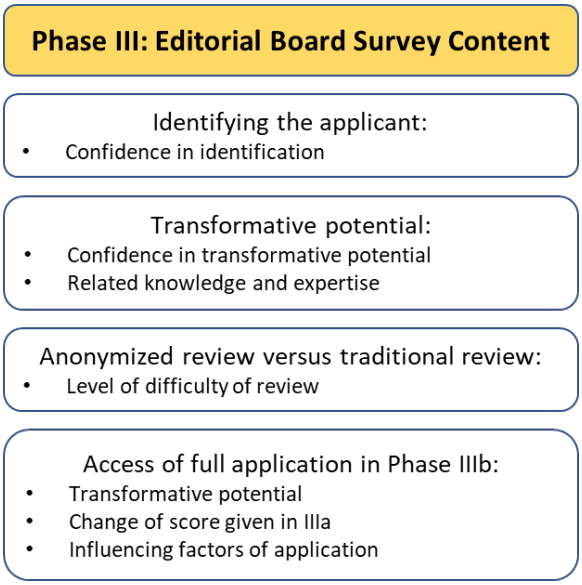


Figure 40. Phase III: Editorial Board Review Survey Content

1. Survey Administration

STPI sent an email invitation to participate in the survey to 25 EB members[15] on April 13, 2021 using a generalized link to the survey on the Alchemer platform. A reminder email was sent on April 19, and the survey closed on April 23. The data were downloaded from Alchemer on April 29.

2. Results

a. Response Rate

Twenty of the 25 EB members completed the Phase III EB survey. The response rate is 80%.

b. Survey Data

Question: Among the anonymized applications you reviewed in Phase IIIa of the review process, please indicate if there were any applications for which you believe you could identify the applicant or collaborator(s), lab group, or institution.

Nineteen of 19 respondents (100%) reported there were no applications for which they could identify the applicant or collaborator(s), lab group, or institution ($p < 0.001$).

Question: Please indicate, on average, how confident you were in your ability to determine if the anonymized applications you reviewed in Phase IIIa were transformative.

Seventeen of the 19 respondents (89.5%) indicated that they were *somewhat* or *very confident* in their ability to determine if the research in the anonymized application was transformative, with two selecting *not very confident* and no respondents selecting *not at all confident* (Figure 41).

[15] Two EB members did not participate in Phase III of the review and therefore were not contacted from the full list of 27 EB members.

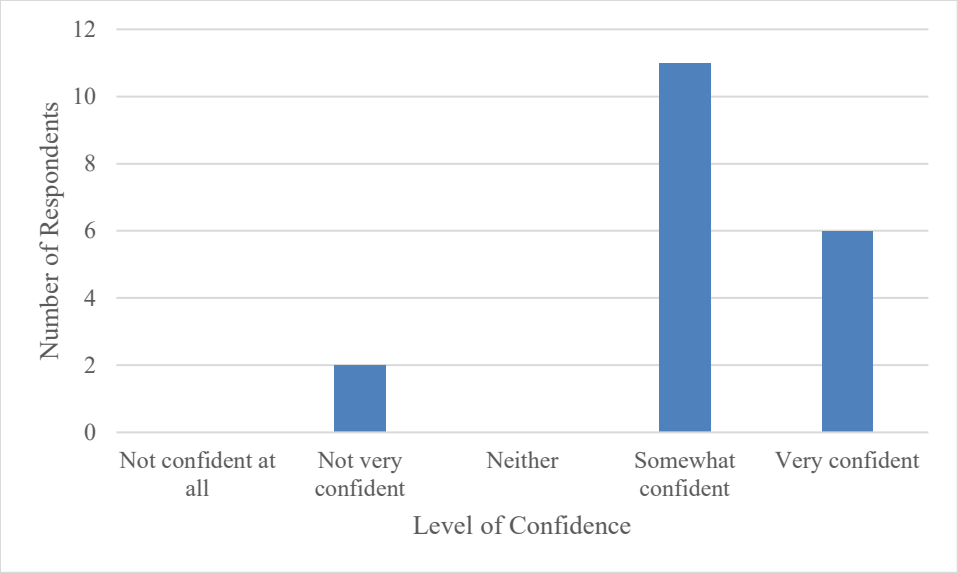


Figure 41. Reviewer Confidence in Ability to Assess Transformative Potential in Anonymized Applications

Question: You indicated that you were [response to previous question inserted here] your ability to evaluate the transformative potential of the proposed research. Please indicate if this is due to a lack of subject expertise.

One respondent indicated the low confidence was not due to a lack of related knowledge and expertise, while the other indicated that it was ($p = 1$).

Question: Please indicate how difficult or easy it was to evaluate anonymized applications in Phase IIIa compared to the traditional NIH review process in which you receive the full, deanonymized application.

Eight of 20 respondents (40%) indicated that anonymized applications in Phase IIIa were *more difficult* or *much more difficult* compared to the traditional NIH review process. Seven respondents (35%) felt that anonymized review was *about the same* in level of difficulty, and five respondents (25%) indicated that anonymized review was *easier* or *much easier* than NIH's traditional review (Figure 42).

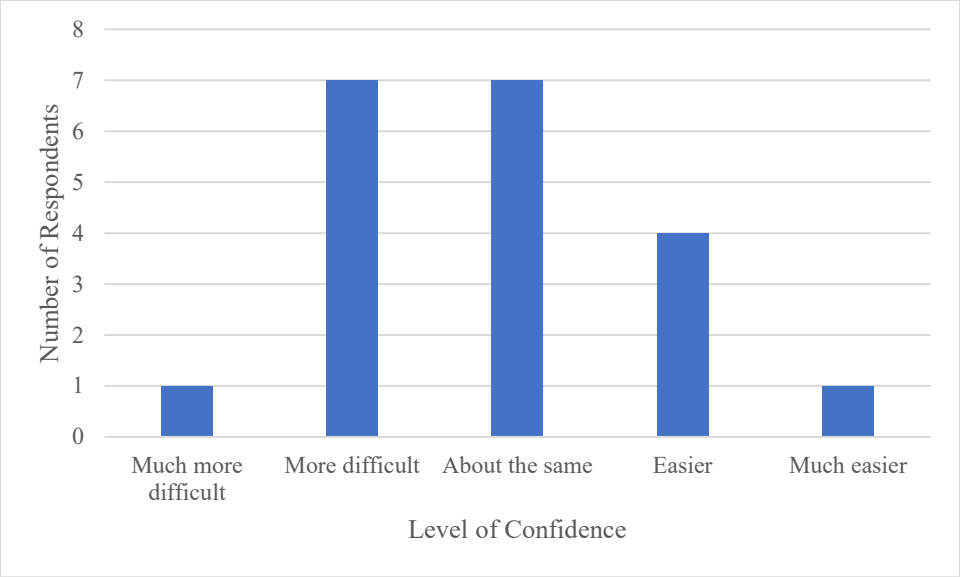


Figure 42. Survey Responses of Phase III Reviewers When Comparing the Difficulty of Anonymized Review to the Regular NIH Review Process

Respondents who indicated anonymized review was *easier* or *much easier* reported that anonymized review was more focused on the science and included less distracting information, or that the technical review notes were especially helpful.

The seven respondents who indicated the anonymized review was *more difficult* or *much more difficult* reported the following reasons: the review took more time; the assigned applications were outside of the reviewers' area of expertise; and did not know collaborators or resources (Figure 43).

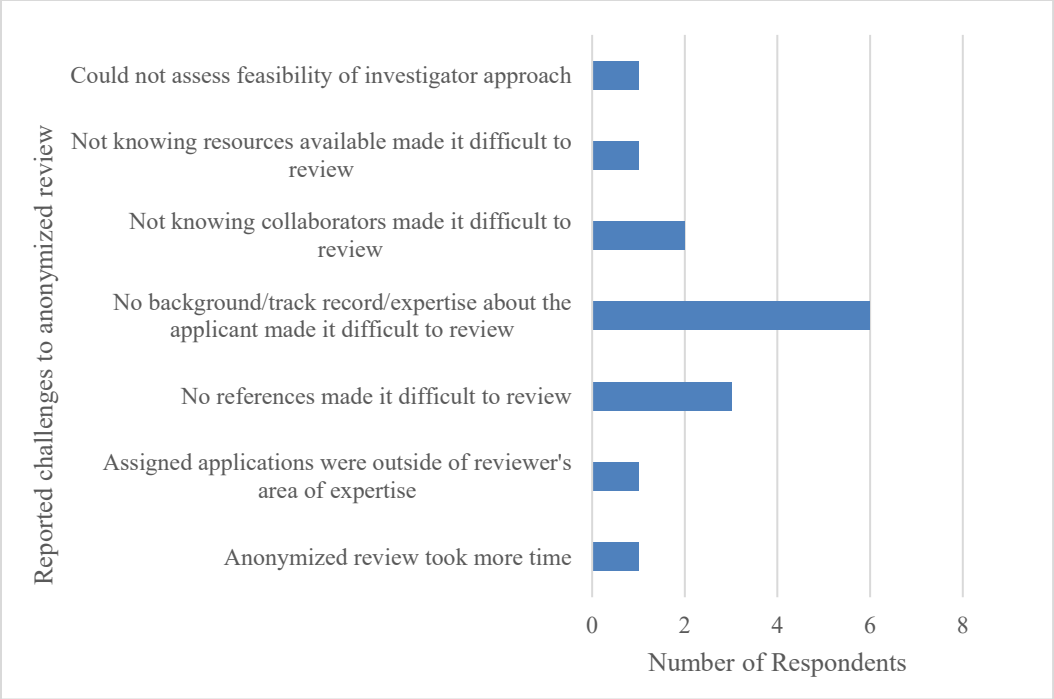


Figure 43. Challenges of Reviewers Who Felt Anonymized Review Was More or Much More Difficult than Traditional NIH Review

Question: Please indicate if you reviewed applications in Phase IIIa that were discussed in Phase IIIb.

Twenty of 20 respondents (100%) indicated that they reviewed applications in Phase IIIa that were discussed in Phase IIIb ($p < 0.001$).

Question: For those applications you reviewed in Phase IIIa that were then discussed in Phase IIIb, please indicate if access to the full, deanonymized application allowed you to better assess transformative potential.

Fourteen of 20 respondents (70%) felt that access to the full, deanonymized application allowed them to better assess transformative potential, and 6 respondents did not. ($p = 0.12$).

Twelve of the 14 respondents (86%) commented that the full, deanonymized application allowed them to better assess feasibility, with 9 respondents citing knowledge of the applicant's track record and expertise as the reason. Reliance on trust that the applicant could complete the proposed research was also cited.

Seven respondents provided answers that were grouped into the *Feasibility other* category of less than five comments per topic, with several using the word feasibility in their response. The respondents reported that, with the full deanonymized application they now knew that applicants had resources to complete the proposed research, had a better

idea of the research environment, and had collaborators who would contribute to the research.

Question: For those applications you reviewed in Phase IIIa that were then discussed in Phase IIIb, please indicate if you changed your final score on one or more of the applications following the discussion

Eighteen of 19 respondents (95%) indicated that they changed their score for 1 or more of the applications that were reviewed in Phase IIIa and then discussed in Phase IIIb ($p < 0.001$). One respondent did not change their score.

Question: Please indicate which of the following influenced you to change your score on one or more applications from Phase IIIa to Phase IIIb (select all that apply).

Eight of 19 respondents (42%) provided additional comments on their decision to change their score. Fourteen of 19 respondents attributed the score change to the discussion in Phase IIIb, and 3 of 19 respondents indicated they changed their score due to other reasons (Figure 44; $h's \quad Q \quad = 10.71, p \quad = 0.005$).

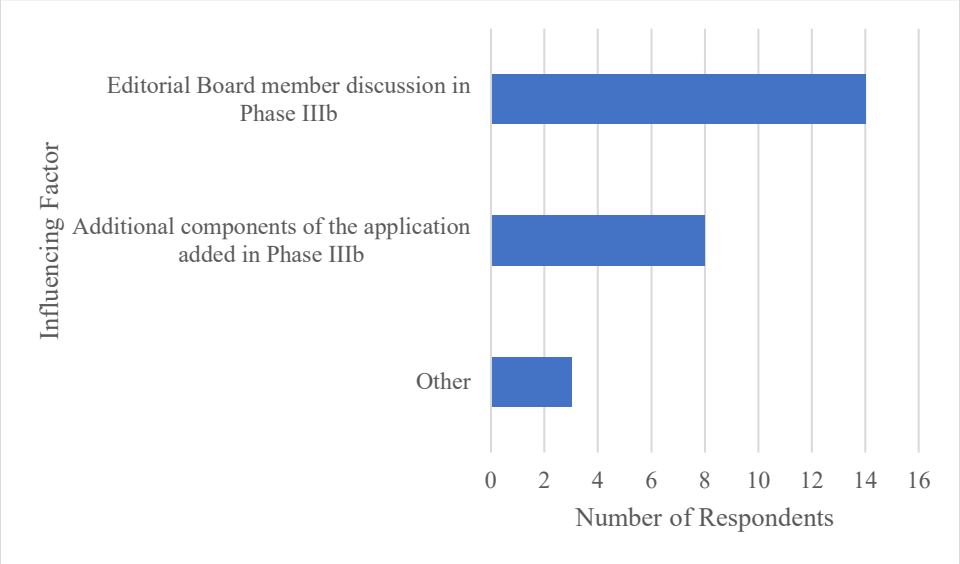


Figure 44. Influencing Factors Cited by Reviewers for Score Changes from Phase IIIa To Phase IIIb

When asked to further specify which additional components added in Phase IIIb influenced the reviewers' decisions to change their score, respondents cited knowledge of the environment, consortium/contractual arrangements, biosketches, bibliography and references, evidence of independence and institutional support, information about the investigators, the leadership plan, and letters from consultants and collaborators (Figure 45; $p \quad = 0$).

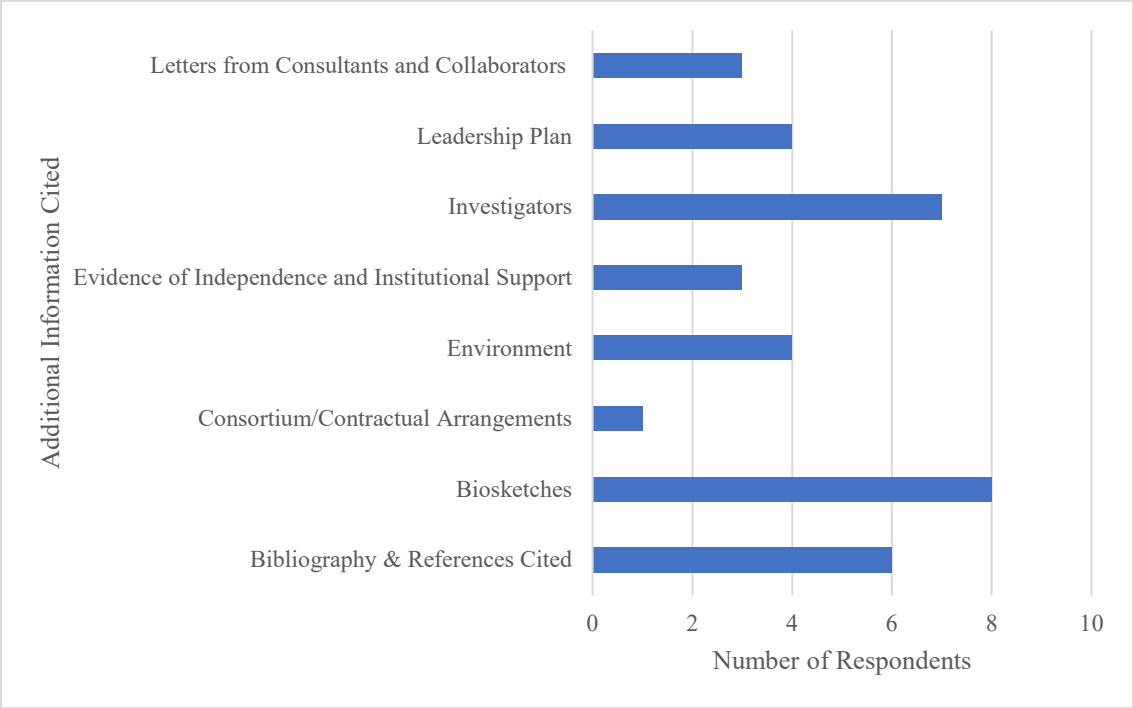


Figure 45. Additional Information That Assisted Reviewers in Phase IIIb

Also cited as rationale for changing preliminary scores were reviewer discussions, viewing other applications, knowing the applicant had access to special resources, and knowing the budget.

Question: Are there any additional comments you would like to provide to NIH about this final step of the review process?

Eleven of 20 survey respondents (55%) included additional comments about the review process. Five respondents indicated that anonymized review is good and should be continued, and all other comments were placed into the *other* category. Within this *other* category, some respondents remarked that identifying information is not needed for many applications to judge transformative potential, whereas others thought that applications should be deanonymized in Phase I to reduce time spent reviewing the same information again in Phase III. Others reported that anonymized review takes longer than a standard review and was awkward and reviewers should be trusted not to communicate with each other about the grants they are reviewing. Several positive comments noted a preference for anonymized review and that it would reduce bias.

Phase III respondents also noted the importance of technical reviews and that some respondents had to rely upon them when they did not have expertise in a research subject.

Although respondents noted reviewer discussion helped to provide a balanced view of applications score, and that changes made after Phase IIIb discussion were minor, others

reported that identifying information revealed in Phase IIIb changed the application scores dramatically.

3. Diversity Analysis

A. Introduction

This assessment of the TRA anonymized application and review process was preceded by many NIH diversity activities, including the ACD HRHR WG and the ACD WGD. To assess the impact of the TRA anonymized application and review process on diversity, STPI examined both demographic (i.e., gender, race, and ethnicity) as well as institutional changes among the TRA applicant and awardee pools through time (Figure 46).

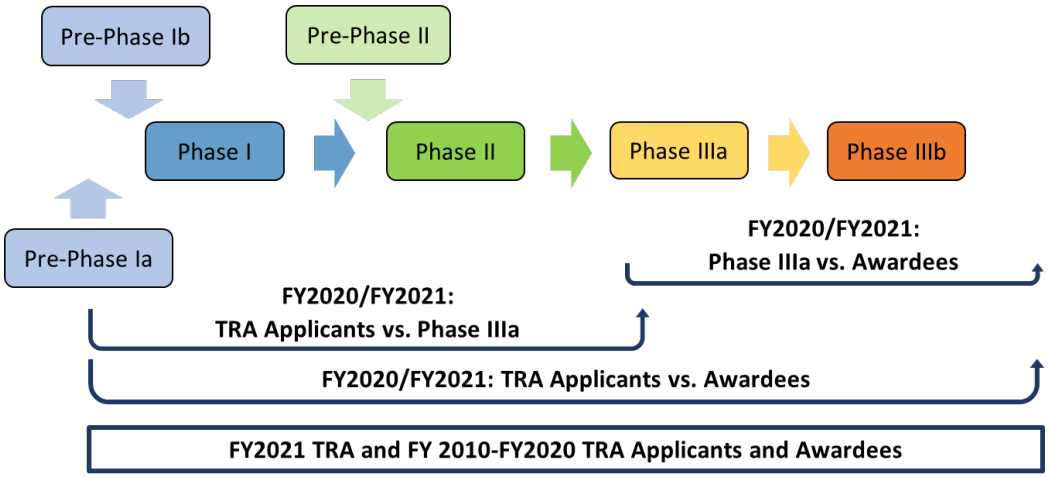


Figure 46. Diversity Analysis of TRA Applicants and Awardees Over Time

B. Methodology

Diversity in this report was assessed in two ways: demographic diversity in which gender, race, and ethnicity are considered; and institutional diversity in which the geographic location and name of institution are considered.[16] To assess whether the FY2021 anonymized application and review process increased demographic or

[16] Eighty percent of institutions observed during FY2010-FY2020 were categorized by NIH as institutions of higher education; 9% were research organizations; 6% were independent hospitals; 1% were other health, human resources, or environment/community services; and 4% were *other* institutions. Because institutions of higher education accounted for four-fifths of all institution types observed, the data were not discernible and institution type was not considered in any further analyses.

institutional diversity, FY2021 TRA applicants and awardees were compared to previous TRA applicants and awardees (i.e., FY2010–FY2020).[17]

STPI applied for and obtained demographic data of the FY2010–FY2021 TRA applicants and preliminary and final score data for the FY2020 and FY2021 TRA cohorts through the CSR, Office of Planning, Analysis and Evaluation. Institutional data were obtained from NIH's QVR system.

1. Demographic Diversity

Demographic diversity in this report is confined specifically to gender, race, and ethnicity. For gender, individuals were categorized based on their self-identification from their NIH TRA application as *female*, *male*, or *Other* (i.e., those who selected *withheld* or *unknown*). For race, individuals were similarly categorized as *Asian*, *Black or African American*, *White*, or *Other* (i.e., those who selected *American Indian or Alaska Native; Native Hawaiian or other Pacific Islander; Person reporting more than 1 race; unknown;* or *withheld*). For ethnicity, individuals were categorized as *Hispanic or Latino*, *not Hispanic or not Latino*, or *Other* (i.e., those who selected *unknown* or *withheld*).

a. FY2010–FY2020 TRA Applicants

To assess the demographic diversity of FY2010–FY2020 TRA applicants, only Type 1 applications were included in the analysis. Because TRA applications allow for multiple PIs, all PIs (i.e., contact PI and co-PIs) are included in these analyses. In addition, because an individual may submit more than one application within a TRA FOA and may also submit applications other TRA FOAs offered in the same fiscal year, STPI defines *applicant* as the number of instances in which an individual applied to a TRA FOA. In other words, the demographic characteristics (i.e., gender, race, and ethnicity) of an individual is represented once for each application that individual is listed as a contact PI or co-PI.

Linear regression models were used to assess if the percentage of applicants and awardees changed through time for each group within gender, race, and ethnicity. A type-II sum of squares was used to determine if rate of change through time (i.e., slope) was significantly different from zero.

Two-sample proportion tests were used to determine whether the percentage of male applicants and awardees differed from female applicants and awardees, respectively. Similarly, two-sample proportion tests were used to determine whether the percentage of Hispanic or Latino applicants and awardees differed from non-Hispanic or Latino

[17] FY2009 TRA applicants and awardees were not included in any analysis because instructions in the FOA changed significantly after the first year of the TRA program.

applicants and awardees, respectively. To assess how the percentage of Asian, Black or African American, and White applicants and awardees differ from one another, pairwise comparison proportion tests were performed. As a reminder, because pairwise comparisons increase the likelihood of a Type I error (i.e., rejecting the null hypothesis when it should not be rejected), the p-values need to be corrected so that a false positive is not made. To account for this, the p-values for the pairwise proportion tests were adjusted using the Bonferroni correction method. The Other categories for gender, race, and ethnicity were not included in the two-sample proportion tests or the pairwise comparison proportion tests because the sample sizes for the Other categories are small, as is the case for gender and ethnicity; and it would increase the likelihood of a Type I error if included in the analyses. Excluding the Other categories from these analyses allows the focus to be on assessing how the main groups within gender, race, and ethnicity differ from one another.

b. Comparison of FY2021 with Baseline (FY2010–2020) TRA Applicants

To assess whether demographic diversity increased in the FY2021 TRA applicant cohort relative to the baseline (i.e., FY2010–2020) TRA applicant population, one-sided, one-sample proportion tests were performed for each group within gender, race, and ethnicity. The FY2021 proportion of each group is compared to the known proportion of each group calculated from the FY2010–2020 TRA applicant data. To test whether diversity increased for gender in FY2021, the one-sided alternative hypotheses is that the FY2021 proportions for females and Others are higher than their respective FY2010–2020 known proportions, and the FY2021 proportion for males is lower than the FY2010–2020 known proportion for males. To test whether diversity increased for race in FY2021, the one-sided alternative hypotheses is that the FY2021 proportions for Asians, Black or African Americans are higher than their respective FY2010–2020 known proportions, and the FY2021 proportion for Whites is lower than the FY2010–2020 known proportion for Whites. To test whether diversity increased for ethnicity in FY2021, the one-sided alternative hypotheses is that the FY2021 proportions for Hispanic or Latinos and others are higher than their respective FY2010–2020 known proportions, and the FY2021 proportion for not Hispanic or not Latinos is lower than the FY2010–2020 known proportion for not Hispanic or not Latinos.

A logistic regression followed by a type-II sum of squares analysis of deviance was used to assess whether individuals' funding status (dependent variable in the regression where 1 indicated an individual with a funded application and 0 indicated an individual with an unfunded application) was significantly influenced by their gender or race. Ethnicity was not included in the logistic regression as it is highly correlated with race. In addition, STPI also performed Fisher's exact test to assess whether there was a significant relationship (i.e., difference) in the breakdown of FY2021 applicants and awardees by gender, race, and ethnicity.

To assess whether demographic diversity increased in the FY2021 TRA awardee pool relative to the baseline (i.e., FY2010–2020) TRA awardee pool, one-sided, one-sample proportion tests were performed for each group within gender, race, and ethnicity as described in the previous paragraph.

c. Comparison of Preliminary and Final Scores between FY2020 and FY2021 TRA Applicants

STPI assessed the average (± SE) preliminary and final scores for FY2020 and FY2021 TRA applications by gender, race, and ethnicity.

To compare changes in final scores between FY2020 and FY2021, STPI used a generalized linear model (GLM) with a Poisson distribution with the final score as the response variable to assess whether final scores differed by gender, race, or ethnicity for FY2020 and FY2021 applicants who made it to Phase III. A type-II sum of squares was used to determine whether gender, race, or ethnicity were significant factors in influencing the final scores received by applicants in Phase III.

STPI also calculated the absolute difference in final scores between each group in gender (i.e., females and males; females and Others; males and Others), race (i.e., Asians and Black or African Americans; Asians and Whites; Asians and Others; Black or African Americans and Whites; Black or African Americans and Others; and Whites and Others), and ethnicity (i.e., Hispanic or Latino and not Hispanic or not Latino; Hispanic or Latino and Others; not Hispanic or not Latino and Others) for FY2020 and FY2021. These multiple comparisons within each of the three demographic factors are the final score gap. STPI then assessed how the final score gap for each comparison changed between FY2021 and FY2020 by subtracting the FY2020 gap value from the FY2021 gap value to determine if the gap decreased, stayed the same, or increased between the 2 years. For instance, if the final score gap between female and male applicants was 0.02 points in FY2020 and 0.28 points in FY2021, then the gap between female and male applicants increased 0.26 from FY2020 to FY2021. No statistical analyses were performed to determine if the changes in final score gaps are significantly different from zero because with only FY2020 and FY2021 data available, there are no repeated measures, and this analysis provides only a snapshot in time.

d. Stratification of FY2020 TRA Applicants

STPI considered changes in demographic diversity for the following comparisons: applicant pool versus individuals who made it to Phase III; individuals who made it to Phase III versus funded awardees (i.e., awardee pool); and the applicant pool versus the awardee pool. Proportional differences for each comparison are calculated for each group within gender, race, and ethnicity using a two-sample proportion test.

A power analysis for a two-sample proportion test with unequal sample sizes was performed for each group within gender, race, and ethnicity to determine the amount of power (i.e., the probability that the null hypothesis will be rejected when it should) present to detect statistically significant differences. As a reminder, it is generally accepted that power should be 0.80 or greater. In other words, the probability of detecting a statistically significant difference when one actually exists is 80%. The power analysis was performed using the *pwr* package in R[18] where the effect size was calculated using the arcsine transformation for proportions, and the significance level was 0.05. To comply with the Data Use Agreement signed by STPI and NIH regarding the use of applicant Personally Identifiable Information (PII) (i.e., race, gender, and ethnicity) and to prevent the identification of individuals through PII, STPI withheld any data where the number of individuals presented by demographic factors is less than 11. Cases in which the number of individuals can be calculated are also withheld.

e. Stratification of FY2021 TRA Applicants

STPI considered changes in demographic diversity for the following comparisons: applicant pool versus individuals who made it to Phase III; individuals who made it to Phase III versus funded awardees (i.e., awardee pool); and the applicant pool versus the awardee pool. Proportional differences for each comparison are calculated for each group within gender, race, and ethnicity using a two-sample proportion test.

Power analyses for two-sample proportion tests with unequal sample sizes were performed for each group within gender, race, and ethnicity to determine the amount of power present to detect statistically significant differences.

f. Simpson's Diversity Index for Demographic Diversity

STPI assessed the changes in demographic diversity over time using Simpson's Diversity Index (commonly referred to as D). Commonly used in ecology to measure community diversity, we applied the same concept to measure demographic diversity for gender, race, and ethnicity through time. Simpson's index is calculated using the following equation:

$$D = \frac{1}{\sum_{i=1}^{s} p_i^2} \qquad \text{eq. 1}$$

where p is the proportion of individuals observed for one particular species divided by the total number of individuals found. The complement of the index:

$$1 - D \qquad \text{eq. 2}$$

[18] See 2020 "Package 'pwr'.", available at: https://github.com/heliosdrm/pwr.

is what is commonly used and also referred to as Simpson's Diversity Index because of the ease in its interpretation. From here on out, all references to Simpson's Diversity Index refers to the formula specified in equation 2. Simpson's Diversity Index ranges from 0 to 1 where low indices (i.e., closer to 0) indicate low diversity and high indices (i.e., closer to 1) indicate high diversity. For instance, an index of 0.85 in a given fiscal year for, say gender, has higher gender diversity than a fiscal year with an index of 0.65. The SDI is a dominance index in that it gives more weight to common or dominant species (or in this case, demographic factors). What this means is that each group within gender, race, or ethnicity with a smaller number of TRA applicants in a given year has less impact on the overall diversity (for that given year and demographic factor) than a group with a larger number of TRA applicants in the same given year. All TRA applicants between FY2010–2021 were used to calculate Simpson's Diversity Index for each demographic factor and fiscal year. The index was calculated using the *diversity* function in the *vegan* package in R.[19]

2. Geographic and Institutional Diversity

Only institutions associated with the contact PI of an application are used in the institutional diversity analyses because the name of the institution associated with an application is that of the contact PI, not the co-PI. Geographic diversity was derived from the location of the contact PI's institution.

a. FY2010–2020 TRA Cohorts

1) Geographic Diversity

STPI defined geographic diversity to be at the State level. Specifically, STPI analyzed which States had one or more contact PIs from TRA applications for each year during FY2010–2020, and which States had one or more TRA awardees for the same time period. For these analyses, STPI only considered contact PIs and awardees of TRA applications that were affiliated with institutions located in the United States (i.e., the 50 States and the District of Columbia). Contact PIs and awardees of TRA applications from outside the United States were excluded from these analyses.

To assess how geographic diversity for contact PIs of TRA applications changed during FY2010–2020, a normal linear regression model was used where the dependent variable was the number of States (along with the District of Columbia) had at least one or more contact PIs and the independent variable was the fiscal year. A type-II sum of squares

[19] See vegan: Community Ecology Package. R package version 2.5-6, available at https://CRAN.R-project.org/package=vegan

was used to determine whether the independent variable was a significant factor in influencing the dependent variable. Similarly, a normal linear regression model with type-II sum of squares was used to assess how geographic diversity for TRA awardees changed during FY2010–2020.

2) Institutional Diversity

STPI assessed the number of unique institutions that were affiliated with one or more contact PIs each year between FY2010–2020. A normal linear regression model was used to assess how the number of total institutions with contact PIs changed over time. A type-II sum of squares was used to determine whether time (i.e., fiscal year) influenced the number of institutions with contact PIs. Similarly, STPI calculated the number of unique institutions that were affiliated with one or more TRA awardees each year between FY2010–2020.

STPI also assessed the number of new institutions with contact PIs between FY2010–2020. A new institution was defined as one without a previous contact PI for each of the previous years until the fiscal year when a contact PI was affiliated with that institution. The number of new institutions was not calculated for FY2010 because STPI was instructed not to consider applications from FY2009 due to differences in TRA FOA directions. STPI also calculated the success rate for institutions by taking the number of unique institutions with at least one TRA awardee between FY2010–2020 and dividing it by the total number of unique institutions that had at least one contact PI during the same time period.

STPI also considered institutional diversity with regard to Historically Black Colleges and Universities (HBCUs). An institution was considered an HBCU if it matched to one of the 102 HBCUs identified by the National Center for Education Statistics as having a specialized mission of being a historically black college or university.[20] The number of HBCUs with at least one TRA applicant between FY2010–2020 was calculated, as was the number of HBCUs with at least one TRA awardee during the same time period.

Lastly, STPI calculated institutional diversity using Simpson's Diversity Index. All institutions observed between FY2010–2021 were used to calculate Simpson's Diversity Index for each fiscal year. All institutions located in U.S. and U.S. territories were used in institutional diversity analyses.

[20] See 2021 National Center for Education Statistics "College Navigator.", available at: https://nces.ed.gov/COLLEGENAVIGATOR/?s=all&sp=4&pg=1.

b. Comparison of FY2021 with Baseline (FY2010–2020) TRA Cohorts

To determine whether geographic diversity increased in FY2021, STPI compared the States represented by the FY2021 TRA contact PIs against the States represented by the FY2010–2020 TRA contact PIs. STPI also calculated the number of new institutions observed in FY2021 and compared this to the number of new institutions observed in previous years. Lastly, the Simpson's Diversity Index was calculated for FY2021 and this was compared to the diversity indices observed in previous years.

C. Results

1. Demographic Diversity

a. FY2010–2020 TRA Applicants

A total of 3,381 TRA applications were received from 4,281 unique individuals between FY2010–2020. Because individuals may submit more than one application in response to a TRA FOA as well as submit applications across multiple years, STPI counted each instance when an individual submitted a TRA application. Using this calculation, there were 5,428 applicants during FY2010–2020. Overall, 134 of the 3,381 applications (4.0%), and 216 of the 5,428 (4.0%) applicants were funded.[21]

1) Gender

Of the 5,428 applicants, 1,107 were female (20.4%), 4,132 were male (76.1%), and 189 were those who were categorized as Other (3.5%). The overall percentage of applicants who were male was significantly higher than the percentage of applicants who were female ($\chi^2_1 = 3{,}374,\ p < 0.001$). While the number of female, male, and Other applicants decreased significantly over time ($p < 0.001$ for all), the proportion of applicants by gender over time has remained constant ($p = 0.28$ for females, $p = 0.20$ for males, and $p = 0.32$ for Others; Figure 47). Specifically, the average (± SE) percentage of applicants who were female is 20.7 (± 0.7)%; 75.7 (± 0.9)% were male; and 3.4 (± 0.7)% were those who were categorized as Other.

[21] Of the 216 applicants who were funded, there were 208 unique individuals.

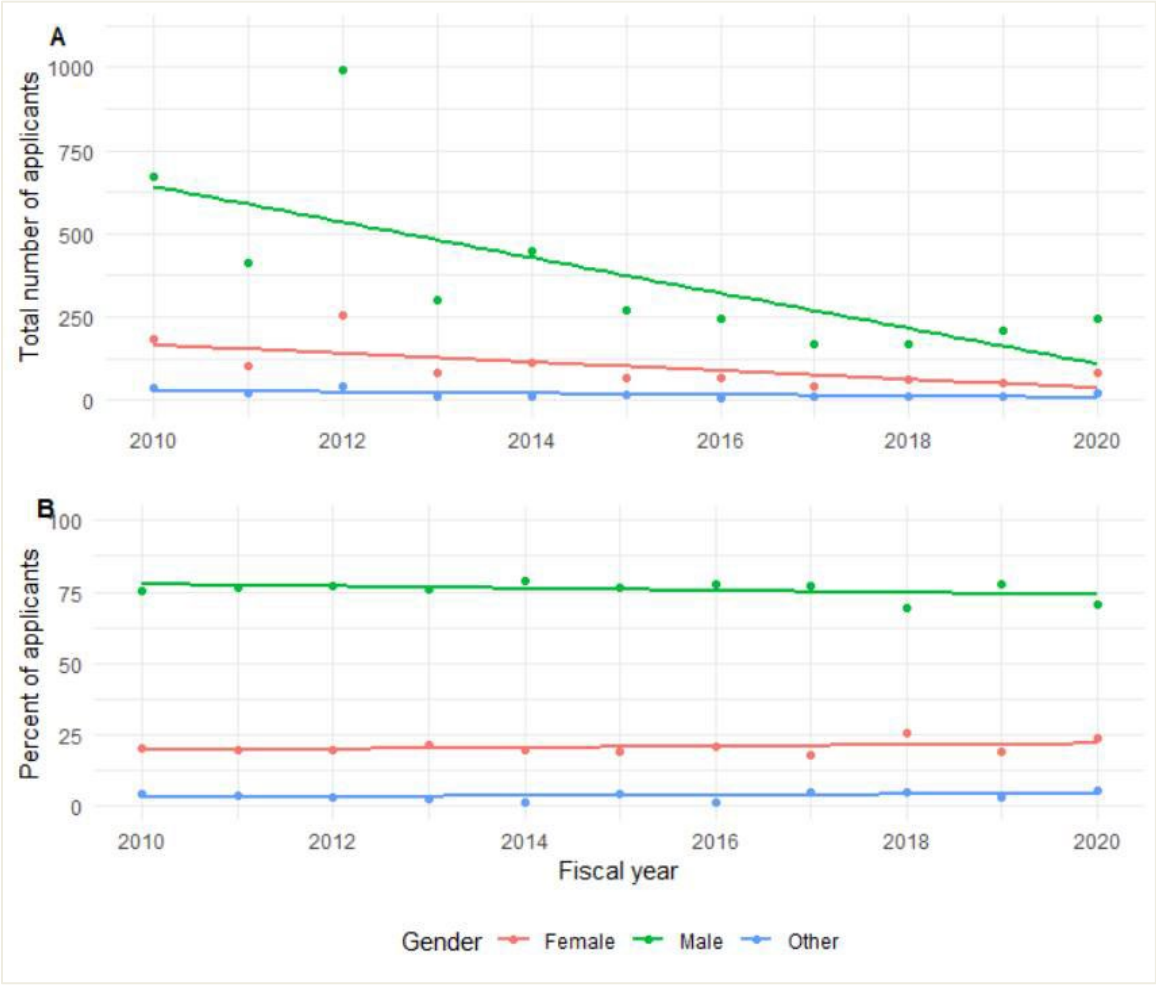


Figure 47. (A) Total Number of TRA Applicants by Gender from FY2010 to FY2020 (B) Percentage of Applicants by Gender through the Same Time Period

A total of 216 awardees were funded between FY2010 and FY2020. We do not provide the breakdown of awardees by gender because the number of individuals who were categorized as Other was less than 11. The overall percentage of awardees who were male was significantly higher than the percentage of awardees who were female ($\chi^2_1 = 136, < 0.001$). The number of awardees who are male has decreased significantly over time (< 0.001) but the number of awardees who are female or are categorized as Other did not change significantly over time ($= 1.0$ and $= 0.43$ for females and Others, respectively; Figure 48). The proportion of awardees has remained constant over time for each gender ($= 0.17$ for females, $= 0.14$ for males, and $= 0.62$ for Others; Figure 48). Specifically, the average (± SE) percent of awardees who were female is 23.2 (± 2.7)%; 77.0 (± 2.7)% were male; and 5.2 (± 1.2)% were those who were categorized as Other.

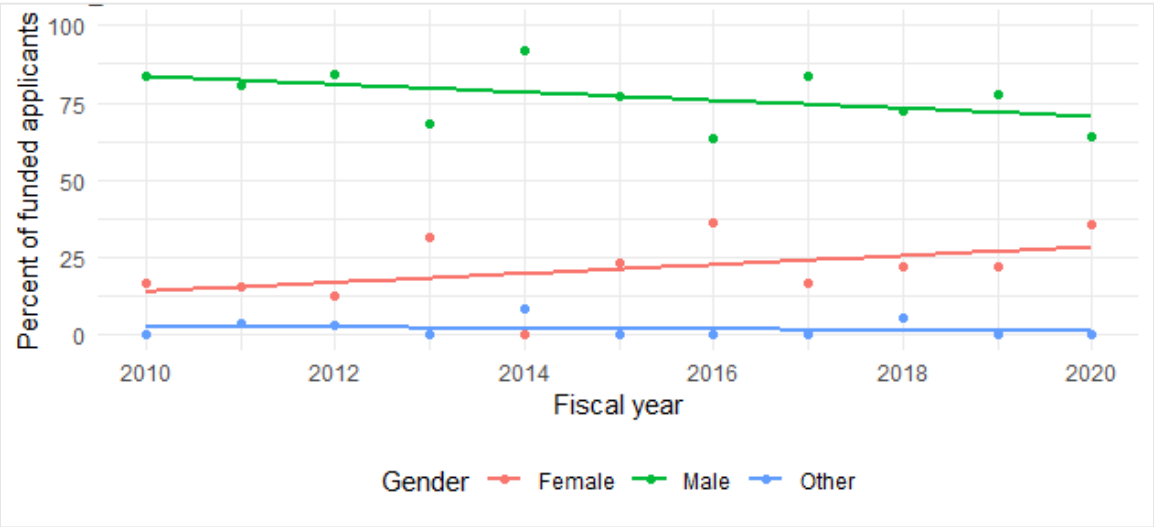


Figure 48. Percentage of Funded Applicants by Gender from FY2010 to FY2020

Comparing the percentage of applicants and awardees by gender from FY2010 to FY2020, the percentage of awardees who are female is not significantly different from the percentage of applicants who are female ($\chi_1^2 = 0.01, p = 0.94$; Figure 48). Similarly, the percentage of awardees who are male is not significantly different from the percentage of applicants who are male ($\chi_1^2 = 0.10, p = 0.75$); and the percentage of awardees who are categorized as Other is not significantly different from the percentage of applicants who are categorized as Other ($\chi_1^2 = 1.21, p = 0.27$).

The overall success rate among female applicants is 4.1%, 4.0% among male applicants, and 2.1% among applicants who were categorized as Other. The success rate among males does not differ significantly from that of females ($\chi_1^2 < 0.001, p = 1.0$). The average (± SE) success rate among female applicants is 5.2 (± 1.0)%; 4.7 (± 0.5)% for male applicants; and 2.4 (± 1.2)% for applicants who were categorized as Other.

2) Race

Of the 5,428 applicants from FY2010 to FY2020, 1,223 were Asian (22.5%), 75 were Black or African American (1.4%), 3,388 were White (62.4%), and 742 were those who were categorized as Other (13.7%). The overall percentage of White applicants was significantly higher than that of all other races, followed by Asian applicants, and Black or African American applicants ($p < 0.001$ for all pairwise comparisons). The number of applicants has significantly decreased over time for Asian applicants ($p < 0.001$), Black or African American applicants ($p < 0.001$), White applicants ($p < 0.001$), and Other applicants ($p < 0.01$) but proportionally has maintained constant through time for Asian ($p = 0.06$), Black or African American ($p = 0.95$), and White applicants ($p = 0.35$;

Figure 49) applicants. The percentage of applicants who were categorized as Other has experienced a small but significant decrease of 0.003 percentage points for each year between FY2010 and FY2020 ($p = 0.04$; Figure 49). Specifically, the average (± SE) percentage of applicants who were Asian is 23.8 (± 1.0)%; 1.8 (± 0.1)% were Black or African American; 61.5 (± 0.9)% were White; and 13.5 (± 0.4)% were those categorized as Other.

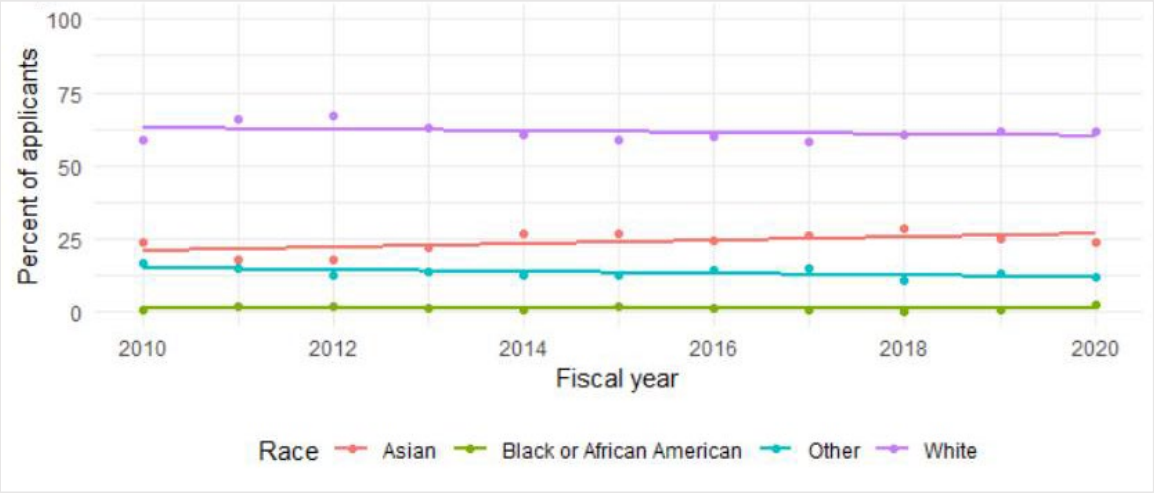


Figure 49. Percentage of Applicants by Race from FY2010 to FY2020

Of the 216 awardees, 34 were Asian (15.7%), 154 were White (71.3%), and 28 were either Black or African American or those who were categorized as Other (13.0%). The percentage of White awardees was significantly higher than the percentage of Asian awardees as well as the percentage of Black or African American awardees ($p < 0.001$ for both); and the percentage of Asian awardees is significantly higher than the percentage of Black or African American awardees ($p < 0.001$). The number of funded applicants who were Asian ($p = 0.11$), Black or African American ($p = 0.09$), and Other ($p = 0.21$) did not change significantly over time but the number of funded applicants who were White decreased significantly over time ($p < 0.01$; Figure 50). The proportion of funded applicants by race has remained constant over time ($p = 0.65$ for Asian awardees, $p = 0.10$ for Black or African American awardees, $p = 0.46$ for White awardees, and $p = 0.94$ for Other awardees; Figure 50). Specifically, the average (± SE) percentage of funded applicants who were Asian is 15.9 (± 1.8)%; 3.5 (± 0.4)% were Black or African American; 72.3 (± 2.3)% were White; and 12.3 (± 2.2)% were Other.

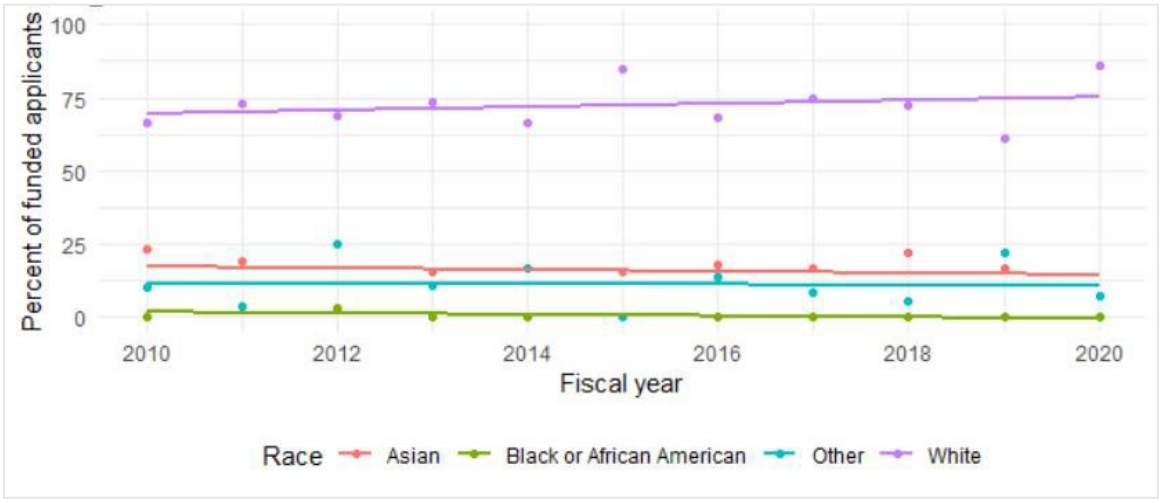


Figure 50. Percentage of Awardees by Race from FY2010 to FY2020

Comparing the percentage of applicants and awardees by race from FY2010 to FY2020, Asian applicants were funded at a significantly lower rate (15.7%) than that at which they apply (22.5%; $\chi_1^2 = 5.15, p = 0.02$); and White applicants were funded at a significantly higher percentage (71.3%) than that at which they apply (62.4%; $\chi_1^2 = 6.63, p = 0.01$; Figure 50). There was no significant difference between percentage of applicants and awardees for Black or African American applicants (1.4% and 0.9%, respectively; $\chi_1^2 = 0.07, p = 0.79$) or those who were categorized as Other (13.7% and 12.0%, respectively) ($\chi_1^2 = 0.34, p = 0.56$).

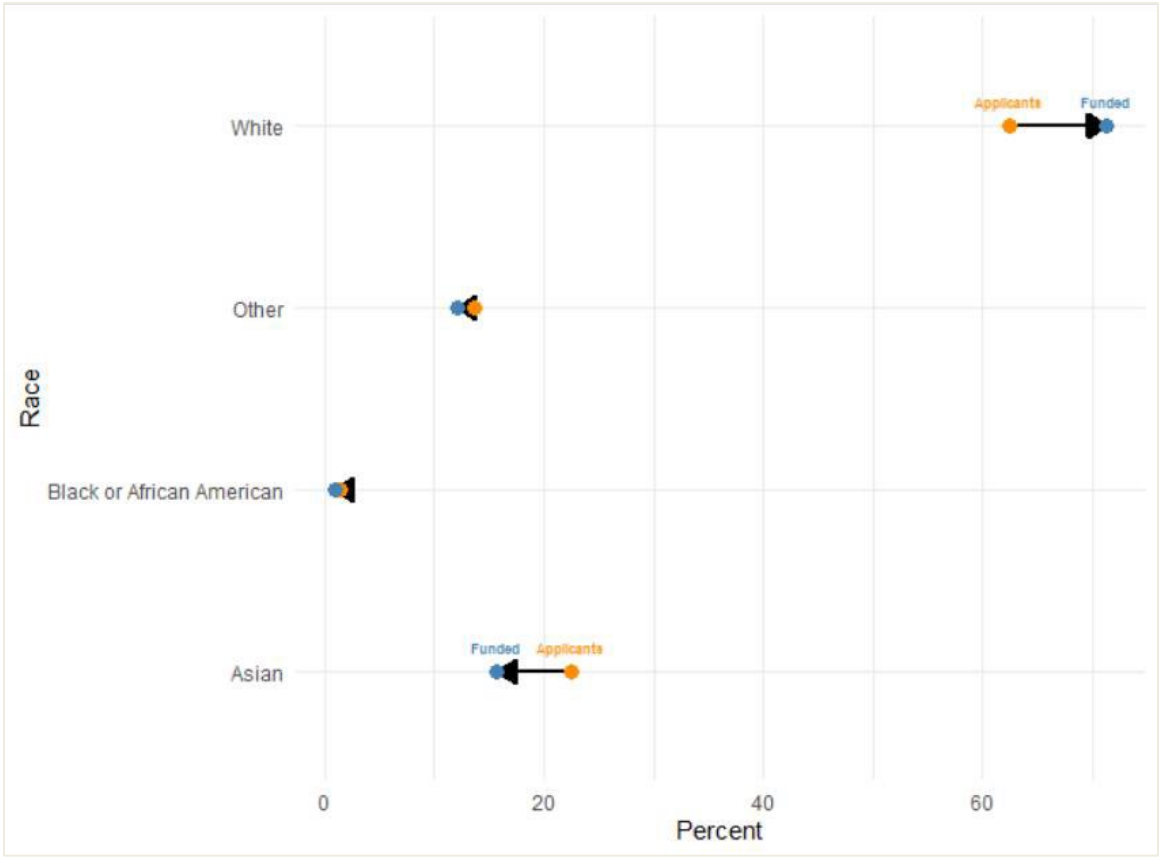


Figure 51. Percentage of Total (Orange) and Funded Applicants (Blue) by Race from FY2010 to FY2020

The overall success rate among Asian applicants from FY2010 to FY2020 is 2.8%, 2.7% among Black or African American applicants, 4.5% among White applicants, and 3.5% among applicants who were categorized as Other. Success rate did not differ between Asian and Black or African American applicants ($p = 1.0$), or between Black or African American and White applicants ($p = 1.0$) (Figure 51). The success rate of Asian applicants, however, was significantly lower than that of White applicants ($p = 0.03$). The average (± SE) success rate among Asian applicants is 3.3 (± 0.5)%; 1.4 (± 1.0)% for Black or African American applicants; 5.5 (± 0.6)% for White applicants; and 3.8 (± 0.9)% for applicants who were categorized as Other.

3) Ethnicity

Of the 5,428 applicants from FY2010 to FY2020, 156 were Hispanic or Latino (2.9%), 4,318 were not Hispanic or not Latino (79.6%), and 954 were those who were categorized as Other (17.6%). The overall percentage of not Hispanic or not Latino applicants was significantly higher than that of Hispanic or Latino applicants ($\chi_1^2 = 6{,}583, p < 0.001$). The number of Hispanic or Latino applicants has decreased significantly over time ($p < 0.001$) but proportionally, has remained constant ($p = 0.56$;

Figure 52). The number of not Hispanic or not Latino applicants decreased significantly over time ($p < 0.001$) while the proportion of not Hispanic or not Latino applicants observed a small but significant increase of 0.01 percentage points each year between FY2010 and FY2020 ($p < 0.01$). The number of applicants who were categorized as Other also decreased significantly over time ($p < 0.001$) while the proportion of applicants who were categorized as Other had a small but significant decrease of 0.01 percentage points each year between FY2010 and FY2020 ($p < 0.001$). On average (± SE), 3.2 (± 0.5)% of applicants were not Hispanic or not Latino; 80.8 (± 1.2)% were Hispanic or Latino; and 16.4 (± 1.3)% were those who were categorized as Other .

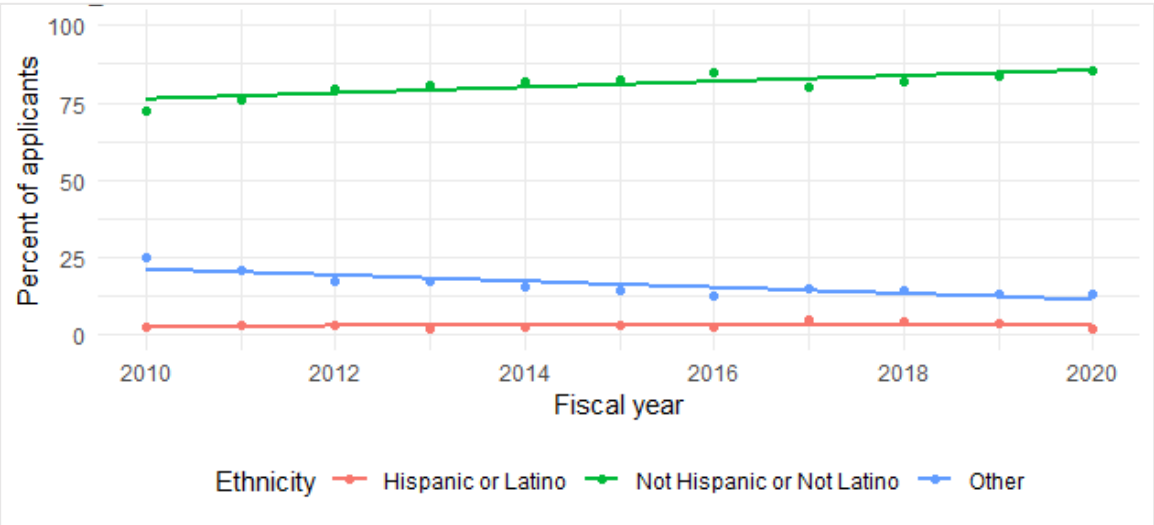


Figure 52. Percentage of Applicants by Ethnicity from FY2010 to FY2020

Of the 216 awardees from FY2010 to FY2020, 180 were not Hispanic or not Latino (83.3%), and 36 were either Hispanic or Latino or those who were categorized as Other (16.7%). The overall percentage of not Hispanic or not Latino awardees was significantly higher than the percentage of Hispanic or Latino awardees ($\chi_1^2 = 275, p < 0.001$). The number of funded applicants who are Hispanic or Latino did not change significantly over time ($p = 1.0$) while the number of funded applicants who are not Hispanic or not Latino and those who are Other decreased significantly over time ($p < 0.01$ and $p = 0.03$, respectively; Figure 52). The proportion of funded applicants has remained approximately constant by ethnicity ($p = 0.60$ for Hispanic or Latinos, $p = 0.59$ for not Hispanic or not Latinos, and $p = 0.23$ for Others; Figure 53). Specifically, the average (± SE) percentage of funded applicants who were Hispanic or Latino is 9.5 (± 2.0)%; 83.8 (± 2.1)% of funded applicants were not Hispanic or not Latino; and 13.1 (± 1.4)% of funded applicants were those who were categorized as Other .

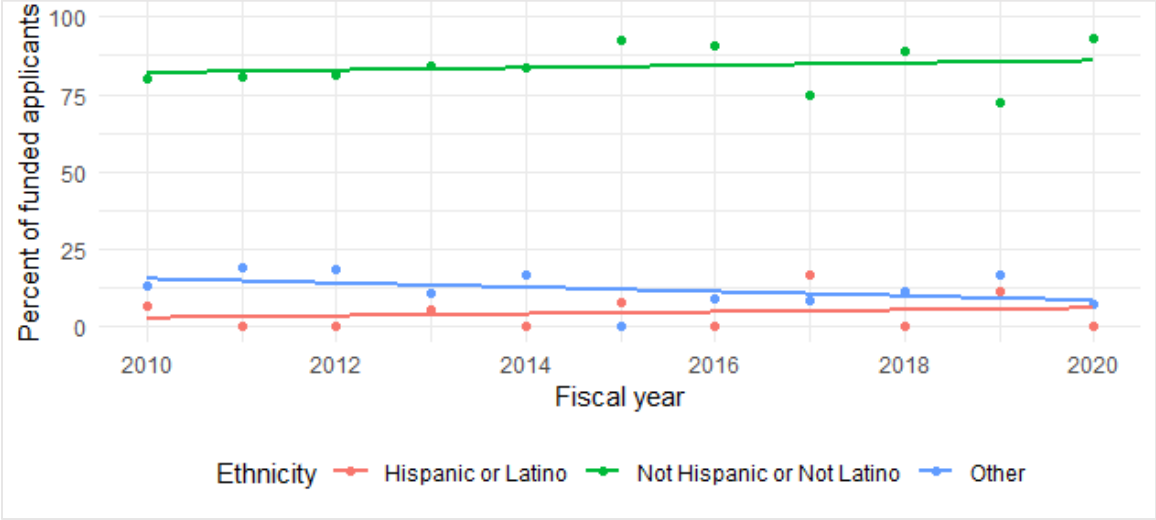


Figure 53. Percent of Funded Applicants by Ethnicity from FY2010 to FY2020

There was no significant difference between the percentage of applicants who are Hispanic or Latino (2.9%) and the percentage of awardees who are Hispanic or Latino (3.7%; $\chi_1^2 = 0.26, p = 0.61$) (Figure 53). Similarly, there were no significant differences between the percentage of applicants who are not Hispanic or not Latino (79.6%) and the percentage of awardees who are not Hispanic or not Latino (83.3%; $\chi_1^2 = 1.61, p = 0.20$), nor between the percentage of applicants who were categorized as Other (17.6%) and the percentage of awardees who were categorized as Other (13.0%; $\chi_1^2 = 2.77, p = 0.10$).

From FY2010 to FY2020, the overall success rate is 5.1% among Hispanic or Latino applicants, 4.2% among not Hispanic or not Latino applicants, and 2.9% among applicants who were categorized as Other. Success rate did not differ between Hispanic or Latino and not Hispanic or not Latino applicants ($\chi_1^2 = 0.15, p = 0.70$). The average (± SE) success rate among Hispanic or Latino applicants is 6.5 (± 2.5)%; 4.9 (± 0.5)% for not Hispanic or not Latino applicants; and 3.5 (± 0.7)% for applicants who were categorized as Other.

b. Comparison of FY2021 with Baseline (FY2010–2020) TRA Applicants

1) Gender

A total of 295 applicants applied to the general TRA FOA in FY2021. Of the 295 applicants, 73 (24.3%) were female, 209 (71.2%) were male, and 13 (4.5%) were those categorized as Other. As a reminder, the proportion of female, male, and Other TRA applicants is 20.4%, 76.1%, and 3.5%, respectively, from FY2010–2020. The percentage of female applicants in FY2021 (24.3%) was significantly higher than the baseline (FY2010–2020) percentage of 20.4% ($\chi_1^2 = 3.17, p = 0.04$). The percentage of male

applicants in FY2021 (71.2%) was significantly lower than the baseline percentage of 76.1% ($\chi_1^2 = 4.19, p = 0.02$). The percentage of applicants who were categorized as Other in FY2021 (4.5%) was not significantly different from the baseline percentage of 3.5% ($\chi_1^2 = 0.47, p = 0.25$).

A total of 10 applications and 20 applicants were funded under the general TRA FOA for FY2021. Results from the logistic regression and analysis of deviance indicated that whether an individual had a funded application in FY2021 was not significantly influenced by that individual's gender ($p = 0.59$) or race ($p = 0.42$). In addition, results from the Fisher's exact test also indicated that there were no significant differences between the breakdown of FY2021 applicants and awardees by gender ($p = 1.0$).

The percentage of awardees who were female in FY2021 was not significantly higher than the percentage of awardees who were female (20.3%) from FY2010–2020 ($\chi_1^2 = 0.35, p = 0.72$). The percentage of awardees who were male in FY2021 was not significantly less than the percentage of awardees who were male (77.5%) from FY2010–2020 ($\chi_1^2 = 0.07, p = 0.60$). The percentage of awardees who were in the Other gender category in FY2021 was not significantly higher than the percentage of awardees who were in the Other gender category (2.1%) from FY2010–2020 ($\chi_1^2 = 0.01, p = 0.45$).

2) Race

Of the 295 applicants in FY2021, 81 (27.5%) were Asian, 12 (4.1%) were Black or African American, 167 (56.6%) were White, and 35 (11.9%) were those categorized as Other. The proportion of Asian, Black or African American, White, and Other TRA applicants is 22.5%, 1.4%, 62.4% and 13.7%, respectively, for FY2010–2020. The percentage of Asian applicants in FY2021 (27.5%) was significantly higher than the baseline percentage of 22.5% ($\chi_1^2 = 3.88, p = 0.02$). The percentage of Black or African American applicants in FY2021 (4.1%) was significantly higher than the baseline percentage of 1.4% ($\chi_1^2 = 13.3, p < 0.001$). The percentage of White applicants in FY2021 (56.6%) was significantly lower than the baseline percentage of 62.4% ($\chi_1^2 = 3.97, p = 0.02$). The percentage of Other applicants in FY2021 (11.9%) was not significantly different from the baseline percentage of 13.7% ($\chi_1^2 = 0.69, p = 0.80$).

Results from the Fisher's exact test indicated that there were no significant differences between the breakdown of FY2021 applicants and awardees by race ($p = 1.0$). The percentage of awardees who were Asian in FY2021 was not significantly higher than the percentage of awardees who were Asian (15.7%) from FY2010–2020 ($\chi_1^2 = 0.01, p = 0.53$). The percentage of awardees who were Black or African American in FY2021 was significantly greater than the percentage of awardees who were Black or African American (1.7%) from FY2010–2020 ($\chi_1^2 = 4.05, p = 0.02$). The percentage of awardees who were White in FY2021 was not significantly less than the percentage of awardees who were White (70.3%) from FY2010–2020 ($\chi_1^2 = 1.02, p = 0.16$). The percentage of awardees

who were in the Other race category in FY2021 was not significantly higher than the percentage of awardees who were in the Other race category (12.3%) from FY2010–2020 ($\chi_1^2 = 0.14, p = 0.36$).

3) Ethnicity

Of the 295 applicants in FY2021, 15 (5.1%) were Hispanic or Latino, 245 (83.1%) were not Hispanic or not Latino, and 35 (11.9%) were those categorized as Other. The proportion of Hispanic or Latino, not Hispanic or not Latino, and Other TRA applicants is 2.9%, 79.6%, and 17.6%, respectively, from FY2010 to FY2020. The percentage of Hispanic or Latino applicants in FY2021 (5.1%) was significantly higher than the baseline percentage of 2.9% ($\chi_1^2 = 4.25, p = 0.02$). The percentage of not Hispanic or not Latino applicants in FY2021 (83.1%) was not significantly different from the baseline percentage of 79.6% ($\chi_1^2 = 1.96, p = 0.92$). The percentage of applicants who were categorized as Other in FY2021 (11.9%) was not significantly different from the baseline percentage of 17.6% ($\chi_1^2 = 6.30, p = 0.99$).

Results from the Fisher's exact test indicated that there were no significant differences between the breakdown of FY2021 applicants and awardees by ethnicity ($p = 1.0$).

The percentage of awardees who were Hispanic or Latino in FY2021 was not significantly higher than the percentage of awardees who were Hispanic or Latino (3.8%) from FY2010–2020 ($\chi_1^2 < 0.001, p = 0.50$). The percentage of awardees who were not Hispanic or not Latino in FY2021 was not significantly less than the percentage of awardees who were not Hispanic or not Latino (83.1%) from FY2010–2020 ($\chi_1^2 = 0.13, p = 0.36$). The percentage of awardees who were in the Other ethnicity category in FY2021 was not significantly higher than the percentage of awardees who were in the Other ethnicity category (13.1%) from FY2010–2020 ($\chi_1^2 = 0.06, p = 0.40$).

c. Comparison of Scores between FY2020 and FY2021 TRA Applicants

As a reminder, the NIH grant application scoring system ranges from 1 to 9 where 1 indicates an exceptional application and 9 indicates a poor application.[22] All applications reviewed in Phase IIIa receive preliminary scores from the EB members who reviewed that application. Applications with lower preliminary scores were discussed in Phase IIIb of the review process and received final scores.

Data in this section report the average preliminary and final scores for categories having one or more scores reported. A dash (-) indicates that an average could not be taken.

[22] See 2016 NIH "Scoring Guidance.", available at: https://grants.nih.gov/grants/policy/review/rev_pre.

1) Gender

Eighty FY2020 TRA applications from 145 applicants received a preliminary score from the assigned reviewers in Phase IIIa of the review process. Preliminary scores in Phase IIIa did not differ significantly by gender in FY2020 ($\chi_2^2 = 0.37, p = 0.83$). The average (± SE) preliminary score among FY2020 female applicants was 5.0 (± 0.16), 4.8 (± 0.09) for male applicants, and 4.9 (± 0.34) for Other applicants (Table 1).

Fifty-four FY2021 TRA applications from 83 applicants received a preliminary score in Phase IIIa of the review process. Preliminary scores in Phase IIIa did not differ significantly by gender in FY2021 ($\chi_2^2 = 1.78, p = 0.41$). The average (± SE) preliminary score among FY2021 female applicants was 4.7 (± 0.16), 4.4 (± 0.11) for male applicants, and 5.0 (± 0.56) for Other applicants.

Thirty FY2020 TRA applications from 45 applicants received a final score in Phase IIIb from all EB reviewers not having a conflict of interest. Final scores did not differ significantly by gender ($\chi_2^2 = 0.12, p = 0.94$). The average (± SE) final score among FY2020 female applicants was 3.8 (± 0.19), 3.8 (± 0.07) for male applicants, and 4.0 (± 0.21) for Other applicants.

Twenty-eight FY2021 TRA applications from 47 applicants received a final score in Phase IIIb of the review process. Final scores in Phase IIIb differed significantly by gender ($\chi_2^2 = 39.8, p < 0.001$). Specifically, final scores did not significantly differ between female and male applicants ($p = 0.10$), but applicants who were categorized as Other had significantly higher final scores than female and male applicants ($p < 0.001$ for comparisons with female and male applicants). The average (± SE) final score among FY2021 female applicants was 3.9 (± 0.29), 3.6 (± 0.30) for male applicants, and 5.5 (± 0.70) for Other applicants.

Table 1. Preliminary and Final Scores for FY2020 and FY2021 Applicants by Gender

Scores	*Female*	*Male*	*Other*
Preliminary FY2020	5.0 (± 0.16)	4.8 (± 0.09)	4.9 (± 0.34)
Preliminary FY2021	4.7 (± 0.16)	4.4 (± 0.11)	5.0 (± 0.56)
Final FY2020	3.8 (± 0.19)	3.8 (± 0.07)	4.0 (± 0.21)
Final FY2021	3.9 (± 0.29)	3.6 (± 0.30)	5.5 (± 0.70)

The difference in average final scores for each multiple comparison within gender (i.e., female and male; female and other; male and other) increased from FY2020 to FY2021 (Figure 54). Specifically, the absolute difference in average final score between female and male applicants was 0.02 in FY2020 and 0.28 in FY2021, resulting in an increased difference of 0.26 between FY2021 and FY2020. For female and Other

applicants, the difference in average final score was 0.12 in FY2020 and 1.69 in FY2021, resulting in an increased difference of 1.57. For male and Other applicants, the difference in average final score was 0.10 in FY2020 and 1.97 in FY2021, resulting in an increased difference of 1.87.

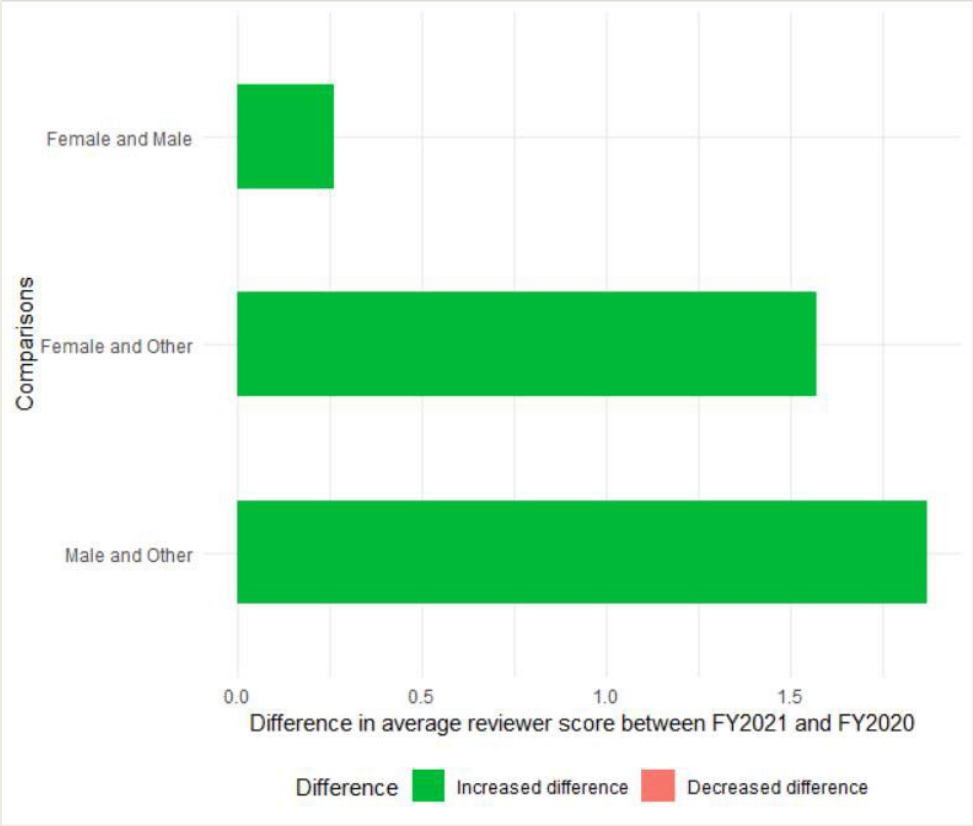


Figure 54. Difference in Average Final Score between FY2021 and FY2020 by Gender

2) Race

Preliminary scores did not differ significantly by race in FY2020 ($\chi_2^2 = 5.42, p = 0.14$) or in FY2021 ($\chi_2^2 = 3.36, p = 0.34$). The average (± SE) preliminary score for Asian applicants in FY2020 was 5.0 (± 0.15), 4.7 (± 0.10) for White applicants, and 5.4 (± 0.20) for Other applicants (Table 2). The average (± SE) preliminary score for Asian applicants in FY2021 was 4.7 (± 0.17), 3.8 (± 0.37) for Black or African American applicants; 4.3 (± 0.12) for White applicants, and 4.7 (± 0.32) for Other applicants.

Final scores differed significantly by race in FY2020 ($\chi_2^2 = 16.4, p < 0.001$). Specifically, White applicants had significantly lower scores than Asian applicants ($p < 0.001$) but did not have significantly different scores from applicants categorized as Other ($p = 0.95$); and Asian applicants did not have significantly different scores from applicants categorized as Other ($p = 0.13$). The average (± SE) final score for Asian applicants in

FY2020 was 4.5 (± 0.12), 3.7 (± 0.07) for White applicants, and 3.8 (± 0.23) for Other applicants.

Final scores differed significantly by race in FY2021 ($\chi_3^2 = 78.2, p < 0.001$). Specifically, White applicants had significantly lower final scores compared to all other races ($p < 0.001$ for all comparisons); Asian applicants did not have significantly different scores from Other applicants ($p = 0.08$) but had significantly lower scores than Black or African American applicants ($p < 0.01$); and Black or African American applicants did not have significantly different scores than Other applicants ($p = 0.25$). The average (± SE) Phase IIIb final score for Asian applicants in FY2021 was 4.1 (± 0.46), 5.4 (± 0.00) for Black or African American applicants, 3.4 (± 0.27) for White applicants, and 4.6 (± 0.96) for Other applicants.

Table 2. Preliminary and Final Scores for FY2020 and FY2021 Applicants by Race

Scores	*Asian*	Black or *African American*	*White*	*Other*
Preliminary FY2020	5.0 (± 0.15)	—	4.7 (± 0.10)	5.4 (± 0.20)
Preliminary FY2021	4.7 (± 0.17)	3.8 (± 0.37)	4.3 (± 0.12)	4.7 (± 0.32)
Final FY2020	4.5 (± 0.12)	—	3.7 (± 0.07)	3.8 (± 0.23)
Final FY2021	4.1 (± 0.46)	5.4 (± 0.00)	3.4 (± 0.27)	4.6 (± 0.96)

Note: Dash (—) denotes data are not reported to comply with DUDA.

The difference in average final scores between FY2020 and FY2021 for each multiple comparison within race differed based on which races were compared to one another (Figure 55). To comply with Data Use and Special Data Access (DUDA) average scores for Black or African Americans are not reported. Differences in average final score are provided for Asian and White, Asian and Other, and White and Other. The absolute difference in average final score between Asian and White applicants was 0.81 in FY2020 and 0.69 in FY2021, resulting in a decreased difference of 0.12 between FY2021 and FY2020. For Asian and Other applicants, the difference in average final score was 0.72 in FY2020 and 0.61 in FY2021, resulting in a decreased difference of 0.11. For White and Other applicants, the difference in average final score was 0.09 in FY2020 and 1.30 in FY2021, resulting in an increased difference of 1.21.

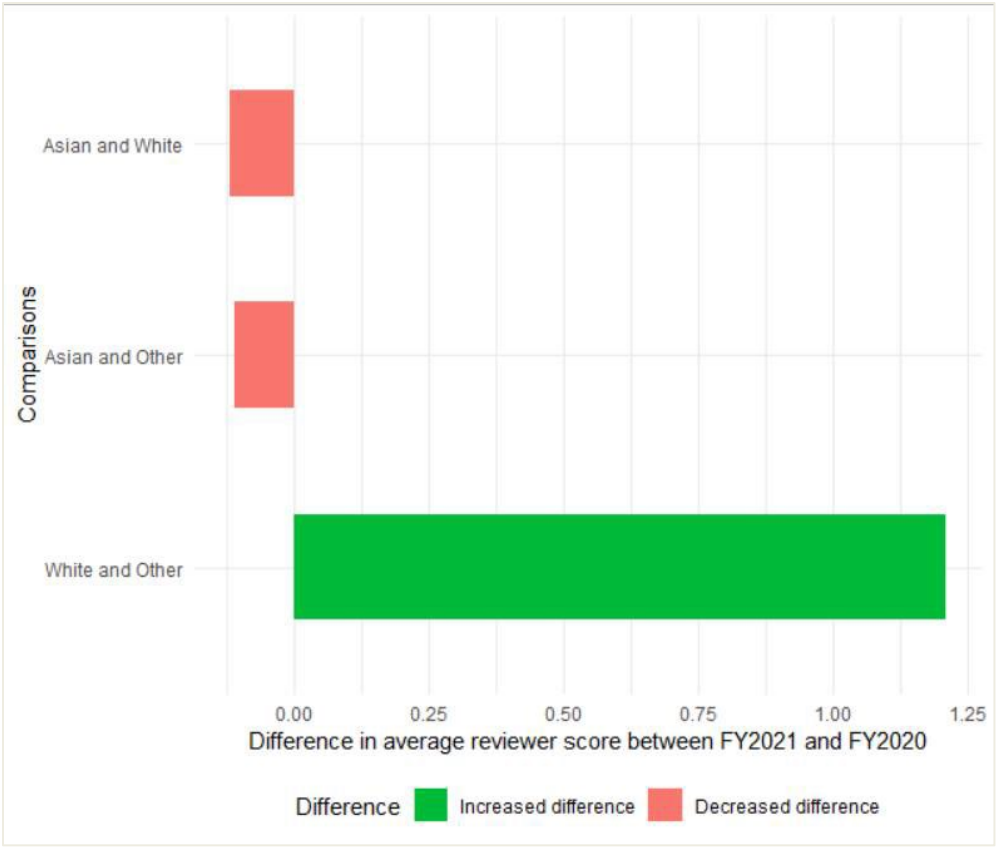


Figure 55. Difference in Average Final Score between FY2021 and FY2020 by Race

3) Ethnicity

Preliminary scores differed significantly by ethnicity in FY2020 ($\chi_2^2 = 6.96, p = 0.03$). Specifically, Hispanic or Latino applicants had significantly higher preliminary scores than not Hispanic or not Latino applicants ($p = 0.03$), but did not differ significantly with Other applicants ($p = 0.24$). Not Hispanic or not Latino applicants did not have significantly different preliminary scores than Other applicants ($p = 0.42$). The average (± SE) preliminary score for not Hispanic or not Latino applicants is 4.8 (± 0.08) and 5.2 (± 0.21) for Other applicants (Table 3). Preliminary scores did not differ significantly by ethnicity in FY2021 ($\chi_2^2 = 1.41, p = 0.49$). The average (± SE) preliminary score for Hispanic or Latino was 3.9 (± 0.35), 4.5 (± 0.10 for not Hispanic or not Latino applicants, and 4.2 (± 0.25) for Other applicants.

Final scores did not differ significantly by ethnicity in FY2020 ($\chi_1^2 = 0.13, p = 0.72$). The average (± SE) final score for not Hispanic or not Latino applicants in FY2020 was 3.8 (± 0.07), and 3.9 (± 0.18) for Other applicants. Final scores differed significantly by ethnicity in FY2021 ($\chi_2^2 = 24.8, p < 0.001$): Hispanic or Latino applicants had significantly higher final scores than not Hispanic or nor Latino ($p < 0.001$) and Other applicants ($p < 0.01$); and there was no significant difference in final scores between not

Hispanic or not Latino and Other applicants ($p = 0.09$). The average (± SE) final score for Hispanic or Latino applicants in FY2021 was 5.1 (± 0.12), 3.6 (± 0.26) for not Hispanic or not Latino applicants, and 4.0 (± 0.54) for Other applicants.

Table 3. Preliminary and Final Scores for FY2020 and FY2021 Applicants by Ethnicity

Scores	*Hispanic or Latino*	*Not Hispanic or not Latino*	Other
Preliminary FY2020	—	4.8 (± 0.08)	5.2 (± 0.21)
Preliminary FY2021	3.9 (± 0.35)	4.5 (± 0.10)	4.2 (± 0.25)
Final FY2020	—	3.8 (± 0.07)	3.9 (± 0.18)
Final FY2021	5.1 (± 0.30)	3.6 (± 0.26)	4.0 (± 0.54)

Note: Dash (—) denotes data are not reported to comply with DUDA.

To comply with the DUDA average scores for Hispanic or Latino are not reported. Differences in average final score are provided for not Hispanic or not Latino and Other. The difference in average final scores between not Hispanic or not Latino and Other increased from FY2020 to FY2021 (Figure 56). Specifically, the absolute difference in average final score between not Hispanic and not Latino and Other applicants was 0.10 in FY2020 and 0.34 in FY2021, resulting in an increased difference of 0.24 between FY2021 and FY2020.

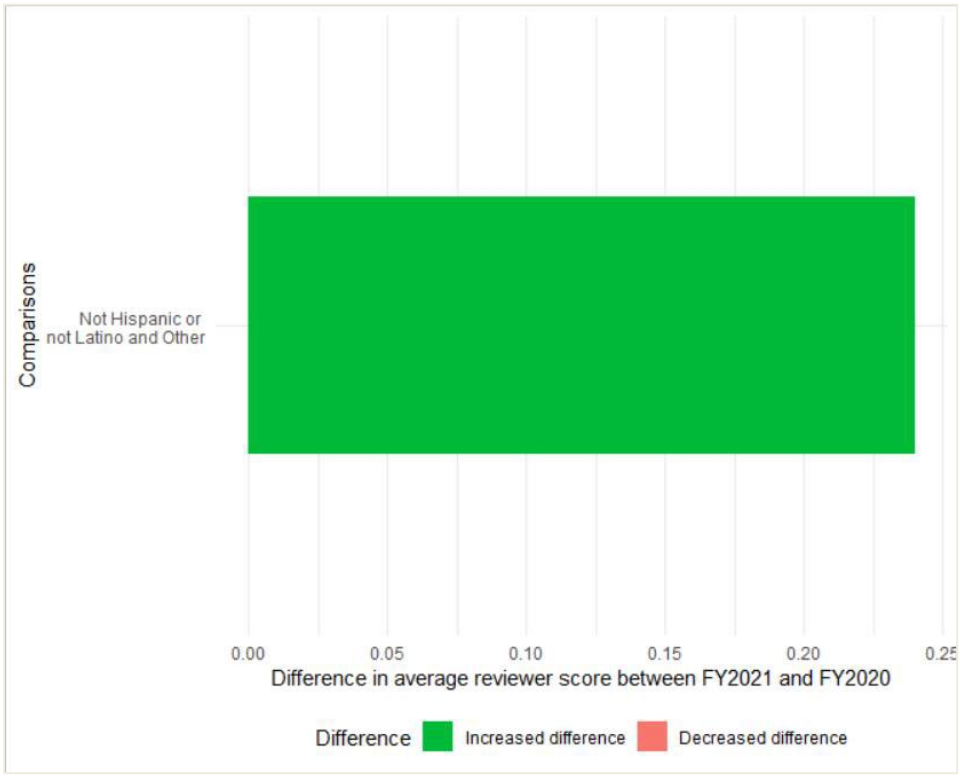


Figure 56. Difference in Average Final Score between FY2021 and FY2020 by Ethnicity

d. Stratification of FY2020 TRA Applicants

1) Gender

Of the 342 TRA applicants in FY2020, 82 (24.0%) were female, 241 (70.5%) were male, and 19 (5.6%) were Other. A total of 45 applicants made it to Phase III and 14 applicants were funded. The gender breakdown of applicants who made it to Phase III and who were awarded is not provided because the number of individuals in at least 2 of the gender groups is less than 11.

Results from the two-sample proportion tests comparing the percentage of applicants to the percentage of applicants who made it to Phase III for each gender showed that there were no significant differences in the two percentages for females ($p = 0.28$), males ($p = 0.40$), or Others ($p = 1.0$). Given the sample sizes of the applicant pool and those who made it to Phase III, the power to detect differences for females was 0.12, 0.18 for males, and 0.06 for Others.[23]

Results from the two-sample proportion tests comparing the percentage of applicants who made it to Phase III and the percentage of funded applicants for each gender showed that there were no significant differences in the two percentages for females ($p = 0.21$), males ($p = 0.51$), and Others ($p = 0.77$). Given the sample sizes of those who made it to Phase III and the awardee pool, the power to detect differences for females was 0.33, 0.16 for males, and 0.40 for Others.

Results from the two-sample proportion tests comparing the percentage of applicants and the percentage of funded applicants for each gender showed that there were no significant differences in the two percentages for females ($p = 0.47$; Figure 57), males ($p = 0.84$), and Others ($p = 0.76$). Given the sample sizes of those who made it to Phase III and the awardee pool, the power to detect differences for females was 0.16, 0.08 for males, and 0.42 for Others.

[23] As a reminder, power ranges between 0 to 1 with 0 being low power and 1 being high power. Low power reduces the likelihood a true effect will be detected. Low power also reduces the likelihood that a statistically significant result reflects a true effect.

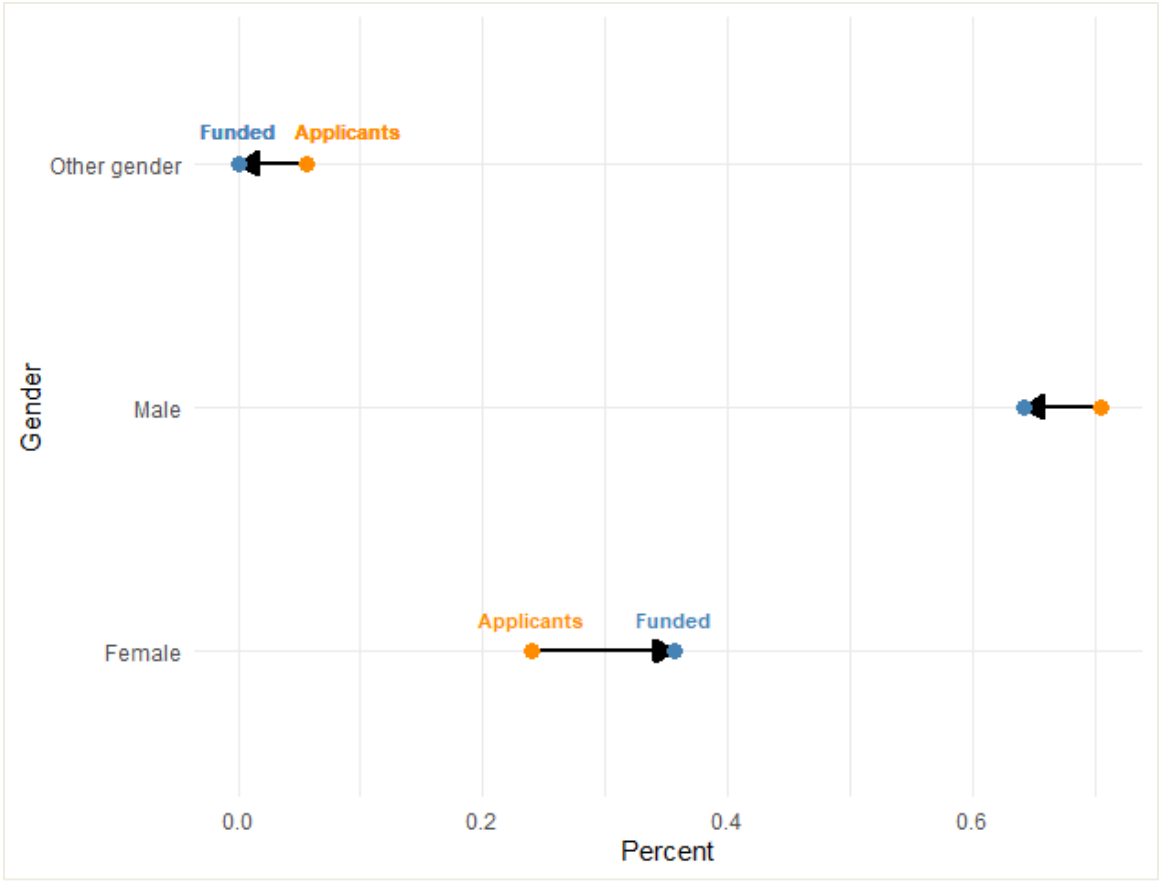


Figure 57. Percentage of Total (orange) and Funded (blue) Applicants by Gender for FY2020

2) Race

Of the 342 TRA applicants in FY2020, 81 (23.7%) were Asian, 212 (62.0%) were White, and 49 (14.3%) were either Black or African American or Other. The racial breakdown of applicants who made it to Phase III and who were awarded is not provided because the number of individuals in at least 2 of the racial groups is less than 11.

Results from the two-sample proportion tests comparing the percentage of applicants to the percentage of applicants who made it to Phase III for each race showed that there were no significant differences in the two percentages for Asian applicants ($p = 0.98$), Black or African American applicants ($p = 0.57$), White applicants ($p = 0.30$), or Other applicants ($p = 0.45$). Given the sample sizes of the applicant pool and those who made it to Phase III, the power to detect differences for Asian applicants was 0.06, 0.54 for Black or African American applicants, 0.23 for White applicants, and 0.20 for Other applicants.

Results from the two-sample proportion tests comparing the percentage of applicants who made it to Phase III and the percentage of funded applicants for each race showed that there were no significant differences in the two percentages for Asians ($p = 0.38$), Whites

(p = 0.46), or Others (p = 1.0).[24] Given the sample sizes of those who made it to Phase III and the awardee pool, the power to detect differences for Asian applicants was 0.30, 0.22 for White applicants, and 0.05 for applicants.

Results from the two-sample proportion tests comparing the percentage of applicants and the percentage of funded applicants for each race showed that there were no significant differences in the two percentages for Asian applicants (p = 0.26; Figure 58), Black or African American applicants (= 1.0), White applicants (= 0.13), or Other applicants (= 0.92). Given the sample sizes of those who made it to Phase III and the awardee pool, the power to detect differences for Asian applicants was 0.41, 0.22 for Black or African American applicants, 0.53 for White applicants, and 0.09 for Other applicants.

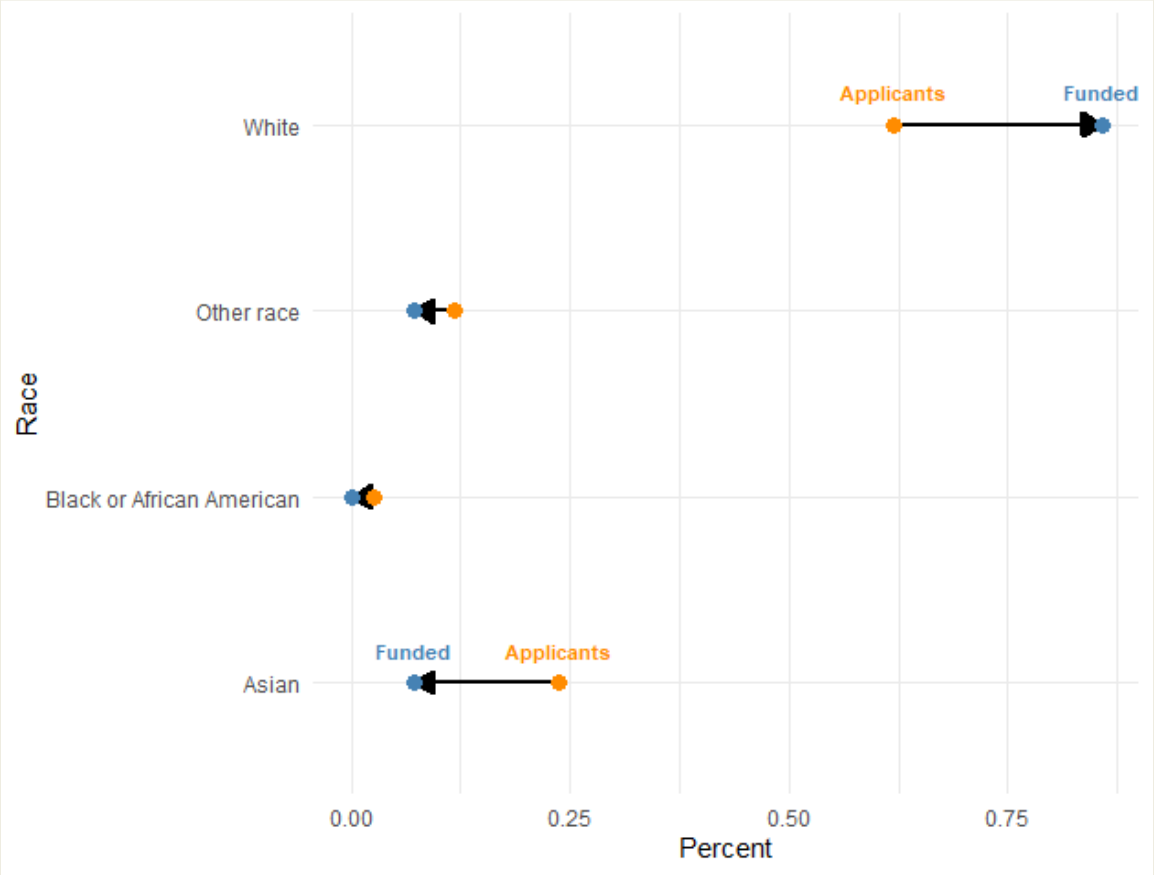


Figure 58. Percent of Total (orange) and Funded (blue) Applicants by Race for FY2020

[24] A two-sample proportion test could not be performed for Black or African American applicants because there were 0 Black or African American applicants who made it to Phase III and therefore, 0 Black or African American applicants who were funded.

3) Ethnicity

Of the 342 TRA applicants in FY2020, 292 (85.4%) were not Hispanic or not Latino, and 50 (14.6%) were either Hispanic or Latino or Other. The ethnicity breakdown of applicants who made it to Phase III and who were awarded is not provided because the number of individuals in at least 2 of the ethnicity groups is less than 11.

Results from the two-sample proportion tests comparing the percentage of applicants to the percentage of applicants who made it to Phase III by ethnicity showed that there were no significant differences in the two percentages for Hispanic or Latino applicants ($p = 0.80$), not Hispanic or not Latino applicants ($p = 0.42$), or Other applicants ($p = 0.60$). Given the sample sizes of the applicant pool and those who made it to Phase III, the power to detect differences for Hispanic or Latino applicants was 0.39, 0.20 for not Hispanic or not Latino applicants, and 0.13 for Other applicants.

Results from the two-sample proportion tests comparing the percentage of applicants who made it to Phase III and the percentage of funded applicants by ethnicity showed that there were no significant differences in the two percentages for not Hispanic or not Latino applicants ($p = 1.0$), or Other applicants ($p = 1.0$).[25] Given the sample sizes of those who made it to Phase III and the awardee pool, the power to detect differences was 0.06 for both not Hispanic or not Latino applicants and applicants categorized as Other.

Results from the two-sample proportion tests comparing the percentage of applicants and the percentage of funded applicants by ethnicity showed that there were no significant differences in the two percentages for Hispanic or Latino applicants ($p = 1.0$), not Hispanic or not Latino applicants ($p = 0.69$), or Other applicants ($p = 0.82$). Given the sample sizes of those who made it to Phase III and the awardee pool, the power to detect differences for Hispanic or Latino applicants was 0.16, 0.15 for not Hispanic or not Latino applicants, and 0.11 for Other applicants.

e. Stratification of FY2021 TRA Applicants

1) Gender

As a reminder, of the 295 TRA applicants in FY2021, 73 (24.7%) were female, 209 (70.8%) were male, and 13 (4.4%) were Other. A total of 47 applicants made it to Phase III, and 20 applicants were funded.

Results from the two-sample proportion tests comparing the percentage of applicants to the percentage of applicants who made it to Phase III for each gender showed that there

[25] A two-sample proportion test could not be performed for Hispanic or Latino applicants because there were 0 Hispanic or Latino applicants who made it to Phase III and therefore, 0 Hispanic or Latino applicants who were funded.

were no significant differences in the two percentages for females, males, or Others ($p = 1.0$ for each). Given the sample sizes of the applicant pool and those who made it to Phase III, the power to detect differences was 0.05 for females, males, and Others. As a reminder, power ranges from 0 to 1, and low statistical power means the likelihood of detecting a true effect is reduced.

Results from the two-sample proportion tests comparing the percentage of applicants who made it to Phase III and the percentage of funded applicants for each gender showed that there were no significant differences in the two percentages for females ($p = 0.21$), males ($p = 0.51$), and Others ($p = 0.77$). Given the sample sizes of those who reached Phase III and the awardee pool, the power to detect differences for females was 0.33, 0.16 for males, and 0.40 for Others.

Results from the two-sample proportion tests comparing the percentage of applicants and the percentage of funded applicants for each gender showed that there were no significant differences in the two percentages for females ($p = 0.47$), males ($p = 0.84$), and Others ($p = 0.76$). Given the sample sizes of those who made it to Phase III and the awardee pool, the power to detect differences for females was 0.16, 0.08 for males, and 0.42 for Others.

2) Race

Of the 295 TRA applicants in FY2021, 81 (27.5%) were Asian, 12 (4.1%) were Black or African American, 167 (56.6%) were White, and 35 (11.9%) were Other. The racial breakdown of applicants who made it to Phase III and who were awarded is not provided because the number of individuals in at least two of the racial groups is less than 11.

Results from the two-sample proportion tests comparing the percentage of applicants to the percentage of applicants who made it to Phase III for each race showed that there were no significant differences in the two percentages for Asian applicants ($p = 0.92$), Black or African American applicants ($p = 1.0$), White applicants ($p = 0.62$), or Other applicants ($p = 0.67$). Given the sample sizes of the applicant pool and those who made it to Phase III, the power to detect differences for Asian applicants was 0.06, 0.05 for Black or African American applicants, 0.10 for White applicants, and 0.11 for Other applicants.

Results from the two-sample proportion tests comparing the percentage of applicants who made it to Phase III and the percentage of funded applicants for each race showed that there were no significant differences in the two percentages for Asians ($p = 0.53$), Black or African Americans ($p = 0.73$), Whites ($p = 1.0$), or Others ($p = 0.72$).[26] Given the sample sizes of those who made it to Phase III and the awardee pool, the power to detect

[26] A two-sample proportion test could not be performed for Black or African American applicants because there were 0 Black or African American applicants who made it to Phase III and therefore, 0 Black or African American applicants who were funded.

differences for Asian applicants was 0.17, 0.14 for Black or African American applicants, 0.05 for White applicants, and 0.12 for Other applicants.

Results from the two-sample proportion tests comparing the percentage of applicants and the percentage of funded applicants for each race showed that there were no significant differences in the two percentages for Asian applicants ($p = 0.34$), Black or African American applicants ($p = 0.49$), White applicants ($p = 0.95$), or Other applicants ($p = 0.95$). Given the sample sizes of those who made it to Phase III and the awardee pool, the power to detect differences for Asian applicants was 0.27, 0.18 for Black or African American applicants, 0.06 for White applicants, and 0.07 for Other applicants.

3) Ethnicity

As a reminder, of the 295 TRA applicants in FY2020, 15 (5.1%) were Hispanic or Latino, 245 (83.1%) were not Hispanic or not Latino, and 35 (11.9%) were Other. The ethnicity breakdown of applicants who made it to Phase III and who were awarded is not provided because the number of individuals in at least 2 of the ethnicity groups is less than 11.

Results from the two-sample proportion tests comparing the percentage of applicants to the percentage of applicants who made it to Phase III by ethnicity showed that there were no significant differences in the two percentages for Hispanic or Latino applicants ($p = 1.0$), not Hispanic or not Latino applicants ($p = 0.60$), or Other applicants ($p = 0.45$). Given the sample sizes of the applicant pool and those who made it to Phase III, the power to detect differences for Hispanic or Latino applicants was 0.06, 0.11 for not Hispanic or not Latino applicants, and 0.16 for Other applicants.

Results from the two-sample proportion tests comparing the percentage of applicants who made it to Phase III and the percentage of funded applicants by ethnicity showed that there were no significant differences in the two percentages for Hispanic or Latino applicants ($p = 1.0$), not Hispanic or not Latino applicants ($p = 1.0$), or Other applicants ($p = 1.0$).[27] Given the sample sizes of those who made it to Phase III and the awardee pool, the power to detect differences was 0.05 for Hispanic or Latino applicants, not Hispanic or not Latino applicants, and applicants categorized as Other.

Results from the two-sample proportion tests comparing the percentage of applicants and the percentage of funded applicants by ethnicity showed that there were no significant differences in the two percentages for Hispanic or Latino applicants ($p = 1.0$), not Hispanic or not Latino applicants ($p = 0.97$), or Other applicants ($p = 0.95$). Given the sample sizes of those who made it to Phase III and the awardee pool, the power to detect

[27] A two-sample proportion test could not be performed for Hispanic or Latino applicants because there were 0 Hispanic or Latino applicants who made it to Phase III and therefore, 0 Hispanic or Latino applicants who were funded.

differences for Hispanic or Latino applicants was 0.05, 0.06 for not Hispanic or not Latino applicants, and 0.07 for Other applicants.

f. Simpson's Diversity Index for TRA Applicants

As a reminder, Simpson's Diversity Index ranges between 0 and 1 with values closer to 0 indicating low diversity and values closer to 1 indicating high diversity.

1) Gender

Simpson's Diversity Index for gender ranged between 0.34 and 0.45 for FY2010–2020, and was 0.43 in FY2021 (Figure 59). Overall, gender diversity was medium-low with the highest relative diversities observed in FY2018 (0.45), FY2020 (0.44), and FY2021 (0.43). Gender diversity among TRA applicants did not increase in FY2021 compared to previous years but was the third highest between FY2010–2021.

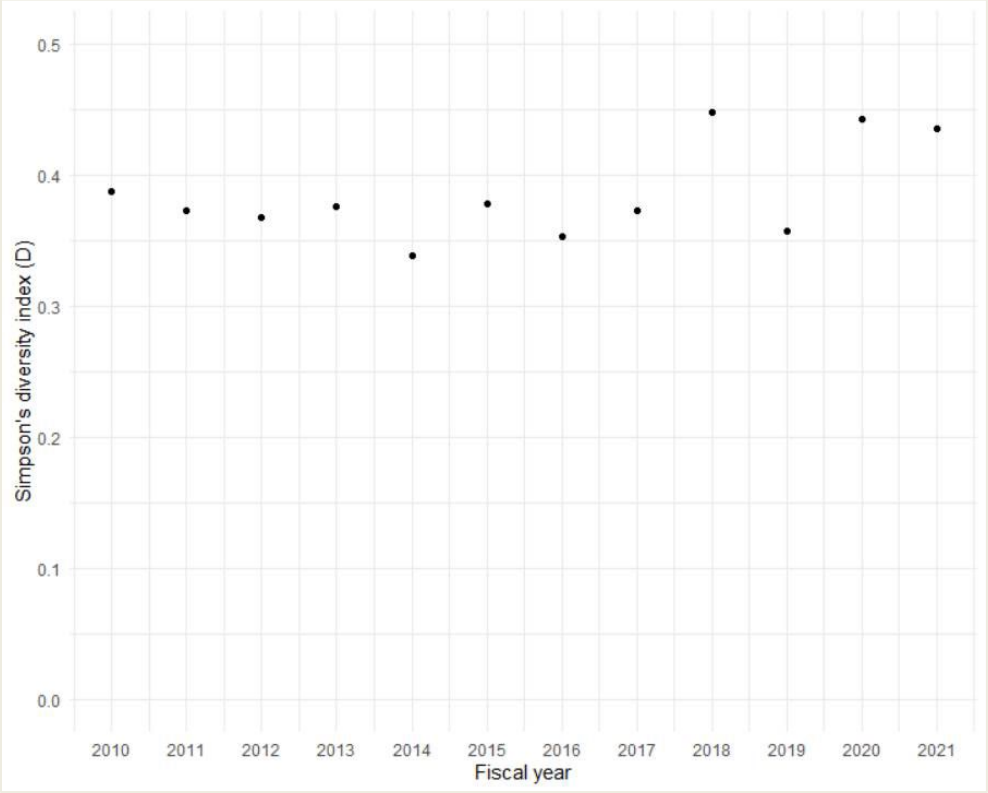


Figure 59. Simpson's Diversity Indices for TRA Applicants by Gender for FY2010–2021

2) Race

Simpson's Diversity Index for race ranged between 0.50 and 0.57 for FY2010–2020, and was 0.59 in FY2021 (Figure 60). Overall, racial diversity was medium with the highest

relative diversities observed in FY2021 (0.59), FY2010 (0.57), and FY2017 (0.57). Racial diversity among TRA applicants increased in FY2021 compared to all previous years.

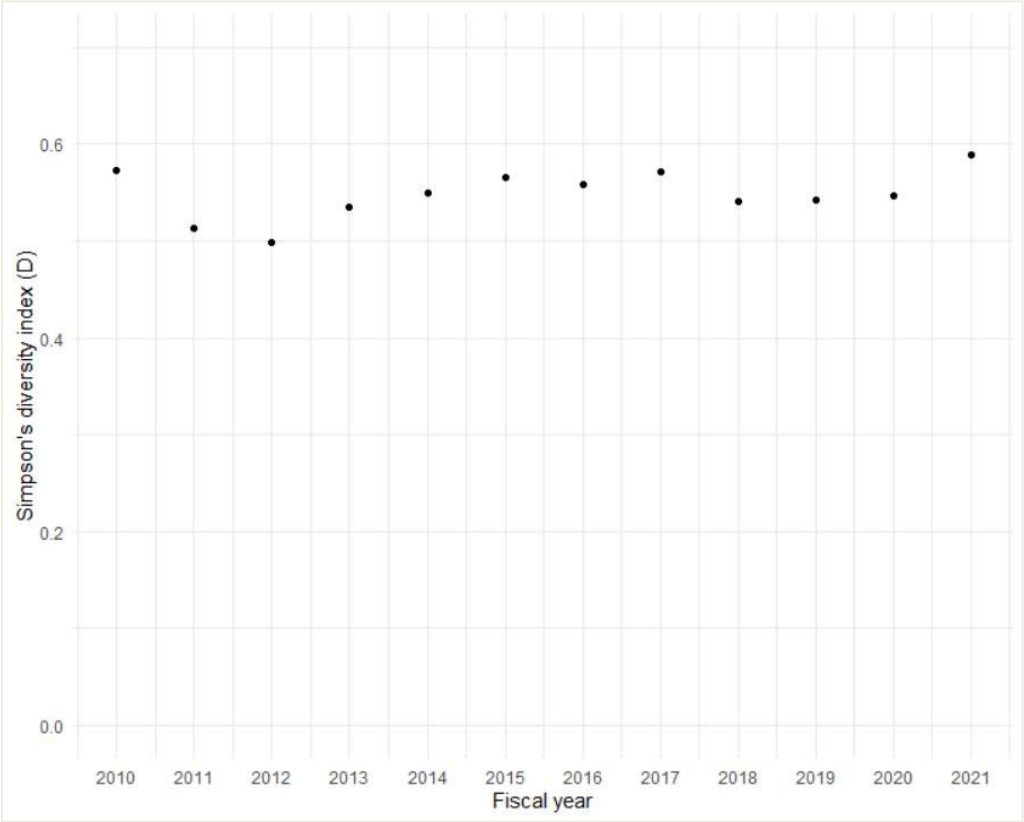


Figure 60. Simpson's Diversity Indices for TRA Applicants by Race for FY2010–2021

3) Ethnicity

Simpson's Diversity Index for ethnicity ranged between 0.25 and 0.42 for FY2010–2020, and was 0.29 in FY2021 (Figure 61). Overall, ethnic diversity was low to medium-low with the highest relative diversities observed in FY2010 (0.42), FY2011 (0.38), and FY2012 (0.33). Racial diversity among TRA applicants increased in FY2021 relative to FY2020 but is firmly within the observed values for Simpson's Diversity Index for previous years, and has the fourth lowest Simpson's Diversity Index across all years.

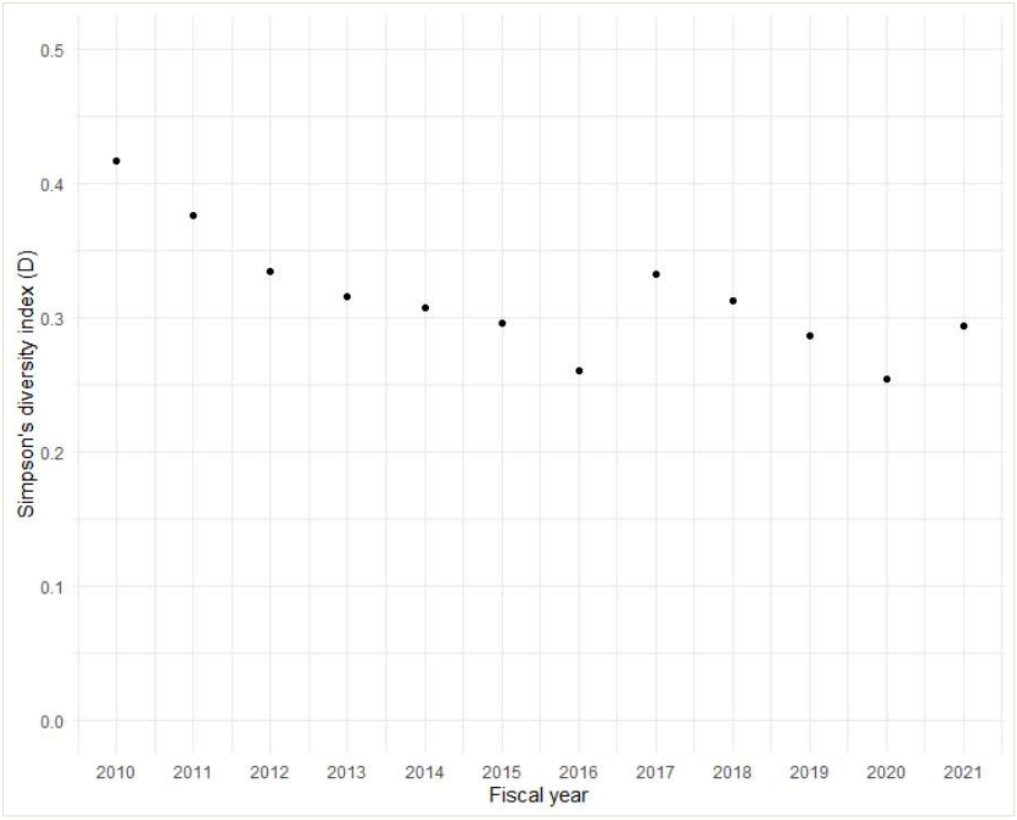


Figure 61. Simpson's Diversity Indices for TRA Applicants by Ethnicity for FY2010–2021

2. Geographic and Institutional Diversity

a. FY2010–2020 TRA Cohorts

1) Geographic Diversity

Of the 3,381 unique TRA applications received between FY2010–2020, applicants came from 495 different institutions across 49 States and the District of Columbia within the United States (Figure 62).[28] Fifteen States had TRA applicants every year between FY2010–2020;[29] and 10 States along with the District of Columbia had TRA applicants 10 out of the 11 years between FY2010–2020.[30] On the other end, four States had TRA

[28] FY2010–2020 TRA applicants also hailed from outside the United States. As a reminder, the United States in this analysis comprises the 50 States and the District of Columbia. FY2010–2020 TRA applicants from outside the United States included those from U.S. Territories (i.e., Puerto Rico) as well as non-U.S. territories (i.e., Australia, Canada, China, Israel, Italy, Japan, South Africa, Switzerland, and the United Kingdom).

[29] The 15 States were California, Connecticut, Florida, Georgia, Illinois, Maryland, Massachusetts, Michigan, New Jersey, New York, Ohio, Pennsylvania, Texas, Virginia, and Washington.

[30] The 10 States were Alabama, Arizona, Colorado, Indiana, Iowa, Minnesota, Missouri, North Carolina, Tennessee, and Utah.

applicants 2 out of the 11 years (Arkansas, Idaho, South Dakota, and Montana); 2 States had TRA applicants 1 out of the 11 years between FY2010–2020 (Alaska and Wyoming); and 1 State has never had a TRA applicant (North Dakota).

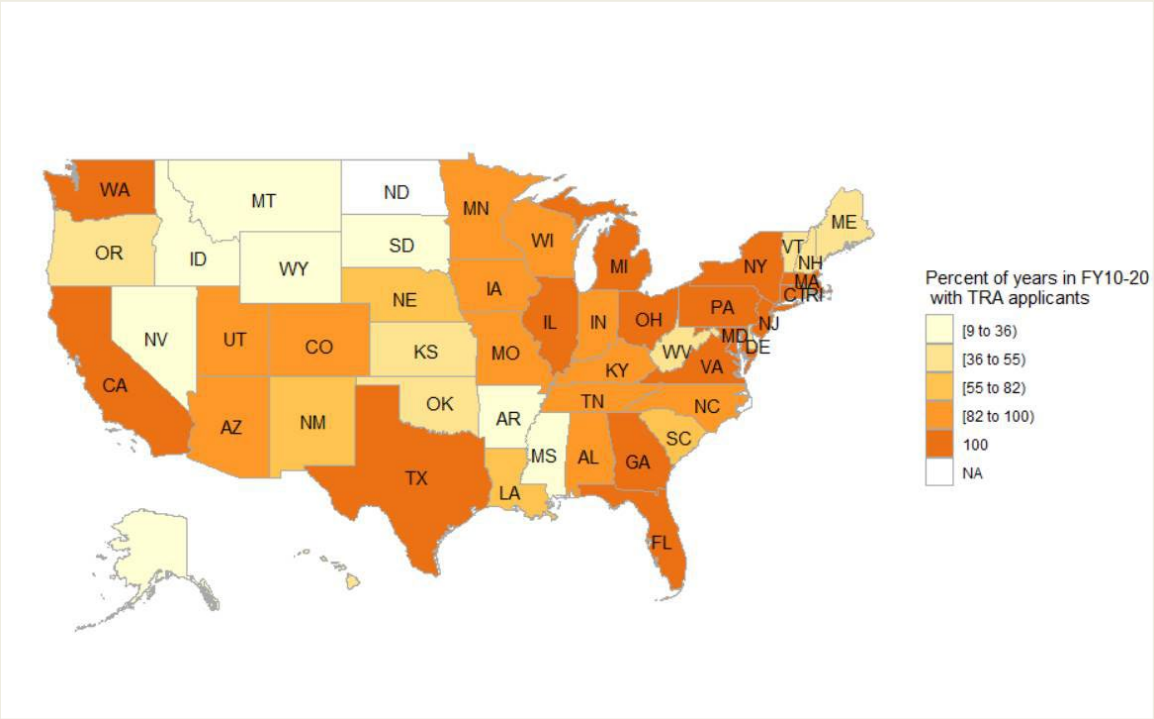


Figure 62. Percentage of Years between FY2010 and FY2020 That Each State Had at Least One TRA Applicant

The number of States represented by TRA applicants was highest in 2010 (49) followed by 2012 (48). Overall, the number of States represented by TRA applicants has decreased through time ($_{1,9} = 34.9, < 0.001$) with an average decrease of two States per year (Figure 63).

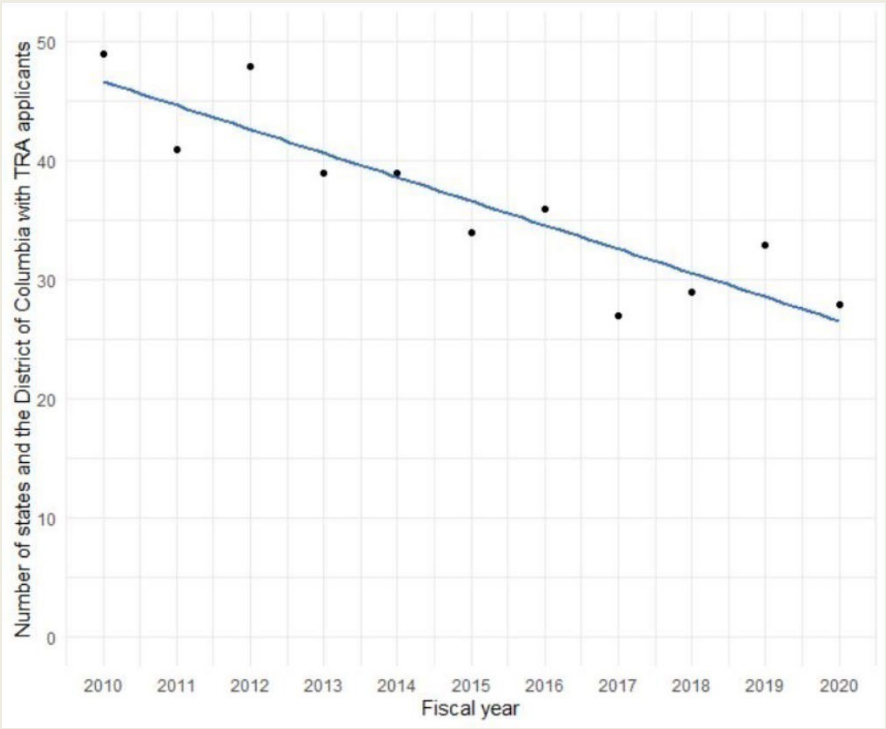


Figure 63. Number of States and the District of Columbia with TRA Applicants through FY2010–2020

Twenty-six States have had at least one TRA awardee between FY2010–2020 for an overall success rate of 51% (Table 4). The success rate differs widely across States. For instance, both California and Massachusetts have had at least one TRA awardee every year between FY2010–2020 (Figure 64). On the other hand, Ohio, Michigan, and New Jersey have had TRA applicants every year but have only received one TRA awardee between FY2010–2020. One explanation for the difference in success rates across States is that more TRA applications were received from applicants in California and Massachusetts (660 and 390, respectively) between FY2010–2020 than those received from Ohio, Michigan, and New Jersey (106, 58, and 112, respectively). The overall success rate of awarded applications where the contact PI is at an institution located in California is 5.2%, compared to a 7.7% success rate for applications from Massachusetts, 0.9% success rate for applications from Ohio, 0.9% success rate for applications from Michigan, and 1.7% success rate for applications from New Jersey. The full list of the number of years each State had at least one TRA applicant and received at least one TRA award can be found in Appendix O.

Table 4. States with at Least One TRA Awardee between FY2010 and FY2020

Number of years with at least one TRA awardee between FY2010–20	States
11	California, Massachusetts
9	New York
4	Illinois
3	Georgia, Maryland, Missouri, Pennsylvania, Texas
2	Alabama, Connecticut, Florida, Louisiana, North Carolina, Utah, Washington, Wisconsin
1	Arizona, Colorado, Delaware, Michigan, Minnesota, New Hampshire, New Jersey, Ohio, Oregon

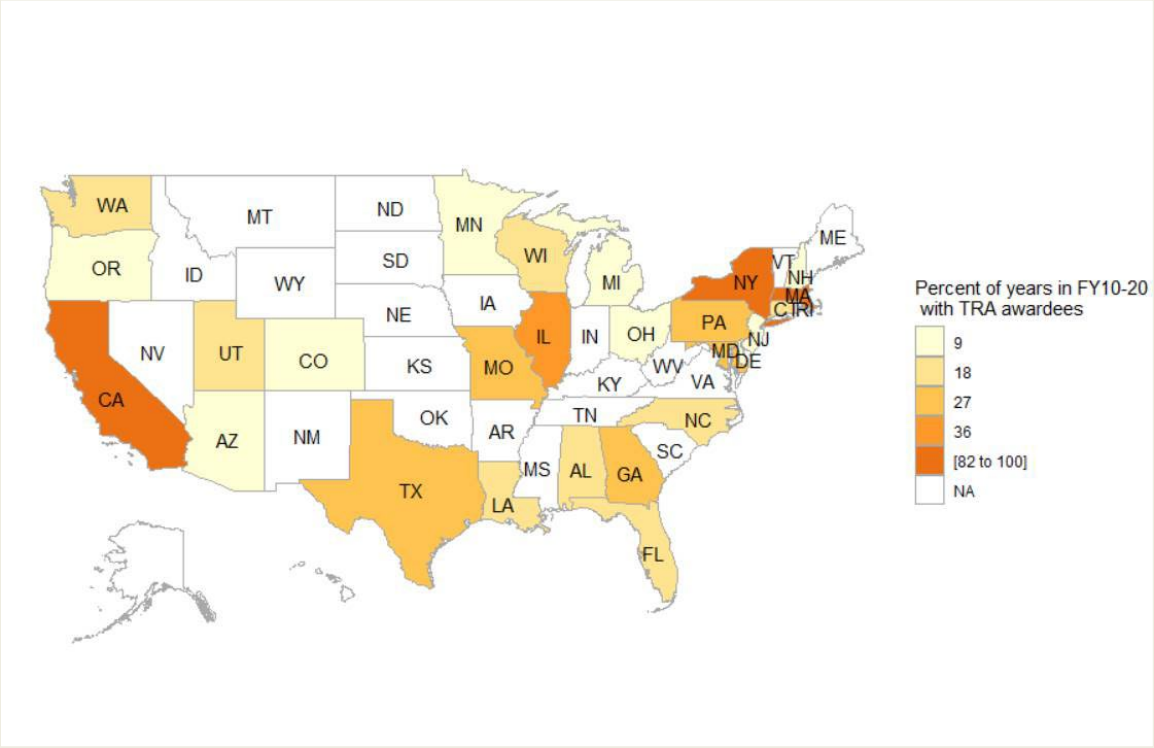


Figure 64. Percentage of Years between FY2010 and FY2020 That Each State Had at Least One TRA Awardee

2) Institution Diversity

Of the 485 unique institutions from which TRA applicants applied, 12 institutions had at least 1 TRA applicant each year between FY2010–2020. The number of institutions with TRA applicants was largest in FY2012 when applicants from 261 institutions applied to the TRA FOA, followed by FY2010 with 212 institutions. Overall, the number of institutions with TRA applicants has decreased through time but appears to have stabilized since FY2017 ($_{1,9} = 17.1, p < 0.01$; Figure 65).

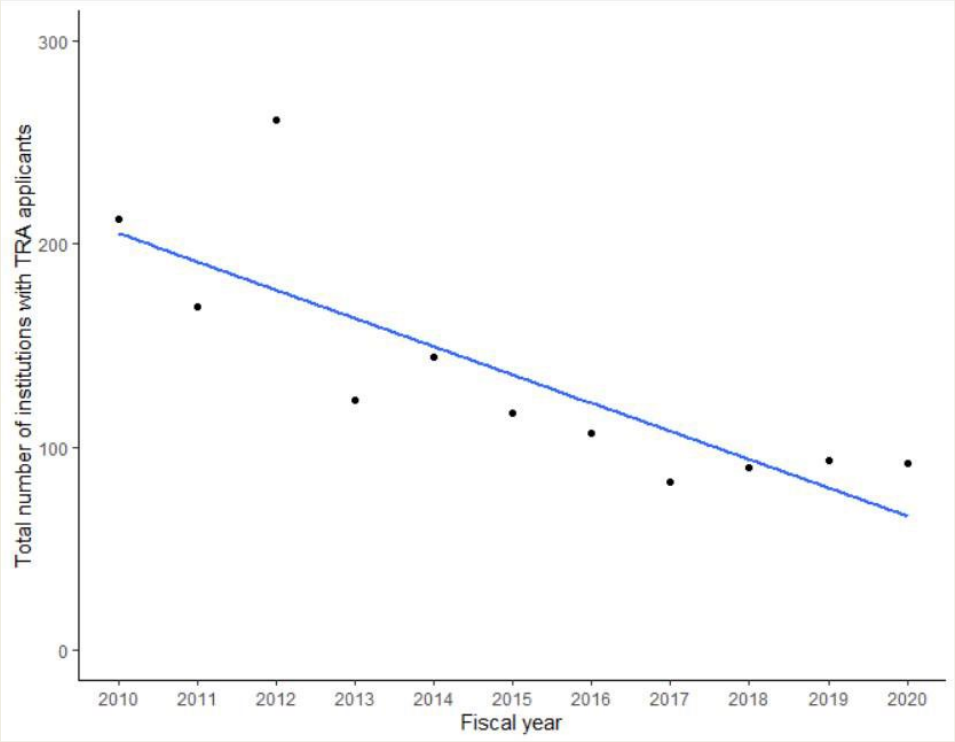


Figure 65. Number of Total Institutions with TRA Applicants between FY2010 and FY2020

The number of new institutions applying to the TRA program for the first time has changed over time (Table 5). STPI acknowledges that as new institutions apply, the pool of potential new institutions decreases, a limitation to consider when interpreting the analysis. The largest number of new institutions was 107, observed in FY2012. The number of new institutions has ranged from 0 to 9 between FY2016 and FY2020.

Table 5. Number of New Institutions Applying to the TRA Program

Fiscal year	Number of new institutions
2011	84
2012	107
2013	12
2014	21
2015	13
2016	0
2017	6
2018	4
2019	9
2020	9

Sixty-eight of the 485 institutions from which TRA applicants have applied have had at least 1 TRA awardee, for an overall success rate of 14%.

Of the 102 HBCUs recognized by the Department of Education as of July 2021,[31] there were TRA applicants from 5 HBCUs between FY2010–2020, and no awards were made. One applicant from an HBCU received a TRA award in FY2021.

Institutional diversity, as calculated by Simpson's Diversity Index, among institutions with TRA applicants was high across all years between FY2010–2020 with an average (± SE) of 0.99 (± 0.001) (Figure 66). Institutional diversity of institutions that have TRA awardees was lower relative to institutional diversity of institutions with TRA applicants with an average (± SE) of 0.89 (± 0.008).

[31] The list of 102 HBCUs can be found on the National Center for Education Statistics, available at https://nces.ed.gov/COLLEGENAVIGATOR/?s=all&sp=4&pg=1

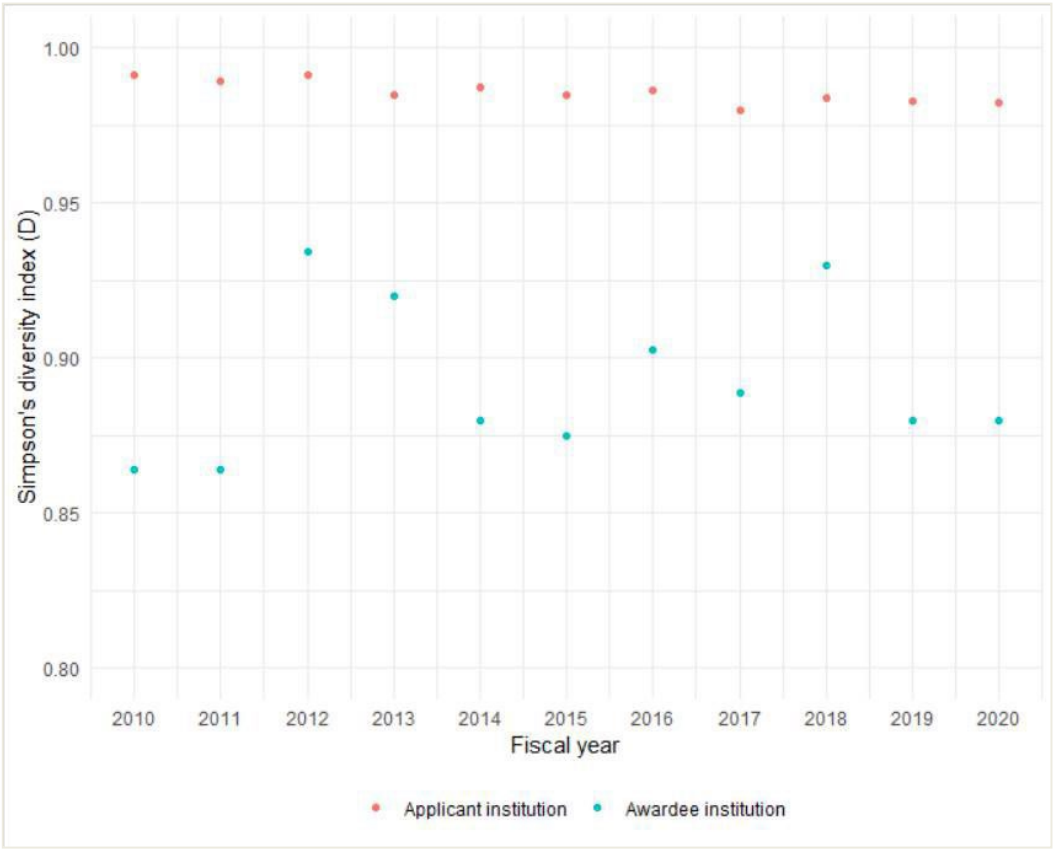


Figure 66. Simpson's Diversity Indices from FY2010 to FY2020 for Institutions with TRA Applicants and TRA Awardees

b. Comparison of FY2021 with Baseline (FY2010–2020) TRA Cohorts

1) Geographic Diversity

The 175 TRA contact PIs from FY2021 came from 102 unique institutions across 37 States and the District of Columbia (Figure 67). The top 5 States, by total applicant number, were California (30, 17%), Massachusetts (21, 12%), New York (15, 9%), Texas (10, 6%), and Missouri (8, 5%).[32] Similar to FY2010–2020, there were no TRA applicants from North Dakota in FY2021.

[32] Institution data was obtained from NIH's QVR database and are therefore not subject to the data use agreement regarding N < 11.

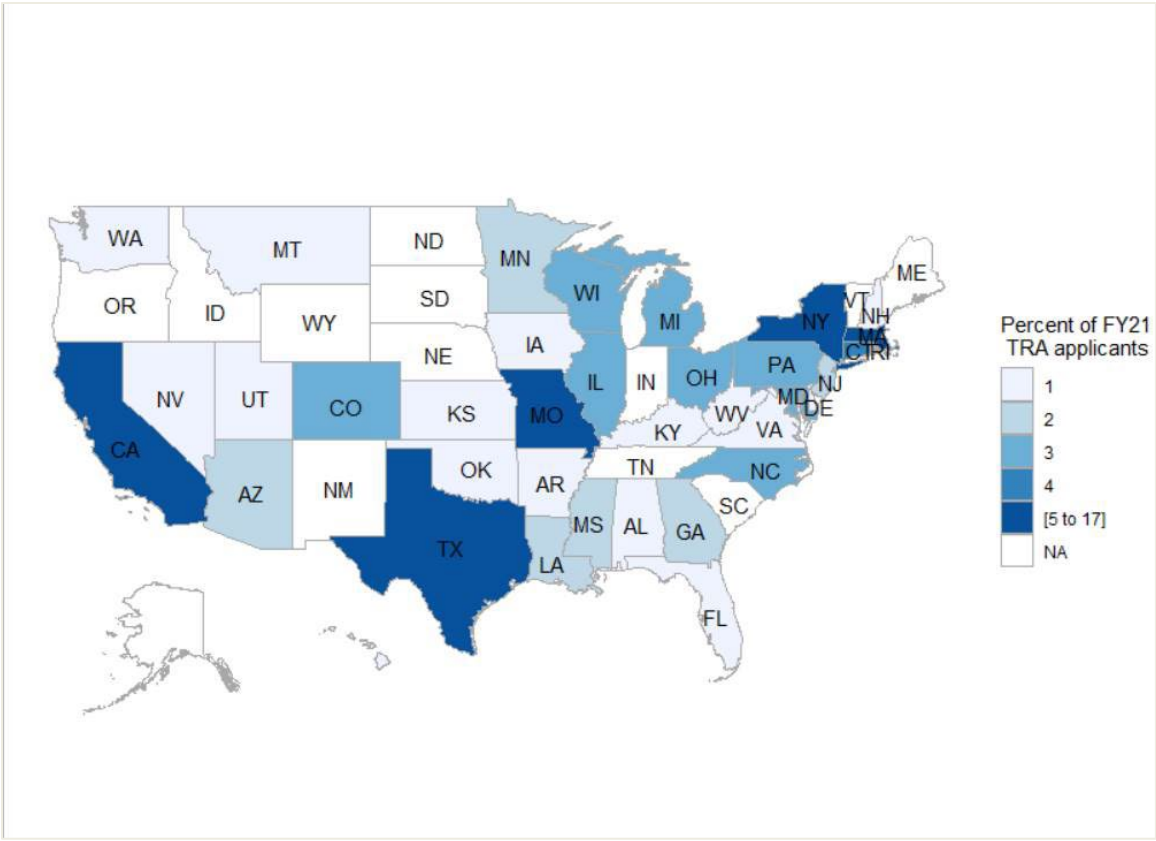


Figure 67. Percent of FY2021 TRA Applicants by State

The 10 contact PIs who had their TRA applications awarded in FY2021 came from 10 institutions across 8 States.

2) Institution Diversity

TRA applications from 10 new institutions were received in FY2021, one of which was an HBCU. Compared to FY2010–2020, the number of new institutions observed in FY2021 is comparable to what was observed in the previous 2 years (9 new institutions in both FY2019 and FY2020) (Figure 68). Similarly, Simpson's Diversity Index in FY2021 (0.99) was also within the range of what has been observed in previous years.

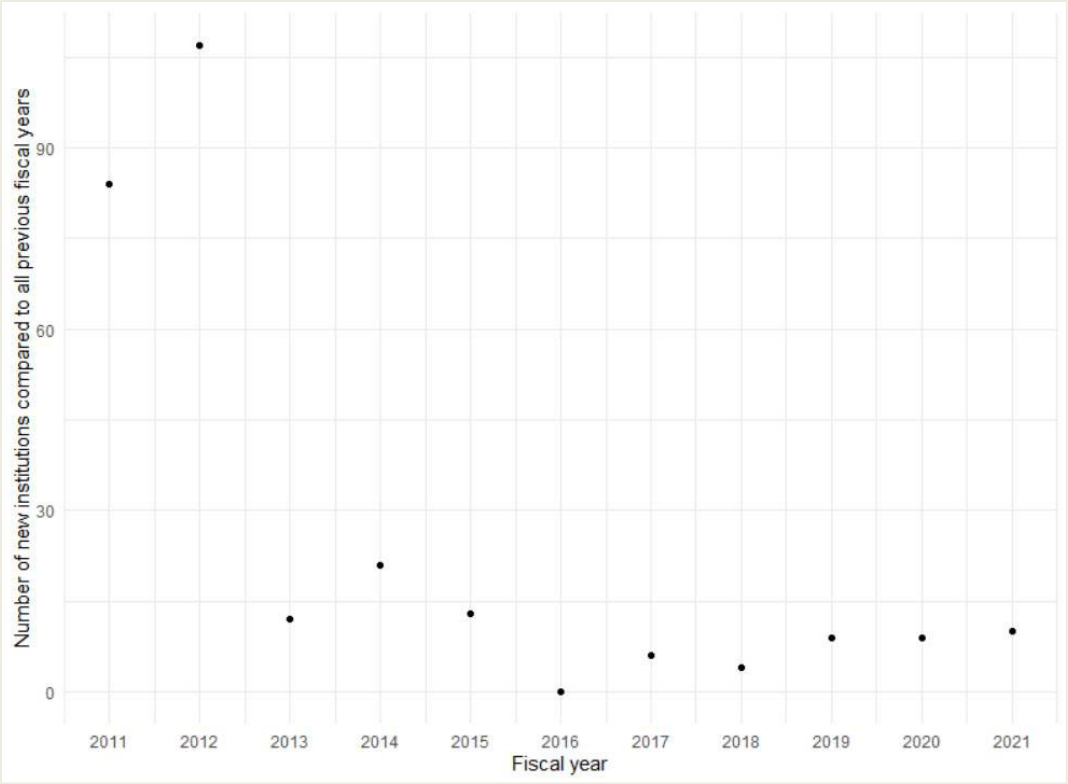


Figure 68. Number of New Institutions Represented by TRA Applicants between FY2011 and FY2021

4. Integration of Findings into Study Questions

Following data analysis, STPI integrated the core survey findings into the study questions and organized this section in two parts. The first is an assessment of anonymized review as it pertained to:

- Reviewers: Was anonymity maintained across the review? Did reviewers have sufficient information to perform a rigorous review? Sufficient information to review of transformative potential of research proposal?
- Applicants: Did anonymization impact the decision to submit an application? Did anonymization impact the ability to prepare a competitive application? Impact the ability to convey transformative research proposal?
- NIH: Are the mechanics of the review process sustainable?

Secondly, STPI examined the demographic and institution diversity of applications submitted and awards received. This assessment focused on changes in diversity as it pertained to:

- FY2021 TRA cohort: Did diversity increase across the phases of the TRA review cycle? Did diversity increase between applicants and funded applicants?
- FY2021 compared to FY2010-2020 TRA cohort: Did applicant or awardee diversity increase in FY2021 relative to the baseline TRA population from FY2010-2020?

A. Assessment of Anonymized Review

As noted in the introduction, the anonymized review process required the removal of any identifiers for the PI, collaborators, laboratory, or institution from the applications.[33] In addition, NIH constrained the anonymized information presented to the reviewers in Phase I (Specific Aims only), Phase II (Specific Aims and Research Strategy), and Phase IIIa (Specific Aims, Research Strategy, technical review). Only Phase IIIb reviewers had

[33] Section IV "names of individuals and institutions, honor and awards, hyperlinks, reference to any investigator attributes or accomplishments, citations that provide specific information about the source, and any other text from which the identity of any participating individual or institution can be reasonably inferred". See 2021 NIH Director's Transformative Research Awards FOA, available at https://grants.nih.gov/grants/guide/rfa-files/RFA-RM-20-013.html

access to the complete, de-anonymized applications; however, Phases IIIa and IIIb were surveyed together.

Through the survey mechanism, STPI obtained data in response to yes or no, select all, numeric, and free response questions. The breadth and variability of the free response questions precludes a thorough discussion here; however, those responses are used to further illustrate core findings, where appropriate.

1. Response Rates

Although the survey populations vary from 25 to 176 individuals, the response rates for the 6 surveys performed in this evaluation are all greater than 40% and exceed 60% for fours surveys and 80% for 2 surveys each (Table 6). These robust response rates reinforce the strength of the findings that STPI presents in this section.

Table 6. Response Rates for All Surveys

Responding Group	Review Phase	Potential Respondents	Number Responding	Response Rate
Applicants	Pre-phase Ia	176	105	62%
NIH staff	Pre-phase Ib	50	43	86%
Editorial Board	Phase I	27	12	44%
NIH staff	Pre-phase II	46	23	50%
Technical Reviewers	Phase II	149	100	67%
Editorial Board	Phase III	25	20	80%

2. Review

NIH asked STPI to assess whether anonymity was maintained and reviewers had sufficient information to perform a rigorous review and assess transformative potential of research proposals.

To determine whether anonymity was maintained across the review process, STPI queried the EB members twice (Phases I and III) and the technical reviewers once (Phase II). No EB survey respondents reported that they could identify an applicant, with 12 respondents reviewing only Specific Aims in Phase I and 19 reviewing the Specific Aims and Research Strategy and technical review for Phase IIIa (Table 7). STPI also notes the NIH efforts to maximize anonymity of the process prior to review by asking NIH staff to

assess the Specific Aims (Pre-phase I) and Research Strategy (Pre-phase II) for noncompliance with anonymization instructions in the FOA. Eleven of 176 applications were identified as noncompliant with anonymity instructions in Pre-phase I and five of 56 applications in Pre-phase II.

Twenty technical review survey respondents reported confidence in being able to identify the applicant, collaborators, laboratory, or institution. The 20 respondents reviewed 18 unique applications, indicating several instances of two or more reviewers noting identifying information in the same Specific Aims and Research Strategy. STPI followed up with 17 of the 20 respondents for whom we could find publicly available contact information. The 17 reviewers receiving the follow-up survey question reviewed 15 unique applications, and 8 follow-up survey respondents reviewed 8 different applications. Seven of the 8 were able to correctly identify the PI or institution. If one accepts that the 20 respondents correctly identified the PI or their institution, 18 of the 54 applications (33%) were not reviewed under anonymized conditions. STPI could confirm PI identification for 7 of 18 applications (39%). If one considers the 7 identified applications in the 54 applications reviewed in this phase, we could confirm technical reviewer identification of 13% of the applications reaching Phase II.

Table 7. Summary of Survey Respondents' Reports of PI Identification

Reviewer Group	Review Phase	Application Content	Number of Survey Responses	Number of Respondents Identifying an Applicant
Editorial Board	Phase I	Specific Aims	12	0
Technical Review	Phase II	Specific Aims, Research Strategy	100	20
Editorial Board	Phase IIIa	Specific Aims, Research Strategy, Technical Review	19	0

When asked what information allowed technical reviewers to identify the PI, most respondents cited the unique technologies, methodologies, resources, and/or specialized research topics or goals**.** These data also suggest that the subject matter expertise of technical reviewers may necessitate additional considerations for anonymization that are not needed for the senior generalist reviewers on the EB.

The next study question for reviewers focused on their ability to conduct a rigorous review with constrained, anonymized information and whether this impacted their ability to evaluate the transformative potential of the research proposal. Through yes/no and free response questions, EB respondents for Phase I (Specific Aims only) and Phase III

(Specific Aims, Research Strategy, technical review) reported sufficient and insufficient information to perform the review. The respondents noting insufficient information cited challenges to the review because they did not know, for example, the qualifications, experience, or publication record of the research team, feasibility of the research, collaborators, resources, or supporting data.

As subject matter experts, the technical reviewers are instructed to assess the anonymized Specific Aims and Research Strategy using the NIH criteria *significance, innovation, and if the research is logical and compelling* (NIH 2020b)—and although not a TRA criterion, feasibility. The questions posed to the technical reviewers reflect this more detailed assessment of review rigor (Table 8). More than 80% of the respondents reported that the anonymized Specific Aims and Research Strategy were sufficient to evaluate significance and innovation and that the approach was logical and compelling. Sixty-one percent reported that the anonymized Specific Aims and Research Strategy were sufficient to evaluate feasibility, a decrease likely attributable to the fact that feasibility is not an official review factor for TRA applications. Despite the high percentages of respondents reporting sufficient data for review, technical review respondents cited anonymization of experience, track record, and references as factors reducing review rigor.

In summary, most technical and EB reviewers considered the information available to them at each review steps sufficient to perform their review but expressed interest in having additional information to support their conclusions.

Table 8. Summary of Technical Review Responses for Sufficient Information

Reviewer Group	Review Phase	Application Content	Review Criteria	Number of Survey Respondents	Number Reporting Sufficient Information	Number Reporting Insufficient Information
Technical Review	Phase II	Specific Aims, Research Strategy	Significance	104	92	12
			Innovation	103	89	14
			Logical and compelling	104	85	19
			Feasibility*	104	64	40

*Not a formal TRA review criterion

To address the second half of this study question, STPI asked EB members about their ability to assess the transformative potential of proposed research for approximately 50 Specific Aims in Phase I and multiple applications in Phase III. EB members are considered senior scientists who have the ability to evaluate a wide range of scientific proposals.

Seven of 12 Phase I survey respondents indicated that they were able to assess the transformative potential of the proposed research using only the anonymized Specific Aims, while 5 reported they were unable to make this assessment for at least some of the applications they were assigned. Follow-up questions indicated that the inability to assess transformative potential applied to approximately 13 Specific Aims. It should be noted that

this was a retrospective survey, thus relying on the respondent's recall. Additionally, the survey format precludes determination if the same Specific Aims was considered difficult to review by more than one EB reviewer. For Phase III review, 17 of 19 respondents reported they could assess transformative potential in Phases IIIa (anonymized Specific Aims, Research Strategy, and technical review) and IIIb (complete, de-anonymized) applications, whereas 14 of the 19 reported that access to the complete application in Phase IIIb allowed them to better assess transformative potential compared to anonymized information in Phase IIIa.

Despite some respondents reporting an inability to assess transformative potential in Phases I and III, 11 of 12 Phase I respondents selected *somewhat* or *very confident* in their ability to determine transformative potential, and 17 of the 19 Phase III respondents indicated that they were *somewhat* or *very confident*. These data could indicate transition challenges for reviewers more accustomed to the traditional NIH review process rather than a flaw in the anonymized review process.

a. Secondary analysis

Given these overall findings, especially the survey results that 20 technical reviewers purported to identify the applicant, collaborators, laboratory or institution, STPI identified two additional questions to consider.

- Did the technical reviews for which a technical reviewer thought they knew the identity of the PI or institution include statements that could allow the Phase IIIA EB to identify the research or researcher?
- Was there consistency or divergence in the 3 technical reviews for those applications for which one or more technical reviewers thought they knew the research or researchers?

For this analysis, STPI reviewed the summary statements of the 18 applications for which a technical reviewer thought they knew the identity of the PI or institution. This analysis identified no explicit information in the technical reviews for the 18 applications that would identify the PI or institution to the Phase IIIa EB. Six of the reviews contained information that could reasonably be interpreted as indicating that the reviewer knew the researcher and research while not giving specific identifying information.

STPI next evaluated the three sections of the technical review (significance, innovation, approach) to determine if the overall content and tone of the reviews were, overall, consistent or divergent. Importantly, STPI did not know which review was provided by the technical reviewer identifying the PI or institution. In all 18 cases, while there is often a difference in the details in the review, technical reviewers produced similar overall comments for 2 of 3 or 3 of 3 of the reviews for an application. Overall, these data indicate that even though 18 of the 54 applications may not have been reviewed under fully

anonymized conditions, there is no evidence that the integrity of the anonymized process was compromised.

Examination of the content of the technical review indicates that technical reviewers were able to identify the research or researcher, not due to inclusion of information that should have been anonymized, but because the reviewers' technical breadth and depth of understanding of their field allowed them to recognize characteristics of the applicant, institution, or lab group, often unique technologies or research niche. The factor that gives technical reviewers unique insight on applications – their technical expertise – simultaneously poses a risk for identification of the PI and institution. This risk would not be solved with additional instructions either to the applicant or reviewer, unless that instruction was for the technical reviewer to disqualify themselves from the review. Additionally, STPI's analysis showing consistency across technical reviews suggests that even though some technical reviewers were able to identify the research or researcher, the integrity of the review was not compromised.

3. Applicants

NIH asked STPI to assess the impact of anonymized review on the applicants, specifically, whether anonymized review impacted their decision to submit an application and prepare a competitive transformative research proposal.

STPI determined that anonymized review was not a decision factor for 79 of 105 respondents submitting a TRA application. Twenty-six of the 105 respondents reported they applied because anonymized review had the potential to be less biased; citing diversity factors and institutional prestige, among others; and the potential for more focus on science and not the scientist.

TRA applicants were challenged in writing their applications by the omission of personal identifiers and the need to convey competitive, transformative research in stand-alone Specific Aims for review Phase I and the Specific Aims and Research Strategy in Phase II.[34] Because the Specific Aims constituted the single information source for Phase I review, respondents cited realignment of content to emphasize transformative potential, while others focused on making the Specific Aims more compelling to reviewers with a broad scientific background. Restrictions on the use of information that demonstrated expertise, feasibility, and unique capabilities and resources were reported as challenges to conveying transformative potential and competitiveness of the research for both parts of the application. The percentage of respondents who reported making changes to the

[34] Ninety-three applicants reported changes to their Specific Aims and 96 reported changes to their Research Strategies.

Specific Aims and Research Strategy did not differ between respondents who had, or had not, previously submitted a TRA application.

4. Sustainability

NIH asked STPI to assess whether the phased anonymized review process is sustainable. STPI has presented data in the previous sections that show that the review process was completed as designed. These data demonstrated the impact of anonymized review on the content and organization of the application and the sufficiency of information to conduct rigorous review. The high levels of confidence cited by the technical and Editorial Board respondents underscore this second finding. STPI also identified additional elements of the review process important to applicants and reviewers, and most likely to differ from the traditional NIH review. Instructions to the applicants, time to complete the review, and willingness of the reviewers to participate in additional anonymized reviews were assessed as additional indicators of sustainability.

STPI surveyed applicants about the NIH instructions in the FOA and webinar. Eighty-five of 104 applicant respondents said the FOA instructions were sufficient to prepare their application with 56 of 99 respondents citing the NIH webinar, institutional resources, collaborators and mentors as supplements to the FOA instructions. Despite the high percentages of applicants and reviewers reporting sufficient information for their respective component of the process, additional information was requested at all stages of the review. For the most part, the requests for additional information requested examples that would clarify anonymization requirements. Applicants reported interest in examples of anonymized information (preliminary data, previous data, collaborators, unique technologies and methods) and technical reviewers requested examples of anonymized applications. All reviewers had a wish list of additional information for each review step, and summing all interests, the list invoked the information that was anonymized or constrained during Phases I - IIIa.

STPI queried Editorial Board members about the difficulty of the Phase I review in which they each reviewed 50 anonymized Specific Aims; Phase IIIa in which they reviewed applications with anonymized, constrained data; and Phase IIIb in which they reviewed the entire application and technical reviews. Overall, Editorial Board respondents reported that anonymized review was similar to or easier than traditional NIH review (8 of the 12 Phase 1 respondents and 12 of 20 Phase III respondents). Respondents who indicated that Phase IIIa anonymized review was *more difficult* or *much more difficult* reported, for example, the review took more time, they were assigned applications outside of their area of expertise, and they did not know the collaborators or resources. Because time commitment is a sustainability factor, the time NIH staff took to perform reviews was queried in the surveys. NIH staff performing administrative review reported spending about 15 minutes reviewing each Specific Aims and about 40 minutes reviewing each

Research Strategies. Most importantly, the majority of reviewers expressed interest in participating in anonymized reviews in the future.

B. Diversity of Applicants

To assess whether the anonymized review process will increase the diversity of TRA applicants and awardees, STPI examined the demographic, and institution and geographic diversity of

- applications submitted and awards received
- changes in diversity across the phases of the FY2021 TRA review cycle
- TRA FY2021 applicants and awardees compared to TRA FY2010 to FY2020 cohorts

a. Demographic Diversity

1) Applications Submitted and Awards Received

For FY2021, no significant differences were detected in the breakdown of applicants and awardees by gender, race, or ethnicity.

2) Changes in Diversity Across the Phases of the FY2021 TRA Review Cycle

For FY2021, no significant difference was detected in the breakdown by gender, race, and ethnicity between the percentage of applicants and those who made it to Phase III, between those who made it to Phase III and those who were funded, or between applicants and funded applicants. However, given the number of applicants who made it to Phase III and the number of funded applicants, there was low power to detect differences between the different phases and those who were funded.

For gender, final scores did not significantly differ between female and male applicants, but applicants who were categorized as Other had significantly higher reviewer scores than both female and male applicants. For race, White applicants had significantly lower final scores compared to all other races; Asian applicants did not have significantly different final scores from Other applicants but had significantly lower final scores than Black or African American applicants; and Black or African American applicants did not have significantly different final scores than Other applicants. For ethnicity, Hispanic or Latino applicants had significantly higher final scores than not Hispanic or nor Latino and Other applicants; and there was no significant difference in final scores between not Hispanic or not Latino and Other applicants.

3) *TRA FY2021* Applicants and Awardees Compared to *TRA FY2010–2020* Cohort

STPI considered diversity to have increased for a demographic factor (i.e., gender, race, and ethnicity) in FY2021 if at least

- one non-dominant group within a demographic factor (i.e, female or Other for gender) increased in percentage relative to FY2010–2020 or
- the dominant group within a demographic factor (i.e., males for gender, Whites for race, and Not Hispanic or Not Latino for ethnicity) decreased in percentage relative to FY2010–2020.

Overall, applicant diversity increased for gender, race, and ethnicity in FY2021 relative to FY2010–2020. With respect to awardees, gender and ethnicity diversity did not increase in FY2021 relative to FY2010-2020 but diversity did increase for race. Specifically, the percentage of Black or African American awardees in FY2021 was significantly higher than that of FY2010-2020. However, given that the total number of Black or African American TRA awardees is small, this one-year observation may not indicate a long-term trend.

b. Institution and Geographic Diversity

The 175 TRA contact PIs from FY2021 came from 102 unique institutions across 37 States and the District of Columbia. No new States were represented by TRA applicants in FY2021. The number of new institutions observed in FY2021 (10) is comparable to what was observed in the previous 2 years (9 each in 2019 and FY2020). In addition, institution diversity as measured by Simpson's Diversity Index for FY2021 was within the range of what has been observed in previous years, suggesting that institution diversity did not increase in FY2021 relative to FY2010–2020. Among the 10 contact PIs who had their TRA applications funded, 2 are from institutions that had never had a contact PI apply to the TRA program prior to FY2021.

5. Final Considerations

The goal of NIH review is to employ a competitive peer-review process that is fair, independent, expert, and timely to identify the most promising biomedical research (NIH Center for Scientific Review 2021). The goal of anonymized review is to increase the diversity of the NIH applicants and awardees by shielding from reviewers any information in the applications that would identify the PI, collaborators, laboratory, or institution.

As a first consideration, a competitive process should attract a robust number of applications. The FY 2021 TRA FOA received 176 applications, which is not significantly different from what has been observed in the previous four fiscal years (FY2016-FY2020). However, in the FY2021 review, 16 applications were identified as noncompliant with anonymization instructions during NIH staff reviews (9% of all applications submitted). The potentially transformative research contained in these proposals was removed from funding consideration due to the anonymization process, although the applicants could seek another NIH grant mechanism with traditional review. Additionally, during technical review, 20 reviewers believed they knew the identity of the PI (17 of 54 applications) and still provided a review for Phase III. Given the novelty of the anonymized review process and its implications for drafting an application, a noncompliance withdrawal rate less than 10% is not unreasonable; however, NIH may wish to consider the management of technical reviews which were not conducted under fully anonymized conditions.

The EB and technical review phases of the anonymized process incorporated a second consideration: independent and expert review. Although analysis of the actual reviews was beyond the scope of this assessment, most EB and technical reviewers reported that they could conduct rigorous reviews of anonymized and constrained information at each step of the process and were confident of their efforts. Redundancy in the review process through the assignment of multiple reviewers to each application also provides evidence for rigorous review.

Anonymization was incorporated into this review to address the issue of fairness raised, in part, by publications demonstrating lack of diversity in the NIH awardee population. For FY2021, no significant differences were detected in the applicants and awardees for gender, race, or ethnicity, suggesting that an increase in applicant diversity could reasonably be expected to translate to the awardee population. STPI analysis identified more applicant diversity for gender, race, and ethnicity in FY2021 relative to FY2010–2020. For awardees, gender and ethnic diversity did not increase in FY2021 relative to FY2010-2020; however, there was greater diversity for race. This finding

indicates that changes in applicant diversity for race, but not gender and ethnicity, were reflected in the awardee cohort. Whether this finding indicates a long-term trend is not clear. Neither geographic or institutional diversity, as measured in this study, increased in FY2021 or in FY2021 relative to FY2010–20.

This report documents an assessment of the first use of an anonymized review process for a Common Fund HRHR initiative. The data confirm that the anonymized process meets the NIH review criteria, and there is evidence for increased diversity for applicants. Because the number of awardees stratified by demographic, institutional, and geographic diversity is small, robust observations of diversity changes will require pooled data from multiple review cycles.

Appendix A. Applicant Submission Criteria for TRA FOA RFA-RM-20-013[35]

Applicants were instructed to not list specific objectives of the proposed research, but to instead describe the overall project and why it is well aligned with the objectives of the TRA award. The applicants were instructed to include the following two sections in their applications:

Significance, Innovation, and Impact

What is the challenge or opportunity that is the focus of your proposed research? Why is this broadly significant? What is the overall approach you are proposing? What are the most innovative aspects of your application? If successful, what would the impact be on our scientific understanding and (ultimately) human health?

Insight and Rationale

What is the fundamental new insight that is motivating the proposed research? What is the underlying logic or rationale that provides support for pursuing this insight despite little or no preliminary data?

The applicants were also instructed to withhold the following information from both the Specific Aims and the Research Strategy:

- Names of any individuals or institutions
- Mention of any honors or awards
- Hyperlinks
- Reference to any investigator attributes or accomplishments, such as "as leaders in this field" or "we have shown"
- Citations that provide specific information about the source. Use numeric citations only, which refer to the corresponding source in the "Bibliography & References Cited" component of the application.
- Any other text from which the identity of any participating individual or institution can be reasonably inferred

[35] See 2020 NIH Director's Transformative Research Award FOA, available at: https://grants.nih.gov/grants/guide/rfa-files/RFA-RM-20-013.html

If the applicant does include this information, their application will be administratively withdrawn.

Appendix B. Review Criteria for TRA FOA RFA-RM-20-013[36]

Reviewers consider the following main review criteria[37] to determine the scientific merit of the applications and give the applicant a score for each:

Significance:

- Does the project address an important problem or a critical barrier to progress in the field? Is the prior research that serves as the key support for the proposed project rigorous? If the aims of the project are achieved, how will scientific knowledge, technical capability, and/or clinical practice be improved? How will successful completion of the aims change the concepts, methods, technologies, treatments, services, or preventative interventions that drive this field?
- Does the study have clear transformative potential? Is the proposed research exceptional in terms of its consequences for the field and size of the community affected? Is the paradigm being challenged or proposed fundamental to the field?

Investigators[38]

- Are the PD(s)/PI(s), collaborators, and other researchers well suited to the project? If early stage investigators or those in the early stages of independent careers, do they have appropriate experience and training? If established, have they demonstrated an ongoing record of accomplishments that have advanced their field(s)? If the project is collaborative or multi-PD/PI, do the investigators have complementary and integrated expertise; are their leadership approach, governance and organizational structure appropriate for the project?
- Do the PD(s)/PI(s) commit sufficient effort to make the proposed research a priority?
- With regard to the proposed leadership for the project, do the PD(s)/PI(s) and key personnel have the expertise, experience, and ability to organize, manage and implement the proposed clinical trial and meet milestones and timelines? Do they have appropriate expertise in study coordination, data management and statistics?

---

[36] See 2020 NIH Director's Transformative Research Award FOA, available at: https://grants.nih.gov/grants/guide/rfa-files/RFA-RM-20-013.html

[37] Main criteria without nuances for clinical trials listed. Additional criteria can be found

[38] The Investigators criteria is only considered in Phase IIIb.

For a multicenter trial, is the organizational structure appropriate and does the application identify a core of potential center investigators and staffing for a coordinating center?

Innovation

- Does the application challenge and seek to shift current research or clinical practice paradigms by utilizing novel theoretical concepts, approaches or methodologies, instrumentation, or interventions? Are the concepts, approaches or methodologies, instrumentation, or interventions novel to one field of research or novel in a broad sense? Is a refinement, improvement, or new application of theoretical concepts, approaches or methodologies, instrumentation, or interventions proposed?

Approach

- Are the overall strategy, methodology, and analyses well-reasoned and appropriate to accomplish the specific aims of the project? Have the investigators included plans to address weaknesses in the rigor of prior research that serves as the key support for the proposed project? Have the investigators presented strategies to ensure a robust and unbiased approach, as appropriate for the work proposed? Are potential problems, alternative strategies, and benchmarks for success presented? If the project is in the early stages of development, will the strategy establish feasibility and will particularly risky aspects be managed? Have the investigators presented adequate plans to address relevant biological variables, such as sex, for studies in vertebrate animals or human subjects?
- Is the logic of the approach compelling despite the lack of experimental details and substantial preliminary data? Is there evidence that the investigators will pursue the project, despite its inherent risks, in a robust, reproducible, and rigorous manner? Does the information in the timeline inspire confidence that the PD(s)/PI(s) will be able to assess progress in each year of the award and either complete the project or demonstrate conclusively that it cannot be completed, despite good-faith efforts, during the term of the award?

Environment

- Will the scientific environment in which the work will be done contribute to the probability of success? Are the institutional support, equipment and other physical resources available to the investigators adequate for the project proposed? Will the project benefit from unique features of the scientific environment, subject populations, or collaborative arrangements?

Appendix C.Applicant Survey

Thank you for taking part in this study conducted by the IDA Science and Technology Policy Institute (STPI) on behalf of the National Institutes of Health Office of the Director (NIH/OD) and the Center for Scientific Review (NIH/CSR). STPI is a federally funded research and development center that provides rigorous, independent research and analysis to the Federal Government.

Purpose of the Survey

This survey solicits your perspectives on preparing an anonymized application for the Transformative Research Award.

Confidentiality Statement

STPI is independent of NIH and has been contracted to collect these data. All responses will be kept confidential and protected to the extent possible by law. Only aggregate data will be presented to NIH. Your decision to participate is voluntary and will have no effect on your current or future relationship with the agency.

The estimated survey completion time is 25-30 minutes. You will be able to move backward through the survey to review or edit responses. Your survey responses are automatically saved up to the last submitted page, so you will be able to pause and return mid-survey. However, once you submit the survey, you will not be able to edit your responses. While completing this survey, you will be asked several questions about your NIH application. You should only consider your TRA award: [invite('custom 1')] when answering these questions.

If you would like to review the relevant TRA Funding Opportunity Announcement (FOA), please click on the following: 2020 TRA FOA

Inquiries and Concerns

If you have questions or concerns about completing this survey, please contact us at anonTRAreview@ida.org.

Thank you for your participation

Your responses are invaluable to the study.

Piping: Piped values from spreadsheet with Applicant application numbers.

1. **Please answer the question below for your TRA application:** ***[invite('custom 1')]***

The 2020 TRA FOA describes an anonymized review process for applications. This process requires applicants to develop Specific Aims and a Research Strategy that do not disclose information identifying the applicant, laboratory, or institution. This survey assesses the applicant's experience in preparing an anonymized application according to the instructions in the FOA.

Please indicate if use of an anonymized review impacted your decision to submit your application.

( ) Yes, anonymized review impacted my decision to submit my application.

( ) No, anonymized review did not impact my decision to submit my application.

Logic: Hidden unless: #1 Question "Please indicate if use of an anonymized review impacted your decision to submit your application." is one of the following answers ("Yes, anonymized review impacted my decision to submit my application.")

Piping: Piped values from spreadsheet with Applicant application numbers.

2. **Please answer the question below for your TRA application:** ***[invite('custom 1')]***

Please describe how anonymized review impacted your decision to submit your application.

_____

_____

_____

Piping: Piped values from spreadsheet with Applicant application numbers.

3. **Please answer the question below for your TRA application:** ***[invite('custom 1')]***

Please indicate if the FOA instructions you received from NIH were sufficient to prepare your application.

( ) Yes, the FOA instructions I received from NIH were sufficient to prepare my application.

( ) No, the FOA instructions I received from NIH were not sufficient to prepare my application.

Logic: Hidden unless: #3 Question "Please indicate if the FOA instructions you received from NIH were sufficient to prepare your application." is one of the following answers ("Yes, the FOA instructions I received from NIH were sufficient to prepare my application.")

Piping: Piped values from spreadsheet with Applicant application numbers.

4. **Please answer the question below for your TRA application: *[invite('custom 1')]***

   Please describe what elements of the instructions in the FOA were helpful in developing your anonymized application.

   _____

   _____

   _____

Logic: Hidden unless: #3 Question "Please indicate if the FOA instructions you received from NIH were sufficient to prepare your application." is one of the following answers ("No, the FOA instructions I received from NIH were not sufficient to prepare my application.")

Piping: Piped values from spreadsheet with Applicant application numbers.

5. **Please answer the question below for your TRA application: *[invite('custom 1')]***

   What additional information would have been useful to help you prepare an anonymized application?

   _____

   _____

   _____

Piping: Piped values from spreadsheet with Applicant application numbers.

6. **Please answer the question below for your TRA application: *[invite('custom 1')]***

   Please indicate which of the additional NIH resources, if any, you used to prepare your application. (select all that apply).

   ( ) Webinar

   ( ) NIH Program Officer

( ) Other (please specify): _____

Logic: Hidden unless: #6 Question "Please indicate which of the additional NIH resources, if any, you used to prepare your application. (select all that apply)." is one of the following answers ("Webinar","NIH Program Officer","Other (please specify)")

Piping: Piped values from spreadsheet with Applicant application numbers.

7. **Please answer the question below for your TRA application:** ***[invite('custom 1')]***

Please describe what additional information these resources provided that helped you develop your anonymized application.

_____

_____

_____

Piping: Piped values from spreadsheet with Applicant application numbers.

8. Please answer the question below for your TRA application: ***[invite('custom 1')]***

In step one of the review an NIH administrative committee assesses the Specific Aims for compliance with the anonymization instructions in the FOA, and those not in compliance are administratively withdrawn. Next, a Center for Scientific Review (CSR) Editorial Board assesses and ranks the Specific Aims based on scientific and technical merit and transformative research potential.

In the following set of questions, three aspects of anonymized review that could impact an investigator's ability to write competitive Specific Aims are considered: anonymization, transformative potential, and use of only Specific Aims in step one of the review process.

Specific Aims: Anonymization

Please indicate if you made changes in your Specific Aims to comply with the anonymization instructions.

( ) Yes, I made changes in my Specific Aims to comply with the anonymization instructions.

( ) No, I did not make changes in my Specific Aims to comply with the anonymization instructions.

Logic: Hidden unless: #8 Question **"Specific Aims: Anonymization**
Please indicate if you made changes in your Specific Aims to comply with the anonymization instructions." is one of the following answers ("Yes, I made changes in my Specific Aims to comply with the anonymization instructions.")

Piping: piped values from spreadsheet with Applicant application numbers.

9. Please answer the question below for your TRA application: [invite('custom 1')]

Please describe the changes you made to your Specific Aims to comply with the anonymization instructions.

_____

_____

_____

_____

Piping: piped values from spreadsheet with Applicant application numbers.

10. Please answer the question below for your TRA application: [invite('custom 1')]

Specific Aims: Transformative Potential
Please indicate if the anonymization instructions affected your ability to convey the transformative potential of your research in the Specific Aims.

( ) Yes, the anonymization instructions affected my ability to convey the transformative potential of my research in the Specific Aims.

( ) No, the anonymization instructions did not affect my ability to convey the transformative potential of my research in the Specific Aims.

Logic: Hidden unless: #10 Question "Specific Aims: Transformative Potential Please indicate if the anonymization instructions affected your ability to convey the transformative potential of your research in the Specific Aims." is one of the following answers ("Yes, the anonymization instructions affected my ability to convey the transformative potential of my research in the Specific Aims.")

Piping: Piped values from spreadsheet with Applicant application numbers.

11. Please answer the question below for your TRA application: [invite('custom 1')]

Specific Aims: Transformative Potential
Please describe how the anonymization instructions affected your ability to convey the transformative potential of your research in the Specific Aims.

_____

_____

_____

_____

Piping: Piped Values From spreadsheet with Applicant application numbers.

12. Please answer the question below for your TRA application: [invite('custom 1')]

Specific Aims: Review
The Specific Aims were the only part of the application considered by the Editorial Board in ranking applications to advance to the next step of the review process. Did the information that only the Specific Aims were used in step one of the review change how you wrote the Specific Aims?

( ) Yes, this information changed how I wrote the Specific Aims.

( ) No, this information did not change how I wrote the Specific Aims.

Page entry logic: This page will show when: #12 Question "**Specific Aims: Review** The Specific Aims were the only part of the application considered by the Editorial Board in ranking applications to advance to the next step of the review process. Did the information that only the Specific Aims were used in step one of the review change how you wrote the Specific Aims?" is one of the following answers ("Yes, this information changed how I wrote the Specific Aims.")

Piping: Piped Values From spreadsheet with Applicant application numbers.

13.* Please answer the question below for your TRA application: *[invite('custom 1')]

Specific Aims: Review
Please describe the changes you made to the Specific Aims knowing that only the Specific Aims were used in step one of the review.

_____

_____

_____

_____

Piping: Piped values from spreadsheet with Applicant application numbers.

14. Please answer the question below for your TRA application: *[invite('custom 1')]*

In step two of the review, the Research Strategy is assessed by an NIH administrative committee for compliance with the anonymization instructions in the FOA, and those not in compliance are administratively withdrawn. Technical reviewers then assess the Specific Aims and Research Strategies of top-ranked applications.

In the following set of questions, two aspects of anonymized review that could impact an investigator's ability to write a competitive Research Strategy are considered: anonymization and transformative potential.

Research Strategy: Anonymization
Please indicate if you made changes in your Research Strategy to comply with the anonymization instructions.

( ) Yes, I made changes in my Research Strategy to comply with the anonymization instructions.

( ) No, I did not make changes in my Research Strategy to comply with the anonymization instructions

Logic: Hidden unless: #14 Question "Research Strategy: Anonymization Please indicate if you made changes in your Research Strategy to comply with the anonymization instructions." is one of the following answers ("Yes, I made changes in my Research Strategy to comply with the anonymization instructions.")

Piping: Piped values from spreadsheet with Applicant application numbers.

15. Please answer the question below for your TRA application: *[invite('custom 1')]*

Research Strategy: Anonymization
Please describe the changes you made to your Research Strategy to comply with the anonymization instructions.

_____

_____

_____

_____

Piping: Piped values from spreadsheet with Applicant application numbers.

16. Please answer the question below for your TRA application: *[invite('custom 1')]*

Research Strategy: Transformative Potential
Please indicate if the anonymization instructions affected your ability to convey the transformative potential of your research in the Research Strategy.

( ) Yes, the anonymization instructions affected my ability to convey the transformative potential of my research in the Research Strategy.

( ) No, the anonymization instructions did not affect my ability to convey the transformative potential of my research in the Research Strategy.

Page entry logic: This page will show when: #16 Question "**Research Strategy: Transformative Potential**
Please indicate if the anonymization instructions affected your ability to convey the transformative potential of your research in the Research Strategy." is one of the following answers ("Yes, the anonymization instructions affected my ability to convey the transformative potential of my research in the Research Strategy.")

Piping: Piped values from spreadsheet with Applicant application numbers.

17. Please answer the question below for your TRA application: *[invite('custom 1')]*

Research Strategy: Transformative Potential
Please describe how the anonymization instructions affected your ability to convey the transformative potential of your research in the Research Strategy.

_____

_____

_____

_____

Piping: Piped values from spreadsheet with Applicant application numbers.

18. Please answer the question below for your TRA application: *[invite('custom 1')]*

Please indicate the number of times that you have submitted or won a TRA award prior to 2020.

	0	**1**	**2**	**3**	**4**	**5 or more**
Number of previous TRA applications submitted	( )	( )	( )	( )	( )	( )
Number of previous TRA applications funded	( )	( )	( )	( )	( )	( )

Piping: piped values from spreadsheet with Applicant application numbers.

19. Please answer the question below for your TRA application: *[invite('custom 1')]*

Please indicate if your institution provided any guidance or resources to assist you in writing your anonymized TRA application.

( ) Yes, my institution provided guidance or resources to assist me in writing an anonymized application.

( ) No, my institution did not provide guidance or resources to assist me in writing an anonymized application.

Page entry logic: This page will show when: #19 Question "Please indicate if your institution provided any guidance or resources to assist you in writing your anonymized TRA application." is one of the following answers ("Yes, my institution provided guidance or resources to assist me in writing an anonymized application.")

Piping: Piped values from spreadsheet with Applicant application numbers.

20. Please answer the question below for your TRA application: *[invite('custom 1')]*

Please describe the specific guidance or resources provided by your institution to assist you in writing your anonymized TRA application.

_____

_____

_____

_____

Piping: Piped values from spreadsheet with Applicant application numbers.

21.* Please answer the question below for your TRA application: *[invite('custom 1')]

Please indicate if your application had one or more collaborator(s).

( ) Yes, my application had one or more collaborator(s).

( ) No, my application did not have one or more collaborator(s).

Page entry logic: This page will show when: #21 Question "Please indicate if your application had one or more collaborator(s)." is one of the following answers ("Yes, my application had one or more collaborator(s).")

Piping: Piped values from spreadsheet with Applicant application numbers.

22. Please answer the question below for your TRA application: *[invite('custom 1')]*

You indicated that you have collaborators in your application. Please select how many collaborators were included in your application.

( ) 1

( ) 2

( ) 3

( ) 4 or more

Page entry logic: This page will show when: #21 Question "Please indicate if your application had one or more collaborator(s)." is one of the following answers ("Yes, my application had one or more collaborator(s).")

Piping: Piped values from spreadsheet with Applicant application numbers.

23. **Please answer the question below for your TRA application:** ***[invite('custom 1')]***

Please indicate if anonymizing the identity of your collaborator(s) was challenging.

( ) Yes, anonymizing the identity of my collaborator(s) was challenging.

( ) No, anonymizing the identity of my collaborator(s) was not challenging.

Logic: Hidden unless: #23 Question "Please indicate if anonymizing the identity of your collaborator(s) was challenging." is one of the following answers ("Yes, anonymizing the identity of my collaborator(s) was challenging.")

Piping: Piped values from spreadsheet with Applicant application numbers.

24. Please answer the question below for your TRA application: ***[invite('custom 1')]***

Collaborator(s): Specific Aims
Please describe the information that was challenging to anonymize in your Specific Aims because you had collaborators.

_____

_____

_____

_____

Logic: Hidden unless: #23 Question "Please indicate if anonymizing the identity of your collaborator(s) was challenging." is one of the following answers ("Yes, anonymizing the identity of my collaborator(s) was challenging.")

Piping: Piped values from spreadsheet with Applicant application numbers.

25. **Please answer the question below for your TRA application: [invite('custom 1')]**

Collaborator(s): Research Strategy
Please describe the information that was challenging to anonymize in your Research Strategy because you had collaborators.

_____

_____

_____

_____

Piping: Piped values from spreadsheet with Applicant application numbers.

26. Please answer the question below for your TRA application: *[invite('custom 1')]*
Preliminary data are not required for TRA applications; however, they may be included at the applicant's discretion.

Please indicate if you included any preliminary data in your application.

( ) Yes, I included preliminary data in my application.

( ) No, I did not include preliminary data in my application.

Page entry logic: Hide unless: #26 Question "Please indicate if you included any preliminary data in your application." is one of the following answers ("Yes, I included preliminary data in my application.")

Piping: Piped values from spreadsheet with applicant application numbers.

27. **Please answer the question below for your TRA application: *[invite('custom 1')]***

Please indicate if anonymized review influenced your decision to include preliminary data in your application.

( ) Yes, anonymized review influenced my decision to include preliminary data in my application.

( ) No, anonymized review did not influence my decision to include preliminary data in my application.

Logic: Hidden unless: #26 Question "Please indicate if you included any preliminary data in your application." is one of the following answers ("Yes, I included preliminary data in my application.")

Piping: Piped values from spreadsheet with applicant application numbers.

28. Please answer the question below for your TRA application: *[invite('custom 1')]*

Please indicate if anonymized review influenced the amount or type of preliminary data included in your application.

( ) Yes, anonymized review influenced the amount or type of preliminary data included in my application.

( ) No, anonymized review did not influence the amount or type of preliminary data included in my application.

Logic: Hidden unless: #28 Question "Please indicate if anonymized review influenced the amount or type of preliminary data included in your application." is one of the following answers ("Yes, anonymized review influenced the amount or type of preliminary data included in my application.")

Piping: Piped values from spreadsheet with applicant application numbers.

29.* Please answer the question below for your TRA application: *[invite('custom 1')]

Please describe how anonymized review influenced the amount or type of preliminary data included in your application.

_____

_____

_____

_____

Page entry logic: This page will show when: #26 Question "Please indicate if you included any preliminary data in your application." is one of the following answers ("No, I did not include preliminary data in my application.")

Piping: Piped values from spreadsheet with applicant application numbers.

30. Please answer the question below for your TRA application: *[invite('custom 1')]*

Please indicate if anonymized review influenced your decision to exclude preliminary data from your application.

( ) Yes, anonymized review influenced my decision to exclude preliminary data

( ) No, anonymized review did not influence my decision to exclude preliminary data

---

Page entry logic: This page will show when: #30 Question "Please indicate if anonymized review influenced your decision to exclude preliminary data from your application." is one of the following answers ("Yes, anonymized review influenced my decision to exclude preliminary data")

Piping: Piped values from spreadsheet with applicant application numbers.

31. Please answer the question below for your TRA application: *[invite('custom 1')]*

Please describe how anonymized review influenced your decision to exclude preliminary data from your application.

_____

_____

_____

_____

Piping: Piped values from spreadsheet with applicant application numbers.

32. Please answer the question below for your TRA application: *[invite('custom 1')]*

Are there any additional comments you would like to provide to NIH about the anonymized TRA application process?

_____

_____

_____

_____

Please be advised that this is the last question in the survey. Once you submit your survey, you will not be able to go back and change your responses.

Thank you for taking our survey. The aggregated results will assist NIH in continuing to offer programs that support transformative research.

Appendix D. Applicant Survey Data

Table 1. Impact of Anonymized Review on Decision to Apply

Variable of interest	Answer options within variable of interest	Total number (percent) of respondents for answer option	p
Please indicate if use of an anonymized review impacted your decision to submit your application	No, anonymized review did not impact my decision to submit my application.	79 (75.2%)	<0.001
	Yes, anonymized review impacted my decision to submit my application.	26 (24.8%)	

Table 2. Sufficiency of Information Provided

Variable of interest	Answer options within variable of interest	Total number (percent) of respondents for answer option	p
Please indicate if the FOA instructions you received from NIH were sufficient to prepare your application	No, the FOA instructions I received from NIH were not sufficient to prepare my application.	19 (18.3%)	<0.001
	Yes, the FOA instructions I received from NIH were sufficient to prepare my application.	85 (81.7%)	
Please indicate if your institution provided any guidance or resources to assist you in writing your anonymized TRA application	No, my institution did not provide guidance or resources to assist me in writing an anonymized application.	84 (80.0%)	<0.001
	Yes, my institution provided guidance or resources to assist me in writing an anonymized application.	21 (20.0%)	

Table 3. Sufficiency of Information Provided:

Please indicate which of the additional NIH resources, if any, you used to prepare your application. (select all that apply).

Item of interest	Total number (percent) of respondents for answer option	Cochran's Q (df)	p
NIH Program Officer	39 (37.9%)	11.68 (df = 2)	0.003
Other	27 (26.2%)		
Webinar	54 (52.4%)		

Table 4. Changes to Specific Aims & Research Strategy to Anonymize

Variable of interest	Answer options within variable of interest	Total number (percent) of respondents for answer option	p
Please indicate if you made changes in your Specific Aims to comply with the anonymization instructions	No, I did not make changes in my Specific Aims to comply with the anonymization instructions.	12 (11.4%)	<0.001
	Yes, I made changes in my Specific Aims to comply with the anonymization instructions.	93 (88.6%)	
Please indicate if you made changes in your Research Strategy to comply with the anonymization instructions	No, I did not make changes in my Research Strategy to comply with the anonymization instructions	9 (8.6%)	<0.001
	Yes, I made changes in my Research Strategy to comply with the anonymization instructions.	96 (91.4%)	

Table 5. Changes to Specific Aims & Research Strategy to Convey Transformative Potential

Variable of interest	Answer options within variable of interest	Total number (percent) of respondents for answer option	p
Please indicate if the anonymization instructions affected your ability to convey the transformative potential of your research in the Specific Aims	No, the anonymization instructions did not affect my ability to convey the transformative potential of my research in the Specific Aims.	73 (69.5%)	<0.001
	Yes, the anonymization instructions affected my ability to convey the transformative potential of my research in the Specific Aims.	32 (30.5%)	
Please indicate if the anonymization instructions affected your ability to convey the transformative potential of your research in the Research Strategy	No, the anonymization instructions did not affect my ability to convey the transformative potential of my research in the Research Strategy.	65 (62.5%)	0.014
	Yes, the anonymization instructions affected my ability to convey the transformative potential of my research in the Research Strategy.	39 (37.5%)	

Table 6. Changes to Specific Aims Given that Specific Aims was Sole Phase I Consideration

Variable of interest	Answer options within variable of interest	Total number (percent) of respondents for answer option	p
The Specific Aims were the only part of the application considered by the Editorial Board in ranking applications to advance to the next step of the review process. Did the information that only the Specific Aims were used in step one of the review change how you wrote the Specific Aims?	No, this information did not change how I wrote the Specific Aims.	52 (50.0%)	1
	Yes, this information changed how I wrote the Specific Aims.	52 (50.0%)	

Table 7. Preliminary Data

Variable of interest	Answer options within variable of interest	Total number (percent) of respondents for answer option	p
Please indicate if you included any preliminary data in your application	No, I did not include preliminary data in my application.	12 (11.5%)	<0.001
	Yes, I included preliminary data in my application.	92 (88.5%)	
Please indicate if anonymized review influenced your decision to include preliminary data in your application	No, anonymized review did not influence my decision to include preliminary data in my application.	64 (69.6%)	<0.001
	Yes, anonymized review influenced my decision to include preliminary data in my application.	28 (30.4%)	
Please indicate if anonymized review influenced the amount or type of preliminary data included in your application	No, anonymized review did not influence the amount or type of preliminary data included in my application.	62 (67.4%)	0.001
	Yes, anonymized review influenced the amount or type of preliminary data included in my application.	30 (32.6%)	
Please indicate if anonymized review influenced your decision to exclude preliminary data from your application	No, anonymized review did not influence my decision to exclude preliminary data	7 (58.3%)	0.774
	Yes, anonymized review influenced my decision to exclude preliminary data	5 (41.7%)	

Table 8. Past TRA History

Variable of interest	Answer options within variable of interest	Total number (percent) of respondents for answer option	p
Please indicate the number of times that you submitted a TRA award prior to 2020	0	74 (70.5%)	-
	1	20 (19.0%)	
	2	8 (7.6%)	
	3	1 (1.0%)	
	5 or more	2 (1.9%)	
Please indicate the number of times that you won a TRA award prior to 2020	0	95 (94.1%)	-
	1	4 (4.0%)	
	2	2 (2.0%)	

Table 9. Anonymizing with Collaborators

Variable of interest	Answer options within variable of interest	Total number (percent) of respondents for answer option	p
Please indicate if your application had one or more collaborator(s)	No, my application did not have one or more collaborator(s).	19 (18.3%)	<0.001
	Yes, my application had one or more collaborator(s).	85 (81.7%)	
You indicated that you have collaborators in your application. Please select how many collaborators were included in your application	1	20 (23.5%)	-
	2	24 (28.2%)	
	3	12 (14.1%)	
	4 or more	29 (34.1%)	
Please indicate if anonymizing the identity of your collaborator(s) was challenging	No, anonymizing the identity of my collaborator(s) was not challenging.	38 (44.7%)	0.386
	Yes, anonymizing the identity of my collaborator(s) was challenging.	47 (55.3%)	

Appendix E. Pre-phase I Administrative Review Survey: Specific Aims

Thank you for taking part in this study conducted by the IDA Science and Technology Policy Institute (STPI) on behalf of the National Institutes of Health Office of the Director (NIH/OD) and the Center for Scientific Review (NIH/CSR). STPI is a federally funded research and development center that provides rigorous, independent research and analysis to the Federal Government.

Purpose of the Survey

This survey solicits your perspectives on reviewing the anonymized Specific Aims for the Transformative Research Award.

Confidentiality Statement

STPI is independent of NIH and has been contracted to collect these data. All responses will be kept confidential and protected to the extent possible by law. Only aggregate data will be presented to NIH. Your decision to participate is voluntary and will have no effect on your current or future relationship with the agency.

The estimated survey completion time is 10-15 minutes. You will be able to move forward and backward through the survey to review or edit responses. Your survey responses are automatically saved up to the last completed page, so you will be able to pause and return mid-survey. However, once you submit the survey, you will not be able to edit your responses. While completing this survey, you will be asked several questions about the anonymized TRA Specific Aims.

If you would like to review the TRA Funding Opportunity Announcements (FOAs), please click on the following: General TRA: **RFA-RM-20-013**, Notice of Special Interest (NOSI) Common Fund TRA FOA for ALS-related research: **NOT-RM-20-019**, and the Emergency TRA FOA for SARS-CoV-2-related research: **RFA-RM-20-020**.

Inquiries and Concerns

If you have questions or concerns about completing this survey, please contact us at **anonTRAreview@ida.org**.

Thank you for your participation

Your responses are invaluable to the study.

The first step in the review of TRA anonymized applications is NIH administrative review of the Specific Aims page of each application. NIH instructed applicants to remove information from their Specific Aims that would allow the reviewers to identify the applicant or collaborators, institution, or laboratory before submitting the application. The applicant's compliance with that instruction was then verified by NIH staff. The following questions address NIH staff experience in reviewing anonymized Specific Aims.

You have been identified as a participant in the NIH administrative review of Specific Aims and received instructions from TRA program managers outlining the review process.

1. **Please indicate if the instructions you received were sufficient for you to assess if the Specific Aims were anonymized.**

   ( ) Yes, the instructions I received were sufficient for me to assess if the Specific Aims were anonymized.

   ( ) No, the instructions I received were not sufficient for me to assess if the Specific Aims were anonymized.

Logic: Hidden unless: #1 Question "Please indicate if the instructions you received were sufficient for you to assess if the Specific Aims were anonymized." is one of the following answers ("No, the instructions I received were not sufficient for me to assess if the Specific Aims were anonymized.")

2. **Please describe what additional instructions would have been helpful for you to determine whether anonymity of the Specific Aims was compromised.**

   _____

   _____

   _____

   _____

Logic: Hidden unless: Invite Variable "custom1" is exactly equal to "yes"

3. **For the anonymized Specific Aims that you believe contained "other" identifying information, please indicate whether you made this determination because you possess related knowledge and expertise.**

   ( ) Yes, I made this determination because I possess related knowledge and expertise.

   ( ) No, I did not make this determination because I possess related knowledge and expertise.

Logic: Hidden unless: Invite Variable "custom2" is exactly equal to "yes"

Piping: Piped values from spreadsheet with reviewer consensus recorded.

4. **Our records show that during the administrative review, you and a fellow reviewer disagreed on the anonymity of a Specific Aims but later reached consensus.**

   Please describe the process through which you and your fellow reviewer reached consensus.

   _____

   _____

   _____

   _____

Page entry logic: This page will show when: Invite Variable "custom3" is exactly equal to "no"

Piping: Piped values from spreadsheet with reviewer consensus recorded.

5. **Our records show that during the administrative review, you and a fellow reviewer disagreed on the anonymity of a Specific Aims and did not reach consensus.**

Please describe what prevented you and the other reviewer from reaching consensus.

_____

_____

_____

_____

Validation: Must be numeric Whole numbers only Positive numbers only
Piping: Piped values from spreadsheet with number of Specific Aims recorded

6. **Our records indicate that you reviewed [invite("custom 4")] Specific Aims. Please estimate, on average, how long (in minutes) it took you to review the Specific Aims for each application.**

   _____

7. **If anonymized review is expanded at NIH, please indicate your level of interest in participating in future exercises.**

   ( ) Not interested at all ( ) Not very interested ( ) Neither ( ) Somewhat interested ( ) Very interested

8. **Please provide any additional information that you would like NIH to consider about administrative review of the anonymized Specific Aims.**

   _____

   _____

   _____

   _____

Please be advised that this is the last question in the survey. Once you submit your survey, you will not be able to go back and change your responses.

Thank you for taking our survey. Your response is important to inform future decisions regarding anonymized review.

Appendix F. Pre-phase I Administrative Review Survey Data: Specific Aims

Table 1. Information Used During Review

Variable of interest	Answer options within variable of interest	Total number (percent) of respondents for answer option	Chisq (df)	p
Please indicate if the instructions you received were sufficient for you to assess if the Specific Aims were anonymized	No, the instructions I received were not sufficient for me to assess if the Specific Aims were anonymized.	6 (14.0%)	41.86 (df =1)	<0.001
	Yes, the instructions I received were sufficient for me to assess if the Specific Aims were anonymized.	37 (86.0%)		
For the anonymized Specific Aims that you believe contained “other” identifying information, please indicate whether you made this determination because you possess related knowledge and expertise	No, I did not make this determination because I possess related knowledge and expertise.	4 (40.0%)	0.2 (df =1)	0.655
	Yes, I made this determination because I possess related knowledge and expertise.	2 (60.0%)		

Table 2. Time of Review

Variable of interest	Answer options within variable of interest	Total number (percent) of respondents for answer option	Mean	SEM
Please estimate, on average, how long (in minutes) it took you to review the Specific Aims for each application	5 or fewer minutes	7 (16.3%)	15.07	1.82
	6 to 10 minutes	16 (37.2%)		
	11 to 15 minutes	8 (18.6%)		
	16 to 30 minutes	10 (23.3%)		
	31 to 60 minutes	0 (4.7%)		

Table 3. Future Interest

Variable of interest	Answer options within variable of interest	Total number (percent) of respondents for answer option	Chisq (df)	p
If anonymized review is expanded at NIH, please indicate your level of interest in participating in future exercises	Not very or not at all interested	5 (11.6%)	39.31 (df =1)	<0.001
	Somewhat or very interested	35 (81.4%)		
	1 - Not interested at all	1 (2.3%)		
	2 - Not very interested	4 (9.3%)		
	3 - Neither	3 (7.0%)		
	4 - Somewhat interested	15 (34.9%)		
	5 - Very interested	20 (46.5%)		

Appendix G. Phase I Editorial Board Review Survey

Thank you for taking part in this study conducted by the IDA Science and Technology Policy Institute (STPI) on behalf of the National Institutes of Health Office of the Director (NIH/OD) and the Center for Scientific Review (NIH/CSR). STPI is a federally funded research and development center that provides rigorous, independent research and analysis to the Federal Government.

Purpose of the Survey

This survey solicits your perspectives on reviewing the Specific Aims in anonymized applications for the Transformative Research Award.

Confidentiality Statement

STPI is independent of NIH and has been contracted to collect these data. All responses will be kept confidential and protected to the extent possible by law. Only aggregate data will be presented to NIH. Your decision to participate is voluntary and will have no effect on your current or future relationship with the agency.

The estimated survey completion time is 10-15 minutes. You will be able to move forward and backward through the survey to review or edit responses. Your survey responses are automatically saved up to the last submitted page, so you will be able to pause and return mid-survey. However, once you submit the survey, you will not be able to edit your responses. While completing this survey, you will be asked several questions about the TRA Specific Aims you reviewed.

If you would like to review the TRA Funding Opportunity Announcements (FOAs), please click on the following: General TRA: **RFA-RM-20-013**, Notice of Special Interest (NOSI) Common Fund TRA FOA for ALS-related research: **NOT-RM-20-019**, and the Emergency TRA FOA for SARS-CoV-2-related research: **RFA-RM-20-020**.

Inquiries and Concerns

If you have questions or concerns about completing this survey, please contact us at **TRAreview2021@ida.org**.

Thank you for your participation

Your responses are invaluable to the study.

Validation: Min = 1 Max = 90 Must be numeric Whole numbers only Positive numbers only Max character count = 2 Min character count = 1

1. **Please indicate the number of anonymized Specific Aims you reviewed.***

_____

Piping: Piped answer values question one

2. **Of the [question('value'), id='4'] Specific Aims you reviewed, please indicate if there were any for which you believe you could identify the applicant or collaborator(s), lab group, or institution.***

( ) Yes, there were Specific Aims for which I believe I could identify the applicant or collaborator(s), lab group, or institution.

( ) No, there were not Specific Aims for which I believe I could identify the applicant or collaborator(s), lab group, or institution.

Logic: Hidden unless: #2 Question "Of the [question('value'), id='4'] Specific Aims you reviewed, please indicate if there were any for which you believe you could identify the applicant or collaborator(s), lab group, or institution." is one of the following answers ("Yes, there were Specific Aims for which I believe I could identify the applicant or collaborator(s), lab group, or institution.")

Piping: Piped answer values question one

Validation: Min = 1 Max = 90 Must be numeric Whole numbers only Positive numbers only Max character count = 2

3. **Of the [question('value'), id='4'] anonymized Specific Aims you reviewed, please estimate the number of anonymized Specific Aims for which you believe you could identify the applicant or collaborator(s), lab group, or institution.**

Please ensure that the total number entered below is less than or equal to [question('value'), id='4'].*

_____

Logic: Hidden unless: #2 Question "Of the [question('value'), id='4'] Specific Aims you reviewed, please indicate if there were any for which you believe you could identify the applicant or collaborator(s), lab group, or institution." is one of the following answers ("Yes, there were Specific Aims for which I believe I could identify the applicant or collaborator(s), lab group, or institution.")

Piping: Piped answer values from question three

Validation: Must be numeric Whole numbers only Positive numbers only

4. **In the previous question, you indicated there were [question('value'), id='134'] anonymized Specific Aims for which you believe you could identify the applicant or collaborator(s), lab group, or institution.**

   Of the anonymized Specific Aims for which you believe you could identify the applicant or collaborator(s), lab group, or institution, please estimate the number **of anonymized Specific Aims for which you felt:**

   Please ensure that the total number of anonymized Specific Aims below is equal to [question('value'), id='134'].

   Very confident in the identification: _____

   Somewhat confident in the identification: _____

   Not confident at all in the identification: _____

5. **Please indicate if there were any applications for which you could *not* assess transformative potential because you only had access to the Specific Aims.**

   ( ) Yes, there were applications for which I could not assess transformative potential because I only had access to the Specific Aims.

   ( ) No, I could assess transformative potential of all applications even though I only access to the Specific Aims.

Page entry logic: This page will show when: #5 Question "Please indicate if there were any applications for which you could *not* assess transformative potential because you only had access to the Specific Aims." is one of the following answers ("Yes, there were applications for which I could not assess transformative potential because I only had access to the Specific Aims.")

Validation: Must be numeric Whole numbers only Positive numbers only

6. **Please estimate how many applications for which you could not assess transformative potential because you only had access to the Specific Aims.**

_____

7. **Please indicate if there were any applications for which you could *not* assess transformative potential because they were anonymized.**

   ( ) Yes, there were applications for which I could not assess transformative potential because they were anonymized.

   ( ) No, I could assess transformative potential of all applications even though they were anonymized.

8. **Please indicate, on average, how confident you were in your ability to determine if the research in the anonymized Specific Aims you reviewed was transformative.***

   ( ) Not confident at all ( ) Not very confident ( ) Neither ( ) Somewhat confident ( ) Very confident

Logic: Show/hide trigger exists.

9. **Please indicate how difficult it was to score anonymized Specific Aims compared to the traditional NIH review process. ***

   ( ) much more difficult ( ) more difficult ( ) about the same ( ) easier ( ) much easier

Logic: Hidden unless: #9 Question "Please indicate how difficult it was to score anonymized Specific Aims compared to the traditional NIH review process. " is one of the following answers ("much more difficult","more difficult","easier","much easier")

Piping: Piped answer values from question nine

10. Please describe the differences between the anonymized Specific Aims review process and the traditional NIH review process that made the anonymized Specific Aims [question("value"), id="155"] to score.*

_____

_____

_____

_____

---

11. Are there any additional comments you would like to provide to NIH about the use of anonymized Specific Aims for this step of the review process?

_____

_____

_____

_____

Please be advised that this is the last question in the survey. Once you submit your survey, you will not be able to go back and change your responses.

Thank you for taking our survey. The aggregated results will assist NIH in continuing to offer programs that support transformative research.

Appendix H. Phase I Editorial Board Review Survey Data

Table 1. Number of Specific Aims Reviewed

Variable of interest	Mean	SEM
Please indicate the number of anonymized Specific Aims you reviewed	50.42	0.92

Table 2. Ability to Identify

Variable of interest	Answer options within variable of interest	Total number (percent) of respondents for answer option	p
Of the X Specific Aims you reviewed, please indicate if there were any for which you believe you could identify the applicant or collaborator(s), lab group, or institution	No, there were not Specific Aims for which I believe I could identify the applicant or collaborator(s), lab group, or institution.	12 (100.0%)	<0.001
	Yes, there were Specific Aims for which I believe I could identify the applicant or collaborator(s), lab group, or institution.	0 (0.0%)	

Table 3. Ability to Assess Based on Only Specific Aims

Variable of interest	Answer options within variable of interest	Total number (percent) of respondents for answer option	p
Please indicate if there were any applications for which you could not assess transformative potential because you only had access to the Specific Aims	No, I could assess transformative potential of all applications even though I only access to the Specific Aims.	7 (58.3%)	0.774
	Yes, there were applications for which I could not assess transformative potential because I only had access to the Specific Aims.	5 (41.7%)	

Table 4. Number of Specific Aims with No Ability to Assess Based on Only Specific Aims

Variable of interest	Mean	SEM
Please estimate how many applications for which you could not assess transformative potential because you only had access to the Specific Aims	13	2

Table 5. Ability to Assess Based on Anonymized Application

Variable of interest	Answer options within variable of interest	Total number (percent) of respondents for answer option	p
Please indicate if there were any applications for which you could not assess transformative potential because they were anonymized	No, I could assess transformative potential of all applications even though they were anonymized.	9 (75.0%)	0.146
	Yes, there were applications for which I could not assess transformative potential because they were anonymized.	3 (25.0%)	

Table 6. Confidence in Assessment of Transformative Potential

Variable of interest	Answer options within variable of interest	Total number (percent) of respondents for answer option
Please indicate, on average, how confident you were in your ability to determine if the research in the anonymized Specific Aims you reviewed was transformative	Not very or not at all confident	1 (8.3%)
	Somewhat or very confident	11 (91.7%)
	1 - Not confident at all	0 (0.0%)
	2 - Not very confident	1 (8.3%)
	3 - Neither	0 (0.0%)
	4 - Somewhat confident	8 (66.7%)
	5 - Very confident	3 (25.0%)

Table 7. Difficulty in Scoring

Variable of interest	Answer options within variable of interest	Total number (percent) of respondents for answer option
Please indicate how difficult it was to score anonymized Specific Aims compared to the traditional NIH review process	More or much more difficult	4 (33.3%)
	Easier or much easier	3 (25.0%)
	1 - Much more difficult	0 (0.0%)
	2 - More difficult	4 (33.3%)
	3 - About the same	5 (41.7%)
	4 - Easier	2 (16.7%)
	5 - Much easier	1 (8.3%)

Appendix I. Pre-phase II Administrative Review Survey: Research Strategy

Thank you for taking part in this study conducted by the IDA Science and Technology Policy Institute (STPI) on behalf of the National Institutes of Health Office of the Director (NIH/OD) and the Center for Scientific Review (NIH/CSR). STPI is a federally funded research and development center that provides rigorous, independent research and analysis to the Federal Government.

Purpose of the Survey

This survey solicits your perspective on reviewing the anonymized Research Strategies from the general Transformative Research Award (TRA) FOA.

Confidentiality Statement

STPI is independent of NIH and has been contracted to collect these data. All responses will be kept confidential and protected to the extent possible by law. Only aggregate data will be presented to NIH. Your decision to participate is voluntary and will have no effect on your current or future relationship with the agency.

The estimated survey completion time is 5 minutes. You will be able to move forward and backward through the survey to review or edit responses. Your survey responses are automatically saved up to the last completed page, so you will be able to pause and return mid-survey. However, once you submit the survey, you will not be able to edit your responses. While completing this survey, you will be asked several questions about the anonymized TRA Research Strategies you reviewed.

If you would like to review the TRA Funding Opportunity Announcements (FOAs), please click on the following: General TRA: **RFA-RM-20-013**, Notice of Special Interest (NOSI) Common Fund TRA FOA for ALS-related research: **NOT-RM-20-019**, and the Emergency TRA FOA for SARS-CoV-2-related research: **RFA-RM-20-020**.

Inquiries and Concerns

If you have questions or concerns about completing this survey, please contact us at **TRAreview2021@ida.org**.

Thank you for your participation

Your responses are invaluable to the study.

In the FOA soliciting anonymized TRA applications, NIH instructed applicants to remove information from their Research Strategy that would allow the reviewers to identify the applicant, lab group, institution, or collaborator(s). You were identified as a participant in the administrative review of these Research Strategies, and received instructions from the TRA program director outlining the administrative review process.

This survey solicits your perspectives on your review of anonymized TRA Research Strategies from the general FOA: RFA-RM-20-013.

1. **Please indicate if the instructions you received were sufficient for you to assess if the Research Strategies from the general TRA FOA were anonymized.**

   ( ) Yes, the instructions I received were sufficient for me to assess if the Research Strategies from the general TRA FOA were anonymized.

   ( ) No, the instructions I received were not sufficient for me to assess if the Research Strategies from the general TRA FOA were anonymized.

Logic: Hidden unless: #1 Question "Please indicate if the instructions you received were sufficient for you to assess if the Research Strategies from the general TRA FOA were anonymized." is one of the following answers ("No, the instructions I received were not sufficient for me to assess if the Research Strategies from the general TRA FOA were anonymized.")

2. **Please describe what additional instructions would have been helpful for you to determine whether anonymity of the Research Strategies were compromised.**

   _____

   _____

   _____

Piping: Piped in data from spreadsheet indicating how many Research Strategies the reviewer reviewed.

Validation: Must be numeric Whole numbers only Positive numbers only

3. **Our records indicate that you reviewed [invite('custom 1')] Research Strategies from the general TRA FOA. Please estimate, on average, how long (in minutes) it took you to review the Research Strategy for each application in the general TRA FOA.**

   _____

4. **If anonymized review is expanded at NIH, please indicate your level of interest in participating in future exercises.**

   ( ) Not interested at all ( ) Not very interested ( ) Neither ( ) Somewhat interested ( ) Very interested

5. **Please provide any additional information that you would like NIH to consider about administrative review of anonymized Research Strategies.**

Appendix J. Pre-phase II Administrative Review Survey Data: Research Strategy

Table 1. Information Used During Review

Variable of interest	Answer options within variable of interest	Total number (percent) of respondents for answer option	Chisq (df)	p
Please indicate if the instructions you received were sufficient for you to assess if the Research Strategies from the general TRA FOA were anonymized	No, the instructions I received were not sufficient for me to assess if the Research Strategies from the general TRA FOA were anonymized.	2 (8.7%)	28.17 (df =1)	<0.001
	Yes, the instructions I received were sufficient for me to assess if the Research Strategies from the general TRA FOA were anonymized.	14 91.3%)		

Table 2. Time of Review

Variable of interest	Answer options within variable of interest	Total number (percent) of respondents for answer option	Mean	SEM
Please estimate, on average, how long (in minutes) it took you to review the Research Strategy for each application in the general TRA FOA.	5 or fewer minutes	2 (8.7%)	39.26	10.12
	6 to 10 minutes	1 (4.3%)		
	11 to 15 minutes	3 (13.0%)		
	16 to 30 minutes	9 (39.1%)		
	31 to 60 minutes	6 (26.1%)		
	More than 60 minutes	2 (8.7%)		

Table 3. Future Interest

Variable of interest	Answer options within variable of interest	Total number (percent) of respondents for answer option	Chisq (df)	p
If anonymized review is expanded at NIH, please indicate your level of interest in participating in future exercises	Not very or not at all interested	4 (17.4%)	10.7 (df =1)	0.001
	Somewhat or very interested	16 (69.6%)		
	1 - Not interested at all	1 (4.3%)		
	2 - Not very interested	3 (13.0%)		
	3 - Neither	3 (13.0%)		
	4 - Somewhat interested	5 (21.7%)		
	5 - Very interested	11 (47.8%)		

Appendix K. Phase II Technical Reviewer Survey

Thank you for taking part in this study conducted by the IDA Science and Technology Policy Institute (STPI) on behalf of the National Institutes of Health Office of the Director (NIH/OD) and the Center for Scientific Review (NIH/CSR). STPI is a federally funded research and development center that provides rigorous, independent research and analysis to the Federal Government.

Purpose of the Survey

This survey solicits your perspectives on reviewing an anonymized application for the Transformative Research Award.

Confidentiality Statement

STPI is independent of NIH and has been contracted to collect these data. All responses will be kept confidential and protected to the extent possible by law. Only aggregate data will be presented to NIH. Your decision to participate is voluntary and will have no effect on your current or future relationship with the agency.

The estimated survey completion time is approximately 10 minutes. You will be able to move forward and backward through the survey to review or edit responses. Your survey responses are automatically saved up to the last completed page, so you will be able to pause and return mid-survey. However, once you submit the survey, you will not be able to edit your responses. While completing this survey, you will be asked several questions about the anonymized TRA Specific Aims and Research Strategy.

If you would like to review the TRA Funding Opportunity Announcements (FOAs), please click on the following: General TRA: RFA-RM-20-013, Notice of Special Interest (NOSI) Common Fund TRA FOA for ALS-related research: NOT-RM-20-019, and the Emergency TRA FOA for SARS-CoV-2-related research: RFA-RM-20-020.

Inquiries and Concerns

If you have questions or concerns about completing this survey, please contact us at TRAreview2021@ida.org.

Thank you for your participation

Your responses are invaluable to the study.

Validation: %s format expected Using custom RegEx pattern

ID: 149

To provide STPI with the appropriate context for your responses, please provide the following information. As a reminder, only aggregate survey results will be provided to NIH.

1) Please enter the the unique 6 digits located in your application's grant number (1R01OD[XXXXXX]-01).
Please note, all subsequent questions will pertain to this application only.*

_____

ID: 150

2) Your full first and last name.*

First Name: _____

Last Name: _____

Please answer the question below for the TRA application: **1R01OD[[question('value'), id='149']]-01**

ID: 108

3) Please indicate if you could identify the applicant, lab group, institution, or the collaborator.

( ) Yes, I could identify the applicant, lab group, institution, or the collaborator.

( ) No, I could not identify the applicant, lab group, institution, or the collaborator.

Page entry logic: This page will show when: #3 Question "Please indicate if you could identify the applicant, lab group, institution, or the collaborator." is one of the following answers ("Yes, I could identify the applicant, lab group, institution, or the collaborator.")

Please answer the question below for the TRA application: **1R01OD[[question('value'), id='149']]-01**

ID: 121

4) Please describe what type of information in the anonymized Specific Aims or Research Strategy suggested the identity of the applicant, lab group, institution, or the collaborator.

_____

_____

_____

_____

Page entry logic: This page will show when: #3 Question "Please indicate if you could identify the applicant, lab group, institution, or the collaborator." is one of the following answers ("Yes, I could identify the applicant, lab group, institution, or the collaborator.")

Please answer the question below for the TRA application: **1R01OD[[question('value'), id='149']]-01**

ID: 146

5) Please indicate how confident you were in your ability to identify the applicant, lab group, institution, or the collaborator.

( ) Not confident at all ( ) Not very confident ( ) Somewhat confident ( ) Very confident

ID: 113

Please answer the question below for the TRA application: 1R01OD[[question('value'), id='149']]-01

NIH describes **Significance** using the following components:

- Does the study have transformative potential; is the study exceptional in terms of its consequences for the field and size of the community affected; or does the study challenge or propose a paradigm fundamental to the field?

6) Please indicate if you felt you had sufficient information in the anonymized Specific Aims and Research Strategy to evaluate if the proposed research was *significant*.

( ) Yes, I felt that I had sufficient information in the anonymized Specific Aims and Research Strategy to evaluate if the proposed research was significant.

( ) No, I did not feel that I had sufficient information in the anonymized Specific Aims and Research Strategy to evaluate if the proposed research was significant.

ID: 114

Please answer the question below for the TRA application: 1R01OD[[question('value'), id='149']]-01

NIH describes **Innovation** with the components below:

- Is the study exceptionally innovative, and/or considered unconventional research with the potential to create new scientific paradigms, establish entirely new and improved clinical approaches, or develop transformative technologies?

7) Please indicate if you felt you had sufficient information in the anonymized Specific Aims and Research Strategy to evaluate if the proposed research was *innovative*.

( ) Yes, I felt that I had sufficient information in the anonymized Specific Aims and Research Strategy to evaluate if the proposed research was innovative.

( ) No, I did not feel that I had sufficient information in the anonymized Specific Aims and Research Strategy to evaluate if the proposed research was innovative.

ID: 127

Please answer the question below for the TRA application: 1R01OD[[question('value'), id='149']]-01

NIH describes **Approach** using the components below:

Logical and compelling
Is the logic of the approach compelling despite the lack of experimental details or substantial preliminary data?
Feasible
Is there evidence that the PI(s) will pursue the risky research topic in a robust, reproducible and rigorous manner, and does the project timeline seem realistic?

8) Please indicate if you felt that you had sufficient information in the Specific Aims and Research Strategy to evaluate if the proposed research was *logical and compelling*.

( ) Yes, I felt that I had sufficient information in the Specific Aims and Research Strategy to evaluate if the proposed research was logical and compelling.

( ) No, I did not feel that I had sufficient information in the Specific Aims and Research Strategy to evaluate if the proposed research was logical and compelling.

ID: 147

9) Please indicate if you felt that you had sufficient information in the Specific Aims and Research Strategy to evaluate if the proposed research was *feasible*.

( ) Yes, I felt that I had sufficient information in the Specific Aims and Research Strategy to evaluate if the proposed research was feasible.

( ) No, I did not feel that I had sufficient information in the Specific Aims and Research Strategy to evaluate if the proposed research was feasible.

(untitled)

Logic: Hidden unless: (((#6 Question "Please indicate if you felt you had sufficient information in the anonymized Specific Aims and Research Strategy to evaluate if

the proposed research was *significant*." is one of the following answers ("No, I did not feel that I had sufficient information in the anonymized Specific Aims and Research Strategy to evaluate if the proposed research was significant.") OR #7 Question "Please indicate if you felt you had sufficient information in the anonymized Specific Aims and Research Strategy to evaluate if the proposed research was *innovative*." is one of the following answers ("No, I did not feel that I had sufficient information in the anonymized Specific Aims and Research Strategy to evaluate if the proposed research was innovative.")) OR #8 Question "Please indicate if you felt that you had sufficient information in the Specific Aims and Research Strategy to evaluate if the proposed research was *logical and compelling*. " is one of the following answers ("No, I did not feel that I had sufficient information in the Specific Aims and Research Strategy to evaluate if the proposed research was logical and compelling.")) OR #9 Question "Please indicate if you felt that you had sufficient information in the Specific Aims and Research Strategy to evaluate if the proposed research was *feasible*." is one of the following answers ("No, I did not feel that I had sufficient information in the Specific Aims and Research Strategy to evaluate if the proposed research was feasible."))

ID: 151

Please answer the question below for the TRA application: 1R01OD[[question('value'), id='149']]-01

10) You indicated that you did not have sufficient information in the Specific Aims and Research Strategy to evaluate one or more aspects of the review criteria (Significance, Innovation, or Approach). Please describe what additional information would have been helpful in the review.

_____

_____

_____

_____

(untitled)

ID: 134

11) Are there any additional comments you would like to provide to NIH about the mail review step in the anonymized TRA review process?

_____

_____

_____

_____

ID: 137

Please be advised that this is the last question in the survey. Once you submit your survey, you will not be able to go back and change your responses.

Thank You!

ID: 1

Thank you for taking our survey. Your response is important to inform future decisions regarding anonymized review.

Appendix L. Phase II Technical Review Survey Data

Table 1. Ability to Identify

Variable of interest	Answer options within variable of interest	Total number (percent) of respondents for answer option	p
Please indicate if you could identify the applicant, lab group, institution, or the collaborator	No, I could not identify the applicant, lab group, institution, or the collaborator.	84 (80.8%)	<0.001
	Yes, I could identify the applicant, lab group, institution, or the collaborator.	20 (19.2%)	

Table 2. Confidence in Identification

Variable of interest	Answer options within variable of interest	Total number (percent) of respondents for answer option
Please indicate how confident you were in your ability to identify the applicant, lab group, institution, or the collaborator.	Not very or not at all confident	0 (0.0%)
	Somewhat or very confident	20 (100.0%)
	1 - Not confident at all	0 (0.0%)
	2 - Not very confident	0 (0.0%)
	3 - Neither	0 (0.0%)
	4 - Somewhat confident	8 (40.0%)
	5 - Very confident	12 (60.0%)

Table 3. Information Sufficiency

Variable of interest	Answer options within variable of interest	Total number (percent) of respondents for answer option	p
Please indicate if you felt you had sufficient information in the anonymized Specific Aims and Research Strategy to evaluate if the proposed research was significant	No, I did not feel that I had sufficient information in the anonymized Specific Aims and Research Strategy to evaluate if the proposed research was significant.	12 (11.5%)	<0.001
	Yes, I felt that I had sufficient information in the anonymized Specific Aims and Research Strategy to evaluate if the proposed research was significant.	92 (88.5%)	
Please indicate if you felt you had sufficient information in the anonymized Specific Aims and Research Strategy to evaluate if the proposed research was innovative	No, I did not feel that I had sufficient information in the anonymized Specific Aims and Research Strategy to evaluate if the proposed research was innovative.	14 (13.6%)	<0.001
	Yes, I felt that I had sufficient information in the anonymized Specific Aims and Research Strategy to evaluate if the proposed research was innovative.	89 (86.4%)	

Please indicate if you felt that you had sufficient information in the Specific Aims and Research Strategy to evaluate if the proposed research was logical and compelling	No, I did not feel that I had sufficient information in the Specific Aims and Research Strategy to evaluate if the proposed research was logical and compelling.	19 (18.3%)	<0.001
	Yes, I felt that I had sufficient information in the Specific Aims and Research Strategy to evaluate if the proposed research was logical and compelling.	85 (81.7%)	
Please indicate if you felt that you had sufficient information in the Specific Aims and Research Strategy to evaluate if the proposed research was feasible	No, I did not feel that I had sufficient information in the Specific Aims and Research Strategy to evaluate if the proposed research was feasible.	40 (38.5%)	0.024
	Yes, I felt that I had sufficient information in the Specific Aims and Research Strategy to evaluate if the proposed research was feasible.	64 (61.5%)	

Appendix M. Phase IIIa and IIIb Editorial Board Survey

Thank you for taking part in this study conducted by the IDA Science and Technology Policy Institute (STPI) on behalf of the National Institutes of Health Office of the Director (NIH/OD) and the Center for Scientific Review (NIH/CSR). STPI is a federally funded research and development center that provides rigorous, independent research and analysis to the Federal Government.

Purpose of the Survey

This survey solicits your perspective on reviewing the Specific Aims, Research Strategies, and mail review notes for anonymized applications in phase IIIa, and the full applications in phase IIIb for the Transformative Research Award (TRA).

Confidentiality Statement

STPI is independent of NIH and has been contracted to collect these data. All responses will be kept confidential and protected to the extent possible by law. Only aggregate data will be presented to NIH. Your decision to participate is voluntary and will have no effect on your current or future relationship with the agency.

The estimated survey completion time is 10-15 minutes. You will be able to move forward and backward through the survey to review or edit responses. Your survey responses are automatically saved up to the last submitted page, so you will be able to pause and return mid-survey. However, once you submit the survey, you will not be able to edit your responses. While completing this survey, you will be asked several questions about the TRA application materials you reviewed.

If you would like to review the TRA Funding Opportunity Announcements (FOAs), please click on the following: General TRA: **RFA-RM-20-013**, Notice of Special Interest (NOSI) Common Fund TRA FOA for ALS-related research: **NOT-RM-20-019**, and the Emergency TRA FOA for SARS-CoV-2-related research: **RFA-RM-20-020**.

Inquiries and Concerns

If you have questions or concerns about completing this survey, please contact us at **TRAreview2021@ida.org**.

Thank you for your participation
Your responses are invaluable to the study.

---

phase III of the TRA anonymized review process has two components:

- **phase IIIa**. review anonymized Specific Aims, Research Strategy, and mail review critiques for each application and provide a preliminary score
- **phase IIIb**. discuss and provide a final score for full, deanonymized applications at the meeting

The following questions pertain to your review experience in phase IIIa, during which the anonymized Specific Aims, Research Strategy, and mail review critiques were provided.

1. **Among the anonymized applications you reviewed in** phase IIIa**, please indicate if there were any applications for which you believe you could identify the applicant or collaborator(s), lab group, or institution.**

   ( ) Yes, there were applications for which I believe I could identify the applicant or collaborator(s), lab group, or institution.

   ( ) No, there were no applications for which I believe I could identify the applicant or collaborator(s), lab group, or institution.

Page entry logic: This page will show when: #1 Question "Among the anonymized applications you reviewed in **phase IIIa**, please indicate if there were any applications for which you believe you could identify the applicant or collaborator(s), lab group, or institution." is one of the following answers ("Yes, there were applications for which I believe I could identify the applicant or collaborator(s), lab group, or institution.")

Validation: Min = 1 Max = 90 Must be numeric Whole numbers only Positive numbers only Max character count = 2

2. **Please estimate the number of anonymized applications for which you believe you could identify the applicant or collaborator(s), lab group, or institution.**

   _____

Logic: This page will show when: #1 Question "Among the anonymized applications you reviewed in **phase IIIa**, please indicate if there were any applications for which you believe you could identify the applicant or collaborator(s), lab group, or institution." is

one of the following answers ("Yes, there were applications for which I believe I could identify the applicant or collaborator(s), lab group, or institution.")

Validation: Must be numeric Whole numbers only Positive numbers only

3. **In the previous question, you indicated there were [question('value'), id='134'] anonymized applications in phase IIIa for which you believe you could identify the applicant or collaborator(s), lab group, or institution.**

   For the applications for which you believed you could identify the applicant or collaborator(s), lab group, or institution, please estimate the number of anonymized applications for which you felt:

   ***Please ensure that the total number of applications below is equal to [question('value'), id='134'].**

   Very confident in the identification: _____

   Somewhat confident in the identification: _____

   Not confident at all in the identification: _____

---

4. **As part of your review in phase IIIa, you were asked to evaluate the transformative potential of the proposed research in the applications assigned to you.**

   Please indicate, on average, how confident you were in your ability to determine if the anonymized applications you reviewed in phase IIIa were transformative.

   ( ) Not confident at all ( ) Not very confident ( ) Neither ( ) Somewhat confident ( ) Very confident

Logic: Hidden unless: #4 Question "As part of your review in phase IIIa, you were asked to evaluate the **transformative potential** of the proposed research in the applications assigned to you.

Please indicate, on average, how confident you were in your ability to determine if the anonymized applications you reviewed in phase IIIa were transformative.
" is one of the following answers ("Not confident at all","Not very confident")

5. **You indicated that you were [question("value"), id="163",case="lower"] in your ability to evaluate the transformative potential of the proposed research. Please indicate if this is due to a lack of related knowledge and expertise.**

( ) Yes, this is due to a lack of related knowledge and expertise.

( ) No, this is not due to a lack of related knowledge and expertise.

Logic: Show/hide trigger exists.

6. **Please indicate how difficult or easy it was to evaluate anonymized applications in phase IIIa compared to the traditional NIH review process in which you receive the full, deanonymized application.**

( ) Much more difficult ( ) More difficult ( ) About the same ( ) Easier ( ) Much easier

---

Logic: Hidden unless: #6 Question "Please indicate how difficult or easy it was to evaluate anonymized applications in phase IIIa compared to the traditional NIH review process in which you receive the full, deanonymized application." is one of the following answers ("Much more difficult","More difficult","Easier","Much easier")

7. **Please describe the factors of the anonymized phase IIIa applications that made them [question("value"),id="155",case="lower"] to score compared to the full, deanonymized applications reviewed in the traditional NIH review.**

_____

_____

_____

8. **Please indicate if you reviewed applications in phase IIIa that were discussed in phase IIIb.**

( ) Yes, I reviewed applications in phase IIIa that were discussed in phase IIIb.

( ) No, none of my assigned applications in phase IIIa were discussed in phase IIIb.

Logic: Hidden unless: #8 Question "Please indicate if you reviewed applications in phase IIIa that were discussed in phase IIIb." is one of the following answers ("Yes, I reviewed applications in phase IIIa that were discussed in phase IIIb.")

The following question pertains to the differences in your review experience in phase IIIa and phase IIIb:

- **phase IIIa**. review anonymized Specific Aims, Research Strategy, and mail review critiques for each application and provide a preliminary score
- **phase IIIb**. discussion and provide a final score for full, deanonymized applications at the meeting

9. **For those applications you reviewed in phase IIIa that were then discussed in phase IIIb, please indicate if access to the full, deanonymized application allowed you to better assess transformative potential.**

   ( ) Yes, access to the full, deanonymized application allowed me to better assess transformative potential.

   ( ) No, access to the full, deanonymized application did not allow me to better assess transformative potential.

Logic: Hidden unless: #9 Question "For those applications you reviewed in phase IIIa that were then discussed in phase IIIb, please indicate if access to the full, deanonymized application allowed you to better assess transformative potential." is one of the following answers ("Yes, access to the full, deanonymized application allowed me to better assess transformative potential.")

10. **Please describe how access to the full, deanonymized application in phase IIIb allowed you to better assess transformative potential.**

    _____

    _____

    _____

Page entry logic: This page will show when: #8 Question "Please indicate if you reviewed applications in phase IIIa that were discussed in phase IIIb." is one of the following answers ("Yes, I reviewed applications in phase IIIa that were discussed in phase IIIb.")

The following questions pertain to your review experience in phase IIIb, where the full, deanonymized application was provided to the reviewer and discussed at the meeting.

11. **For those applications you reviewed in phase IIIa that were then discussed in phase IIIb, please indicate if you changed your final score on one or more of the applications following the discussion.**

    ( ) Yes, I changed my score for one or more of the applications I reviewed in phase IIIa that were then discussed in phase IIIb.

    ( ) No, I did not change my score for any of the applications I reviewed in phase IIIa that were then discussed in phase IIIb.

Page entry logic: This page will show when: #11 Question "For those applications you reviewed in phase IIIa that were then discussed in phase IIIb, please indicate if you changed your final score on one or more of the applications following the discussion." is one of the following answers ("Yes, I changed my score for one or more of the applications I reviewed in phase IIIa that were then discussed in phase IIIb.")

Logic: Show/hide trigger exists.

You indicated that you changed your score on an application for one or more of the applications you reviewed in phase IIIa that were then discussed in phase IIIb.

12. **Please indicate which of the following influenced you to change your score on one or more applications from phase IIIa to phase IIIb (select all that apply).**

    [ ] Editorial Board member discussion in phase IIIb

    [ ] Additional components added in phase IIIb

    [ ] Other (please specify): _____*

Logic: Hidden unless: #12 Question "Please indicate which of the following influenced you to change your score on one or more applications from phase IIIa to phase IIIb (select all that apply)." is one of the following answers ("Additional components added in phase IIIb")

13. **Please indicate which *components* of the full, deanonymized application you reviewed in phase IIIa influenced you to change your score on one or more applications from phase IIIa to phase IIIb (select all that apply).**

[ ] Biosketches

[ ] Bibliography & References Cited

[ ] Care and Use of Vertebrate Animals in Research

[ ] Consortium/Contractual Arrangements

[ ] Environment

[ ] Evidence of Independence and Institutional Support

[ ] Inclusion of Women, Minorities and Children in Research

[ ] Investigators

[ ] Leadership Plan (Multiple Program Directors/Principle Investigators)

[ ] Letters from Consultants and Collaborators (Available Expertise)

[ ] Protection of Human Subjects from Research Risk

[ ] Other (please specify):_____

14. Are there any additional comments you would like to provide to NIH about this final step of the review process?

_____

_____

_____

Please be advised that this is the last question in the survey. Once you submit your survey, you will not be able to go back and change your responses.

Thank you for taking our survey. The aggregated results will assist NIH in continuing to offer programs that support transformative research.

Appendix N. Phase III Editorial Board Survey Data

Table 1. Confidence of Identification

Variable of interest	Answer options within variable of interest	Total number (percent) of respondents for answer option	p
Among the anonymized applications you reviewed in **phase IIIa**, please indicate if there were any applications for which you believe you could identify the applicant or collaborator(s), lab group, or institution.	No, there were no applications for which I believe I could identify the applicant or collaborator(s), lab group, or institution.	19 (100.0%)	<0.001

Table 2. Confidence of Assessment on Transformative Potential

Variable of interest	Answer options within variable of interest	Total number (percent) of respondents for answer option
As part of your review in phase IIIa, you were asked to evaluate the transformative potential of the proposed research in the applications assigned to you. Please indicate, on average, how confident you were in your ability to determine if the anonymized applications you reviewed in phase IIIa were transformative.	Not very or not at all confident	2 (10.0%)
	Somewhat or very confident	18 (90.0%)
	1 - Not confident at all	0 (0.0%)
	2 - Not very confident	2 (10.0%)
	3 - Neither	0 (0.0%)
	4 - Somewhat confident	12 (60.0%)
	5 - Very confident	6 (30.0%)

Table 3. Evaluation of Transformative Potential

Variable of interest	Answer options within variable of interest	Total number (percent) of respondents for answer option	p
You indicated that you were [insert response from previous question] in your ability to evaluate the transformative potential of the proposed research. Please indicate if this is due to a lack of related knowledge and expertise.	No, this is not due to a lack of related knowledge and expertise.	1 (50.0%)	1
	Yes, this is due to a lack of related knowledge and expertise.	1 (50.0%)	

Table 4. Ease of Evaluating Anonymized Applications

Variable of interest	Answer options within variable of interest	Total number (percent) of respondents for answer option
Please indicate how difficult or easy it was to evaluate anonymized applications in phase IIIa compared to the traditional NIH review process in which you receive the full, deanonymized application.	More or much more difficult	8 (40.0%)
	Easier or much easier	5 (25.0%)
	1 - Much more difficult	1 (5.0%)
	2 - More difficult	7 (35.0%)
	3 - About the same	7 (35.0%)
	4 - Easier	4 (20.0%)
	5 - Much easier	1 (5.0%)

Table 5. Review of Discussed Applications

Variable of interest	Answer options within variable of interest	Total number (percent) of respondents for answer option	p
Please indicate if you reviewed applications in phase IIIa that were discussed in phase IIIb	Yes, I reviewed applications in phase IIIa that were discussed in phase IIIb.	20 (100.0%)	<0.001

Table 6. Impact of Full Application on Assessment of Transformative Potential

Variable of interest	Answer options within variable of interest	Total number (percent) of respondents for answer option	p
For those applications you reviewed in phase IIIa that were then discussed in phase IIIb, please indicate if access to the full, deanonymized application allowed you to better assess transformative potential	No, access to the full, deanonymized application did not allow me to better assess transformative potential.	6 (30.0%)	0.115
	Yes, access to the full, deanonymized application allowed me to better assess transformative potential.	70.0%)	

Table 7. Changing of Final Score Following Discussion

Variable of interest	Answer options within variable of interest	Total number (percent) of respondents for answer option	p
For those applications you reviewed in phase IIIa that were then discussed in phase IIIb, please indicate if you changed your final score on one or more of the applications following the discussion	No, I did not change my score for any of the applications I reviewed in phase IIIa that were then discussed in phase IIIb.	1 (5.3%)	<0.001
	Yes, I changed my score for one or more of the applications I reviewed in phase IIIa that were then discussed in phase IIIb.	18 (94.7%)	

Table 8. Information Sources That Influenced Score Changes

Please indicate which of the following influenced you to change your score on one or more applications from phase IIIa to phase IIIb (select all that apply).

Item of interest	Total number (percent) of respondents for answer option	Cochran's Q (df)	p
Additional components added in phase IIIb	8 (42.1%)	10.71 (df = 2)	0.005
Editorial Board member discussion in phase IIIb	14 (73.7%)		
Other (please specify)	3 (15.8%)		

Table 9. Information Type That Influenced Score Changes

Please indicate which components of the full, deanonymized application you reviewed in phase IIIa influenced you to change your score on one or more applications from phase IIIa to phase IIIb (select all that apply).

Item of interest	Total number (percent) of respondents for answer option	Cochran's Q (df)	p
Bibliography & References Cited	6 (30.0%)	45.05 (df = 11)	0
Biosketches	8 (40.0%)		
Consortium/Contractual Arrangements	1 (5.0%)		
Environment	4 (20.0%)		
Evidence of Independence and Institutional Support	3 (15.0%)		
Investigators	7 (35.0%)		
Leadership Plan (Multiple Program Directors/Principle Investigators)	4 (20.0%)		
Letters from Consultants and Collaborators (Available Expertise)	3 (15.0%)		

Appendix O. TRA Applicants and Awardees by State

As part of the demographic analysis of TRA applicants and awardees, STPI assessed the number of a state had a TRA applicant or awardee.

State	Number of years with at least one TRA applicant	Number of years with at least one TRA awardee
California	11	11
Massachusetts	11	11
New York	11	9
Illinois	11	4
Texas	11	3
Pennsylvania	11	3
Maryland	11	3
Georgia	11	3
Florida	11	2
Connecticut	11	2
Washington	11	2
Michigan	11	1
New Jersey	11	1
Ohio	11	1
Virginia	11	NA
Missouri	10	3

State	Number of years with at least one TRA applicant	Number of years with at least one TRA awardee
North Carolina	10	2
Alabama	10	2
Utah	10	2
Arizona	10	1
Minnesota	10	1
Colorado	10	1
Indiana	10	NA
Iowa	10	NA
Tennessee	10	NA
District of Columbia	10	NA
Wisconsin	9	2
Kentucky	9	NA
Louisiana	8	2
New Mexico	8	NA
Rhode Island	7	NA
Nebraska	7	NA
South Carolina	6	NA
Delaware	5	1
New Hampshire	5	1
Oregon	5	1
Kansas	5	NA
Maine	5	NA
Vermont	5	NA

State	Number of years with at least one TRA applicant	Number of years with at least one TRA awardee
West Virginia	5	NA
Hawaii	4	NA
Oklahoma	4	NA
Nevada	3	NA
Mississippi	3	NA
Montana	2	NA
Idaho	2	NA
South Dakota	2	NA
Arkansas	2	NA
Wyoming	1	NA
Alaska	1	NA

The evaluation presented here considers both goals: was anonymization successful and was the anonymized review rigorous and competitive.

NIH publishes FOAs to solicit applications from as many scientists as possible to maximize competitiveness and inclusion. From this perspective, one could consider that the 176 applications submitted to the FY2021 TRA FOA was decreased by 16 applications solely due to noncompliance with anonymization instructions, independent of the proposed research.

Table 6. Changes in Diversity by Demographic Factor for FY2021 Relative to the Baseline Population of FY2010–2020 for applicants and Funded Applicants

Variable of Interest	Demographic Factor	FY2021 Percentage Relative to FY2010–2020
Percentage of applicants	**Gender**	
	Female	Increased
	Male	Decreased
	Other	No change
	Race	
	Asian	Increased
	Black or African American	Increased
	White	Decreased
	Other	No change
	Ethnicity	
	Hispanic or Latino	Increased
	Not Hispanic or Not Latino	No change
	Other	No change
Percentage of funded applicants	**Gender**	
	Female	No change
	Male	No change
	Other	No change
	Race	
	Asian	No change
	Black or African American	Increased
	White	No change
	Other	No change
	Ethnicity	
	Hispanic or Latino	No change
	Not Hispanic or Not Latino	No change
	Other	No change

6. References

Mervis, Jeffrey. 2019. "To help the 'disadvantaged,' NIH refines its definition." *Science* 366 (6472): 1430-1430. https://doi.org/doi:10.1126/science.366.6472.1430. https://www.science.org/doi/abs/10.1126/science.366.6472.1430.

NIH, November 22, 2019, 2019, "Notice of NIH's Interest in Diversity," https://grants.nih.gov/grants/guide/notice-files/NOT-OD-20-031.html.

---. 2020a. "Budget." https://www.nih.gov/about-nih/what-we-do/budget#note.

---, 2020b, "NIH Director's Transformative Research Awards FOA," https://grants.nih.gov/grants/guide/rfa-files/RFA-RM-20-013.html.

---. 2021a. "Diversity Matters." https://extramural-diversity.nih.gov/diversity-matters. ---.

2021b. "NIH Director's Transformative Research Awards." https://commonfund.nih.gov/tra.

NIH Advisory Committee to the Director. 2018. "ACD High-Risk, High Reward: Working Group Update." ACD Meeting, June 14, 2018.

NIH Center for Scientific Review. 2018. "Power of NIH Peer Review." https://public.csr.nih.gov/AboutCSR/Mission/PowerofNIHPeerReview.

---. 2021. "CSR Mission." Last Modified 02/23/2021. https://public.csr.nih.gov/AboutCSR/Mission.

Reardon, Sara. 2014. "NIH to probe racial disparity in grant awards." *Nature* 512 (7514): 243-243. https://doi.org/10.1038/512243a. https://doi.org/10.1038/512243a.

Valantine, Hannah A., P. Kay Lund, and Alison E. Gammie. 2016. "From the NIH: A Systems Approach to Increasing the Diversity of the Biomedical Research Workforce." *CBE—Life Sciences Education* 15 (3): fe4. https://doi.org/10.1187/cbe.16-03-0138. https://www.lifescied.org/doi/abs/10.1187/cbe.16-03-0138.

Valantine, Hannah, Elba Serrano, and The Advisory Committee to the Director Working Group on Diversity. 2017. *National Institutes of Health Report on the Progress of Activities.*https://acd.od.nih.gov/documents/presentations/06082017Valantine-Progress.pdf.

Wadman, Meredith. 2012. "A workforce out of balance." *Nature* 486 (7403): 304-304. https://doi.org/10.1038/486304a. https://doi.org/10.1038/486304a.