Westat

Improving Lives Through Research®

Exploring the Efficacy of Live Survey Methods at the National Science Foundation

James McCall, Westat
Kelsey Gray, Westat
Matthew Ring, Westat
Breanna Wakar, Westat
Richard Griffiths, Westat
Rahul Shrivastava, Westat
Yiting Long, Westat
Robin Ferg, Westat
Karmen Perry, Westat
M. Taylor Rhodes, National Science Foundation

WESTAT @ FCSM 2024

The views presented are those of the author(s) and do not represent the views of any government agency/department or Westat.

Disclaimer

This research and the views expressed are those of the authors. They should not be attributed to the U.S. National Science Foundation (NSF), nor does the mention of trade names, commercial products, or organizations imply endorsement of the same by the U.S. government.

Agenda

- Introduction
- Survey Methods
- Survey Results
- Operational Challenges
- Exploratory Analysis of Recency Effects
- Future Research

Introduction

Introduction

Opportunity for Evaluation and Assessment Capability (EAC): Explore the logistics of implementing live survey methods

Prior surveys of the principal investigator (PI) and reviewer populations occurred much less frequently, resulting in delays between when these populations engage with the merit review process (improving recall) and when data-driven decisions can be made.

- Example: The Merit Review Survey–Biannual census survey

Live survey methods (sometimes referred to as rapid surveys or real-time data collection) aim to close the gap in temporal proximity between when PIs and reviewers submitted or reviewed proposals and when they received their first survey request.

Introduction

Research Questions

- Can we reliably collect NSF sample data on a rolling basis?
- Can we establish a workflow to field samples on a rolling basis?
- Do these methods produce higher response rates compared to other NSF surveys?

Survey Methods

- Populations
- Survey administration
- Closing the gap with sample waves

Survey Methods

Populations

PIs: Researchers submitting research proposals to a set of programs in one NSF directorate during fiscal year 2021 (n = 865)

Reviewers: Panelists and ad hoc reviewers of these research proposals (n = 907)

- 481 panelists sat on a panel to review a collection of proposals
- 426 ad hoc reviewers who reviewed a single proposal not in a panel setting

Survey Methods

Survey Administration

- Qualtrics: customized with preloaded sample data, applied display and branching logic, and automated email schedules

- Recruitment
  - Invitation
  - First reminder
  - Second reminder
  - Last chance reminder

NSF

Prenotice

Reminder

Fielding Timelines

PIs

9/27/21–9/25/22

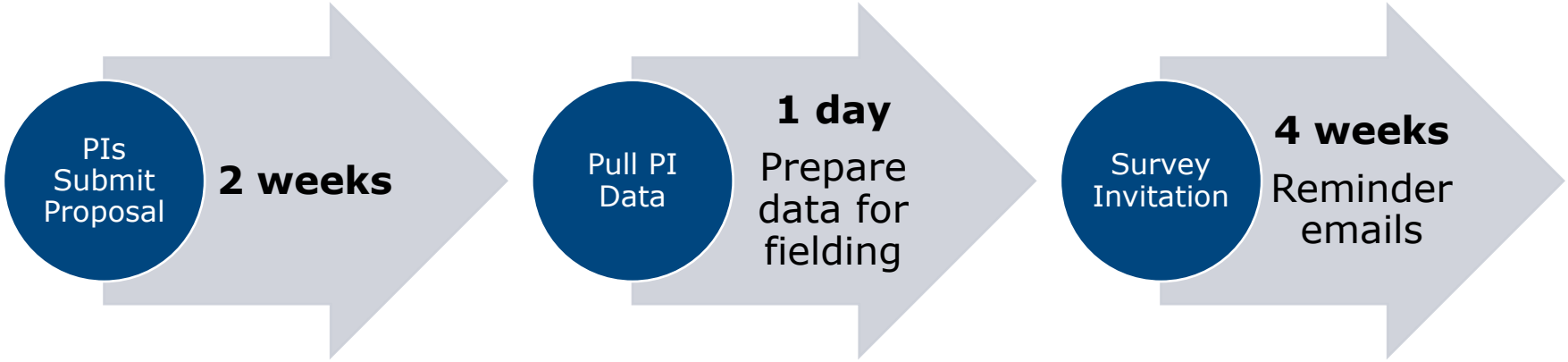
Reviewers

11/1/21–9/25/22

Survey Methods

Closing the Gap With Sample Waves

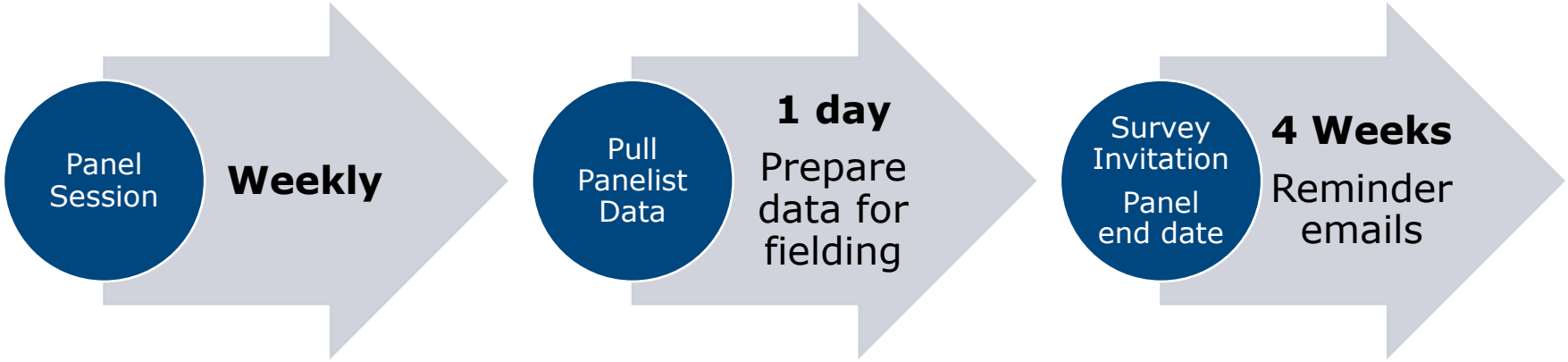
PI Sample Waves (*n* = 26)



Survey Methods

Closing the Gap with Sample Waves

Panelist Sample Waves (*n* = 34)



Survey Results

- Response rates
- Response rates by recruitment email
- Weighting and nonresponse bias analysis

Survey Results

Response Rates

PI: 370 of 865 eligible PIs completed for a response rate of **43.2%**

Panelist: 270 of 481 eligible panelists completed for a response rate of **56.1%**

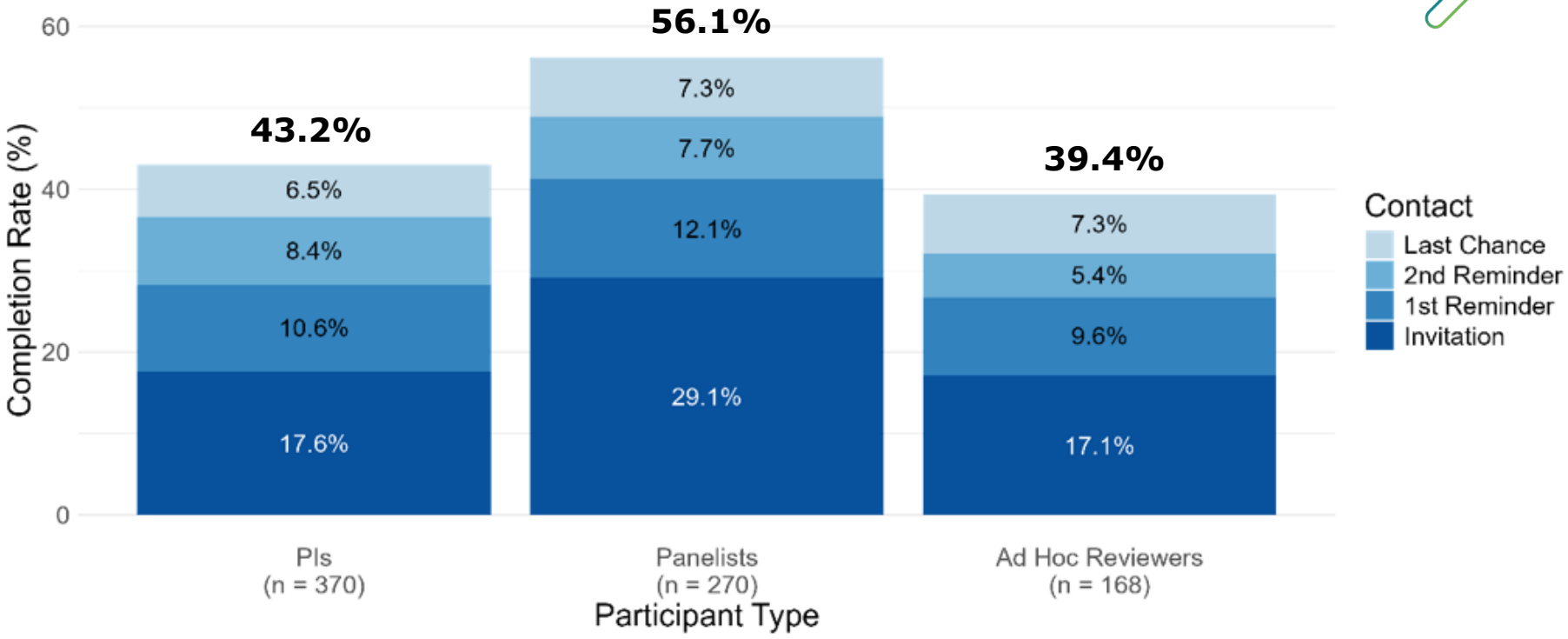
Ad hoc reviewer: 168 of 426 eligible ad hoc reviewers completed for a response rate of **39.4%**

2021 Merit Review Survey

Response rate = 27%

Survey Results

Response Rates by Recruitment Email



Survey Results

Weighting and Nonresponse Bias Analysis (NRBA)

PI population characteristics

- Solicitation number
- Deadline vs. no deadline
- Institution
- Race/ethnicity
- Gender
- Disability status
- Number of proposals since 2010
- Early career status
- Career stage
- PI funding rate since 2010

Reviewer population characteristics

- Ad hoc
- Deadline vs. no deadline
- Institution
- Race/ethnicity
- Gender
- Disability status
- Number of proposals reviewed since 2010

Survey Results

Weighting and NRBA–Propensity to Respond

PIs

- Institution–R1/non-MSI were less likely to respond than non-R1/non-MSI
- Early career PIs were more likely to respond
- Females were less likely to respond

Reviewers

- Ad hoc reviewers were less likely to respond
- Institution–non-R1/non-MSI were more likely to respond

Comparison between weighted and unweighted estimates of survey responses found little evidence of systematic differences. As such, there were no strong indications of nonresponse bias in the data.

Operational Challenges

- Coordinating task labor for rolling data collection
- Establishing eligibility criteria
- Overlapping sample waves

Operational Challenges

Coordinating Task Labor for Rolling Data Collection

- Survey methodologist: Set task schedules, aided data pulls, set wave fielding schedules, and led data management and analysis
- Two data scientists: Pull biweekly and weekly sample waves
- Two research assistants: Sample wave preparation and scheduling and sending emails
- Four statisticians: Weighting and nonresponse bias analysis

Solution: Set days for key fielding tasks–sample pulls, prep, scheduling and sending emails, and quality assurance. New sample usually went out on the same day.

Operational Challenges

Establishing Eligibility Criteria

- **PIs**: Identify unique PI-proposal pairings
- **Reviewers**: Prioritize panel attendance
  - Panelists: Identify unique PI-panel pairings
  - Ad hoc reviewers: Deduplicated within sample waves but not across sample waves
  - Reviewers were not eligible for inclusion in the sample wave if they already had one survey in the field

Solution: Wrote programs to systematically apply the eligibility criteria for all sample PI and reviewer sample waves.

Operational Challenges

Overlapping Sample Waves

- Between two and four sample waves were in the field at any given time
- Each wave was a separate Qualtrics distribution, making it easier to pause fielding for a wave if issues were discovered
- 76 total PI and reviewer sample waves

Solutions: Kept a master fielding schedule on Microsoft Teams detailing the start and end dates for tasks related to drawing and fielding sample waves.

Scheduled and assigned individual tasks for each wave in Microsoft Teams.

Operational Challenges

Overlapping Sample Waves–Example Master Fielding Schedule

	A	B	C	D	E
1	WAVE START/FINISH	Fielding Task	Due Date	Done (Y)	Scheduled in Te
2	**WAVE 1 START**	Wave 1 (PI only) Data pull/sample prep	9/23/2022	Y	
3		Wave 1 (PI only) NSF prenotice (from NSF.gov)	9/27/2022	Y	
4		Wave 1 (PI only) Insight invite (from Qualtrics/Insight)	9/27/2022	Y	
5	**WAVE 2 START**	Wave 2 (PI only) Data pull/sample prep	9/30/2022	Y	
6		Wave 2 (PI only) NSF prenotice (from NSF.gov)	10/3/2022	Y	
7		Wave 2 (PI only) Insight invite (from Qualtrics/Insight)	10/3/2022	Y	
8		Wave 1 (PI only) First reminder (from Qualtrics/Insight)	10/3/2022	Y	
9		Wave 1 (PI only) Second reminder (from Qualtrics/Insight)	10/12/2022	Y	
10		Wave 2 (PI only) First reminder (from Qualtrics/Insight)	10/12/2022	Y	
11	**WAVE 3 START**	Wave 3 (PI only) Data pull/sample prep	10/14/2022	Y	
12		Wave 3 (PI only) NSF prenotice (from NSF.gov)	10/17/2022	Y	
13		Wave 3 (PI only) Insight invite (from Qualtrics/Insight)	10/17/2022	Y	
14		Wave 1 (PI only) NSF reminder (from NSF.gov)	10/17/2022	Y	
15	**WAVE 1 FINISH**	Wave 1 (PI only) Last chance (from Qualtrics/Insight)	10/17/2022	Y	
16		Wave 2 (PI only) Second reminder (from Qualtrics/Insight)	10/17/2022	Y	
17		Wave 2 (PI only) NSF reminder (from NSF.gov)	10/24/2022	Y	
18	**WAVE 2 FINISH**	Wave 2 (PI only) Last chance (from Qualtrics/Insight)	10/24/2022	Y	

Exploratory Analysis of Recency Effects

Exploratory Analysis of Recency Effects

Dependent Variables (survey quality indicators)

- Respondent: Whether a case is a respondent or nonrespondent
- Time to start the survey: The lapsed time between when PIs and reviewers received the first survey invitation and when they clicked on the survey link
- Time to complete the survey: The lapsed time between when PIs and reviewers clicked on the survey link to when they submitted the survey
- Answered the last question: Whether a respondent answered the last question in the survey

Independent Variable

- Recency: The lapsed time between when PIs and reviewers engaged with the merit review process and when they received their first survey invitation

Exploratory Analysis of Recency Effects

Hypothesis

Lower recency values (i.e., less elapsed time between engaging with the merit review process and receiving the first survey invitation) will result in

- Higher response to the survey
- Starting the survey sooner
- Completing the survey sooner
- Answering all survey questions

Exploratory Analysis of Recency Effects

Results

PIs:

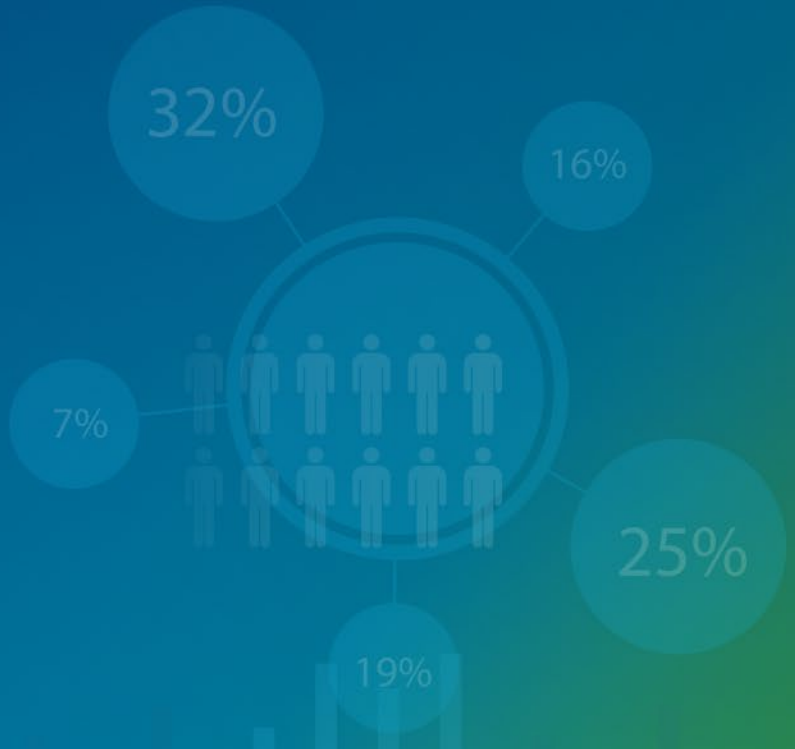
- Recency was not statistically associated with any of the survey quality indicators

Reviewers:

- Survey respondents had lower recency values
- Survey respondents started the survey sooner if they had lower recency values

Many of the associations operated in the hypothesized directions but were not statistically significant (p<.01).

Future Research



Future Research in NSF Surveys

- Does continuing to close the gap in recency improve survey quality?
- Does using prenotices improve response rates?
- Does including one additional email reminder result in a notable response rate increase?
- Can alternative email messaging strategies (e.g., alterations to the subject line and/or including response deadlines) compensate for sending fewer email reminders?

Conclusion

Conclusion

This work highlights survey data collection methods that begin to address a need at NSF to collect more regular survey data to inform data-driven decisions. Our results demonstrated the following:

- The use of programming scripts, a master fielding schedule, and quality assurance procedures allowed us to reliably collect NSF sample wave data and field 76 waves on a rolling basis over the course of 1 year
- Live survey methods can produce higher response rates, unbiased survey data, and may benefit survey quality more generally by encouraging survey response and reducing the time to start the survey

Q&A

Questions?

Study Team Contact Information

James McCall–jamesmccall@westat.com

Kelsey Gray–kelseygray@westat.com

Taylor Rhodes–mrhodes@nsf.gov

Photos are for illustrative purposes only. All persons depicted, unless otherwise stated, are models.