



FY 2004 - 2006 NIH Director's Pioneer Award Process Evaluation - Comprehensive Report

FINAL REPORT

January 2008

Prepared for:

G. Stephane Philogene, Ph.D.
Office of Behavioral and Social Sciences Research
National Institutes of Health
31 Center Drive, Room B1C32
Bethesda, MD 20892-2248

Prepared by

Bhavya Lal - Task Leader
Ritu Chaturvedi, Colleen Horin, Alexis Wilson, Elmer Yglesias
Arthur Fries, Ph.D. - Statistician, Darius Singpurwalla - Statistician,
Brian Zuckerman, Ph.D. - Reviewer

Science and Technology Policy Institute
1899 Pennsylvania Avenue NW, Suite 520
Washington DC 20006

Table of Contents

ACKNOWLEDGEMENTS	III
EXECUTIVE SUMMARY	V
1. INTRODUCTION.....	1
1.1 ORIGIN OF THE NDPA PROGRAM	1
1.2 EVOLUTION OF THE NDPA PROCESS (FY 2004 - FY 2006)	1
1.3 NDPA PROCESS EVALUATION	2
2. METHODOLOGY	3
2.1 PROCESS EVALUATION DESIGN.....	3
2.2 DATA SOURCES, COLLECTION METHODS, AND ANALYSIS	4
3. CHARACTERISTICS OF NDPA CANDIDATES (FY 2004 - FY 2006)	5
3.1 DEMOGRAPHIC CHARACTERISTICS OF FY 2004 - FY 2006 NDPA CANDIDATES	5
3.2 CHARACTERISTICS OF NDPA AWARDEES	11
3.3 REPEAT CANDIDATES	12
3.4 FUNDING SOURCES OF NDPA CANDIDATES	13
3.5 DEMOGRAPHIC BENCHMARKS	15
4. CHARACTERISTICS OF NDPA REVIEWERS (FY 2004 - FY 2006)	19
4.1 NIH LIAISONS	19
4.2 EXTERNAL EVALUATORS	19
4.3 ACCOMPLISHMENTS OF THE NDPA EXTERNAL EVALUATORS	23
5. NDPA PROCESS DESCRIPTION, FY 2004 - FY 2006	25
5.1 PHASE OVERVIEW	25
5.2 CANDIDATE PERCEPTION OF THE NDPA PROCESS.....	31
5.3 MODELING DETERMINANTS OF CANDIDATE SUCCESS FOR COMBINED YEARS	34
6. OVERALL ASSESSMENT	37
6.1 PROGRAM DESIGN.....	37
6.2 PROGRAM IMPLEMENTATION	38
6.3 CHARACTERISTICS OF CANDIDATES.....	39
7. RECOMMENDATIONS FOR FUTURE YEARS.....	43
7.1 PROGRAM DESIGN.....	43
7.2 PROGRAM IMPLEMENTATION	44
7.3 PROGRAM PARTICIPATION	45
APPENDIX: EVOLUTION OF THE NDPA PROGRAM FY 2004 - FY 2006: KEY CHANGES	47
ENDNOTES	51

Acknowledgements

The authors would like to thank the many individuals who have contributed to the design and implementation of this process evaluation. Specifically, the evaluation could not have been completed without the assistance of the many NIH staff and the numerous members of the wider scientific community who agreed to be interviewed about the FY 2004 - FY 2006 process.

We would like to thank the NIH Director's Pioneer Award Evaluation Advisory Committee (EAC) that helped guide the study and its methodology. Advisory Committee members included:

- Lawrence Fine: Leader, Clinical Prevention and Translation Scientific Research Group, Division of Epidemiology and Clinical Applications, National Heart Lung and Blood Institute
- Judith Greenberg: Director, Division of Genetics and Developmental Biology, National Institute of General Medical Sciences
- Teresa Levitin: Director, Office of Extramural Affairs, National Institute on Drug Abuse
- James Onken: Chief, Office of Program Analysis and Evaluation, National Institute of General Medical Sciences
- Walter Schaffer: NIH Research Training Officer, Office of Extramural Research, Office of the Director

We would also like to thank evaluation experts Ms. Margaret Blasinsky and Dr. Mary Dufour of the Madrillon Group who provided thoughtful feedback throughout the project. Lastly, and most importantly, we would like to thank Dr. Stephane Philogene, our technical officer, for his guidance and support.

Executive Summary

The National Institutes of Health (NIH) Director's Pioneer Award (NDPA) was initiated in Fiscal Year (FY) 2004 to support individual investigators who display the creativity and talent to pursue high-impact ideas in the biomedical sciences. As a pilot, the NDPA continues to evolve and change both structurally and conceptually; annual process evaluations are conducted to track program implementation and the selection process and to inform future years of the program. This report summarizes the first three years of NDPA's design, implementation, and participation.

Three main sources of data were considered for the annual process evaluations: (1) NIH administrative data including funding data; (2) interviews with NIH liaisons, external evaluators and panelists; and (3) surveys of all candidates considered for an award.

Program Design

With time, the NDPA program design adapted to incorporate many lessons learned from prior years of implementation. The review criteria and program structure (application materials and outreach) were adjusted to promote consistency in the external evaluators' reviews and the candidate submissions. In FY 2005 the leadership and motivation criteria were dropped; however, discomfort with some of the criteria continues. For example, the newly introduced "suitability" criterion was seen as a "Catch-22" by some evaluators who indicated that they were asked to seek individuals who typically would not be funded under traditional NIH mechanisms, but found that exceptional researchers were frequently already within the NIH fold. Some evaluators were also uneasy with the guideline that requires candidates to demonstrate the ability to devote 51% or more effort on the NDPA project.

In each of the three years, all evaluators and liaisons were trained to implement the new criteria, and care was taken to ensure that all understood the meaning and proper application of the criteria. The majority of the evaluators and candidates who responded to the survey were satisfied with the general effectiveness and relevance of the criteria specified for identifying a pioneer.

The addition of the "scientific problem to be addressed" review criterion in the second year of the program shifted the program away from a purely "people-based" approach to a mix of a people- and "project-based" approach. The shift was intended to make the criteria easier to operationalize - several of the criteria from FY 2004 (e.g., leadership) were subjective and difficult to apply consistently. However, the shift moved the program away from the purely people-based approach, a cornerstone of the program when initially designed. As a result, there was clear division on methods for deciding whether or not an application was competitive. Specifically, evaluators and liaisons alike were unsure if their purpose was primarily to look for a creative, "pioneering" individual or an innovative project idea.

Program Implementation

The FY 2004 - FY 2006 NDPA processes were largely implemented as designed; going forward, however, the process needs still more clarity and consistency. Overall, most evaluators and candidates were satisfied with the application format and materials submitted. Because evaluators in later years were unsure if their purpose was primarily to look for a pioneering

individual or an innovative project, they made their choices differently from each other, introducing unpredictability in the process. The “guidelines” to look for underrepresented groups, those with 51 % time for NDPA, and junior researchers were also applied inconsistently, with some evaluators using them more rigorously and others ignoring them entirely. The essays and the biographical sketches were viewed as the most useful components to display “pioneering” qualities. Evaluators and candidates recommended that the letters of reference should be revised or eliminated altogether in the future.

Despite some perceptions by participants that the selection process favored certain groups, statistical analyses did not indicate significant relationships between candidate characteristics and the selection of finalists, and scoring became more consistent over the years. The results of a series of stepwise logistic regressions indicated that top-4 and average overall score had a significant relationship with proceeding to the interview stage. With one exception (seniority), statistical differences by gender, race, research area, affiliation, degree, and level of NIH funding were not found. Statistical analyses also did not indicate that being a repeat candidate increased the odds of being selected for an interview. The NDPA selection process was not entirely transparent in any given year. The most common criticism for FY 2004 - 2006 NDPA competitions was the lack of evaluator feedback to candidates. NIH typically provides feedback to applicants in other programs; NDPA candidates stated that without evaluator feedback, it is difficult to understand the reasons they did not move on in the process, or to improve their proposals for future resubmissions. This lack of transparency contributes to the perception by some candidates that reviewers have something to hide, such as preference for or against certain groups (even though these perceptions were not supported by statistical analysis). A second transparency concern was raised by panelists about the selection of awardees in FY 2005 and FY 2006. After the interview phase and panelist rankings, the finalists were presented to NIH Institute Directors in order to secure funding for additional awards. While this resulted in nearly twice as many awards in FY 2005 and FY 2006 as would have been possible strictly through the NIH Roadmap funds, it also resulted in the funding of several individuals from the “middle” and “bottom tiers” of the interviewees. Several panelists were concerned that individuals from the bottom tier were funded, as this “defeated the purpose of the panel review.”

Program Participation:

FY 2004 - FY 2006 NDPA Candidates: Although the total number of candidates applying for NDPA has declined steeply over the years, available information shows that the characteristics of the candidates, excluding seniority, have remained largely unchanged. While the average seniority of the candidates has modestly increased, the number and percent of junior investigators (those with fewer than 10 years of experience since first terminal degree) has decreased in each year of the program. The distribution of the candidates across the three years resembled the general distribution in the biomedical research community. The candidates were drawn broadly from a range of biomedical science fields (including behavioral and social sciences, and engineering, in addition to the more traditional biological science disciplines), with the majority applying from universities or university-affiliated medical institutes.

In FY 2005 and FY 2006, more than one fifth of the total candidates were repeats (i.e., had applied in previous years). Many of the candidates who did not reapply stated that there was too much competition for a small number of awards, while others thought the process was unfair and therefore not worth their while to reapply. More than a third of the repeat candidate survey respondents made substantial changes to their ideas, or submitted entirely new ideas in subsequent years of the process.

Candidates were generally well-funded by NIH, and the majority of the survey respondents had applied to the NIH at some point in their career.

FY 2004 - FY 2006 NDPA Evaluators: Overall, the NDPA external evaluators were an accomplished group of researchers in their respective fields. Though it is difficult to judge the “quality” of the evaluators, based on publicly available information, the evaluators were well-known researchers - more than two-thirds have won major awards in their fields or have received prestigious fellowships or honors. In FY 2004 - FY 2006, the evaluators were: predominately male, though the proportion of female evaluators increased in FY 2005 and FY 2006 due to targeted recruitment; primarily drawn from universities or university-affiliated medical institutes; relatively senior (the majority obtained degrees more than 25 years ago); and well-matched to the research area distribution of the candidate pool.

FY 2004 - FY 2006 NDPA Awardees: Through NDPA, NIH aimed to attract researchers who typically did not apply to the NIH as well as new ideas and approaches. Thus far, the program has attracted some new candidates who were not already part of the NIH system, though the majority of candidates and awardees had received or applied for NIH funding in recent years. With respect to the presence of pioneers and pioneering ideas, both candidates surveyed and evaluators interviewed believed that NDPA attracts some higher-risk ideas than would be submitted to a standard study section. In fact, a majority of the candidates - not all of them awardees - who responded to the survey indicated that they would be unlikely to receive funding for their proposed NDPA idea through a traditional mechanism.

Many external evaluators, however, asserted that the candidate and awardee pools were a “mixed-bag.” They saw some “truly pioneering” candidates and awardees, but also others who clearly should have been applying for an R01 or other type of award. Evaluators did affirm that the NDPA process itself was very different from a traditional study section, and conducive to allowing investigators to submit more innovative and creative applications.

Generally the program was viewed as having a favorable inception; however, it will take several years to determine whether awardees' research is pioneering. Perhaps, more importantly than success, participants believed that the program sends an essential message about risk-taking to the research community.

Key Recommendations

Evaluation findings from the first three years of the NDPA program suggest both strategic and structural changes to the program. Principal recommendations include:

Program Design:

- NDPA program leadership should continue to clarify program criteria and operationalization.
 - NDPA program leadership should clearly articulate whether NDPA is primarily intended to identify the best people, the best ideas, or a mixture of both. This would make the program purpose clear to all participants (candidates and evaluators alike).
 - Program leadership should also revisit the “suitability” criterion, and consider operationalizing it better so it can be applied more consistently by reviewers.
 - If NDPA wants to continue the practice of requiring that awardees be able to devote at least 51% of their time to their NDPA research, the requirement should be applied consistently, and at the same stage. Candidates should be asked to indicate at the application stage whether or not they can devote at least 51% of their time to NDPA if they win. After this initial consent, the candidates’ commitment should not be questioned by multiple evaluators at different phases of review.

- Both the absolute number of qualified junior investigators as well as their proportion in the candidate pool has been shrinking; assuming the emphasis on junior investigators persists, new strategies should be developed to attract them in future years.

Program Implementation:

- Given feedback from participants – both candidates and reviewers – regarding concerns about fairness of the program, NDPA leadership may need to take steps to increase consistency, communication, and transparency of the selection process.
 - NDPA leadership should consider providing more feedback to candidates who do not win an award.
 - NDPA leadership should consider requiring evaluators to justify their scores (perhaps using a standardized template to decrease burden, and/or free-form comments). These comments may help NDPA provide feedback to the applicants as well.
 - In future years, if funding is sought from NIH Institutes and Centers (ICs) for additional awardees, then that selection process should seek to support and reward the remaining top-tier and middle-tier finalists before funding bottom-tier finalists.

Program Participation:

- Attracting a diverse pool of investigators to apply to the program should remain a top priority for NDPA leadership. Once a diverse pool of candidates has been attracted, guidelines that encourage evaluators to pay special attention to underrepresented groups, junior candidates, etc. may be redundant.

1. Introduction

1.1 Origin of the NDPA Program

The National Institutes of Health (NIH) Director's Pioneer Award (NDPA) was initiated in Fiscal Year (FY) 2004 to support individual investigators who display the creativity and talent to pursue high-risk, potentially high-impact ideas in biomedical sciences. The program grew out of concerns that the traditional peer review process is overly conservative and that NIH may require additional means by which to fund high-risk research.^{i,ii,iii,iv} On the premise that great ideas are driven by an individual, and not necessarily by a work plan, the program aimed to find researchers who have the skills and the creativity to take productive risks and to make significant contributions to medical research.^v A secondary goal of the program was to identify investigators not typically seeking funding through traditional NIH mechanisms.^{vi}

From the onset of the program, NIH leadership believed that if an alternative mechanism was going to work and convince the outside community that NIH was serious about wanting to fund research in a new way, then it had to have a vehicle that looked different from anything that NIH had previously done. As a result, NIH chose to minimize the role of existing NIH bodies (e.g., study sections, Center for Scientific Review) in the NDPA process and to focus on new methods (e.g., nominations, abbreviated applications) through which to attract high-impact ideas and creative individuals. Several of the distinctive characteristics of the NDPA program initially included:

- It was run centrally, out of the Office of the Director (OD)
- The application process was short relative to other NIH programs, with no requirement for detailed project plans or discussion of preliminary data. Abbreviated candidacy forms were used throughout
- *Ad hoc* committees of extramural evaluators were convened to evaluate applications, as opposed to the traditional study sections used to evaluate investigator-initiated research applications
- A multistage process with a phase-specific scoring system was used instead of the peer review priority scoring system
- External review was conducted electronically with no face-to-face interaction until the final interview phase
- It was designed to be a “people-based” program
- The NIH Director was personally involved in the selection of awardees, with the Advisory Committee to the Director serving as a secondary review body

In September 2004, the first Pioneer Awards were made, with nine individuals receiving funding under the DP1 activity code.

1.2 Evolution of the NDPA Process (FY 2004 - FY 2006)

Because the NDPA program was designed as a pilot, changes were made as lessons were learned from the initial years of implementation. In both FY 2005 and FY 2006, changes were made in the nomination and selection process. The changes have been both structural (e.g., changes in the number of phases, rating system, submission interface) and conceptual (e.g., changes in selection criteria and program emphasis). The Appendix table highlights the

changes across the years in terms of the program emphasis, definition of a “pioneer,” selection process, selection criteria, application materials, and the research areas of candidates.

1.3 NDPA Process Evaluation

Following the first round of awards, the NIH OD commissioned an independent process evaluation of the NDPA program. In subsequent years, two more cohorts of awards were made and process evaluations conducted. The fundamental goals of these annual process evaluations were to:

- Assess the NDPA award selection process
- Determine if the NDPA program was implemented as designed; and
- Determine if the selection process was consistent with program goals

This report summarizes the findings from the first three years (FY 2004 - FY 2006) of the process evaluation and is divided into seven chapters and an Appendix: Chapter 2 summarizes the process evaluation methodology; Chapters 3 and 4 describe the characteristics of the candidates^{vii} and the reviewers^{viii}; Chapter 5 provides an overview of the NDPA process; Chapters 6 and 7 summarize the overall assessment of the program and recommendations; and the Appendix summarizes elements of and changes to the NDPA selection process from FY 2004 to FY 2006.

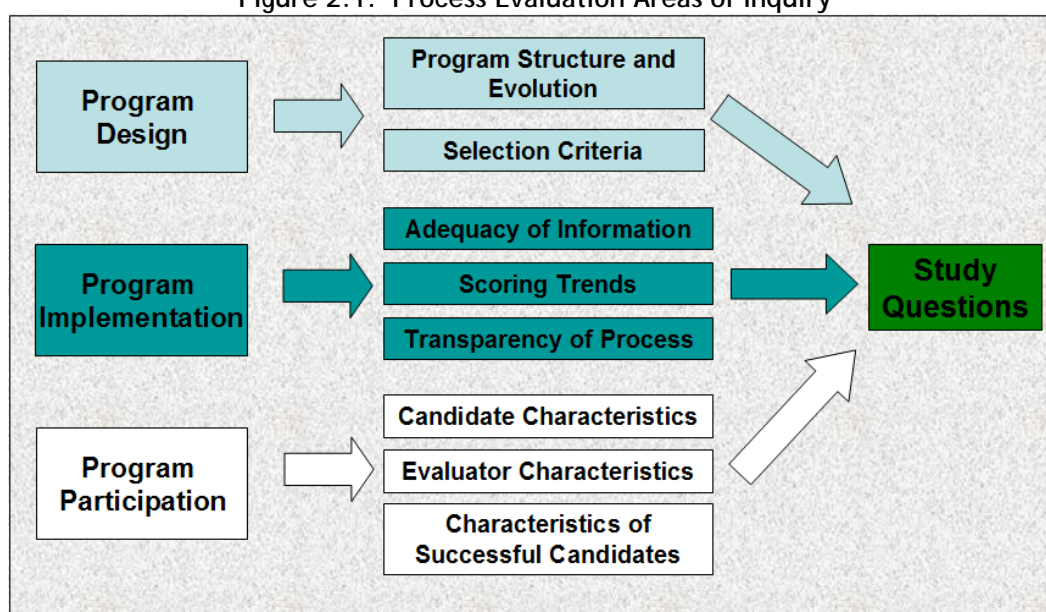
This comprehensive report builds on individual data reports for each of the three years. Each data report includes detailed descriptions of the origin, conduct and evolution of the program during its first three years of implementation. Data reports can be obtained upon request from the NIH.

2. Methodology

2.1 Process Evaluation Design

The process evaluation was designed to study NDPA’s operations in light of program goals and to provide recommendations for how program activities could be improved. Two conceptual models drove this process evaluation: first, a process “flow” that outlined the NDPA process in each individual year; second, a stakeholder map that highlights individuals involved at each phase. Based on the process flow and stakeholder maps, a set of detailed study questions was developed. The study questions and findings were organized by three categories: Program Design, Program Implementation, and Program Participation as illustrated in Figure 2.1.

Figure 2.1: Process Evaluation Areas of Inquiry



In each year, study questions, data collection approach, and interim findings were reviewed by both the NIH Evaluation Officer and the NIH Director’s Pioneer Award Evaluation Advisory Committee¹. The high-level study questions that guided the process evaluation were:

¹ To guide the study and its methodology, the Office of the Director/Office of Behavioral and Social Science Research (OD/OBSSR) convened a six-member NIH Director’s Pioneer Award Evaluation Advisory Committee (EAC). EAC members are: *Lawrence Fine*: Leader, Clinical Prevention and Translation Scientific Research Group, Division of Epidemiology and Clinical Applications, National Heart Lung and Blood Institute; *Judith Greenberg*: Director, Division of Genetics and Developmental Biology, National Institute of General Medical Sciences; *Teresa Levitin*: Director, Office of Extramural Affairs, National Institute on Drug Abuse; *James Onken*: Chief, Office of Program Analysis and Evaluation, National Institute of General Medical Sciences; *Walter Schaffer*: NIH Research Training Officer, Office of Extramural Research, Office of the Director; *Stephane Philogene* (Executive Secretary and Evaluation Officer): Office of Behavioral and Social Science Research, Office of the Director.

Program Design

- *Program Structure and Evolution:* What was the overall structure of the selection process in each individual year? How and why did the NDPA program evolve from the preceding year?
- *Selection Criteria:* How were the characteristics of “pioneering research” defined and operationalized as selection criteria? How did the selection criteria evolve over the years?

Program Implementation

- *Adequacy of Information:* To what extent was the information available to the evaluators adequate to select the best applications?
- *Scoring Trends:* What were the trends in scoring by phase and other attributes of interest?
- *Transparency of Process:* To what extent did the candidates and evaluators understand the process?

Program Participation

- *Candidate Characteristics:* What were the characteristics of the nominees/applicants who applied to the program?
- *Evaluator Characteristics:* What were the basic characteristics of the internal and external evaluators?
- *Characteristics of Successful Candidates:* What were the characteristics of the candidates who were successful in advancing at each phase and who won the award?

2.2 Data Sources, Collection Methods, and Analysis

Three main data sources were used in each year to conduct the annual process evaluation:

1. Administrative data from the NIH - To gain insights into nominee characteristics, reviewer scores, and comparison mechanisms, data were obtained on candidates' demographic and other characteristics, scores, and prior funding history from NIH databases.
2. Interviews with NIH liaisons and external evaluators - During the first three years of the process evaluation, a total of 149 interviews were conducted to gain insights about satisfaction with the process.
3. Surveys of all candidates considered for an award in FY 2004, FY 2005, and FY 2006 - over the years, a total of 1,161 NDPA candidates completed the survey (63% of delivered surveys). The annual breakdown is summarized in Table 2.1.

Table 2.1:
Survey Response Rates by Year FY 2004 - FY 2006

Respondent Type	Surveys Delivered	Surveys Completed	Response Rate
FY 2004 Candidate*	677	411	61%
FY 2005 Candidate	705	420	60%
FY 2006 Candidate	456	330	72%
FY 2004 - FY 2006 Candidate Total	1838	1161	63%

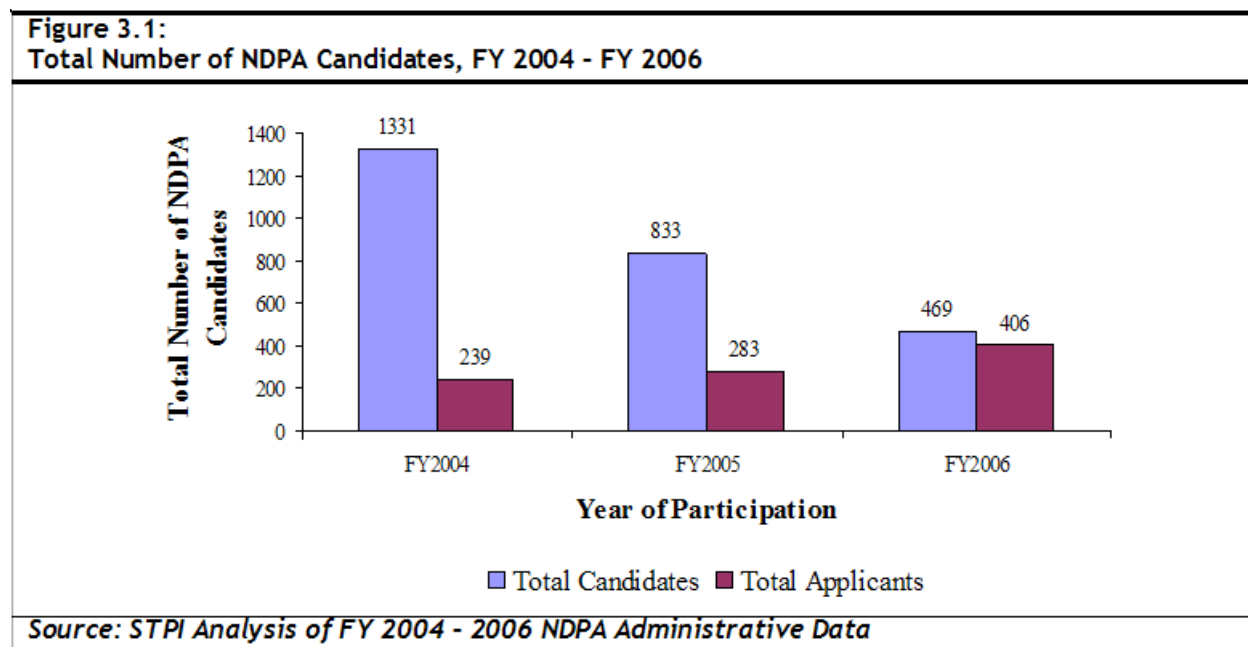
*Note: A total of 1,444 surveys were sent in FY 2004; however, 767 surveys were sent to nominators rather than to the candidates themselves and are not included in this table.

Source: STPI Analysis of FY 2004 - FY 2006 NDPA Survey Data

3. Characteristics of NDPA Candidates (FY 2004 - FY 2006)

3.1 Demographic Characteristics of FY 2004 - FY 2006 NDPA Candidates

In its first three years of implementation, the total number of candidates applying for the NDPA award has steadily declined from 1,331 candidates in FY 2004 to 469 candidates in FY 2006. In all, 2,311 individuals have applied to the program, though some applied in multiple years, for a total of 2,633 individual submissions. Although the total number of “candidates” dropped, the number of “applicants” (candidates whose applications were reviewed by the external evaluators) has increased in every year (Figure 3.1 and Appendix table).

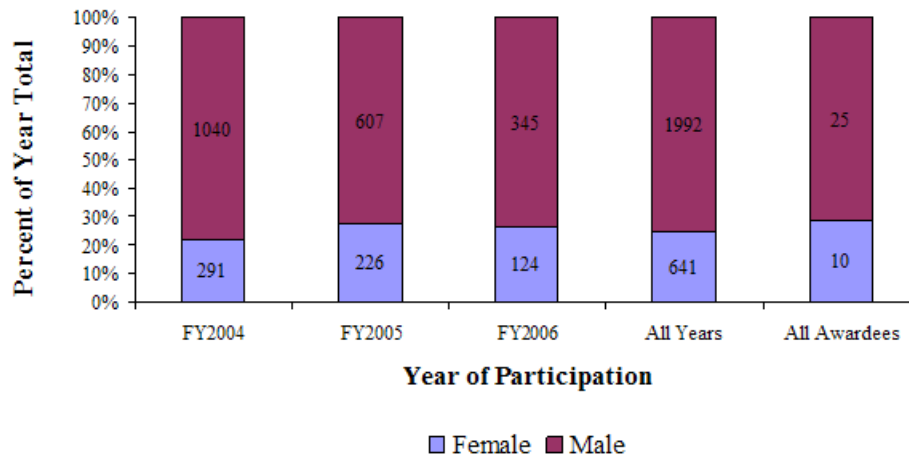


While the total candidate *numbers* have declined over the years, available data show that the demographic *characteristics* of the candidates have remained largely unchanged (with the exception of seniority).

3.1.1 Gender Distribution

In all years of implementation, NDPA was intended to attract as many candidates from as diverse a set of backgrounds as possible. After FY 2004, a new emphasis was placed on encouraging younger investigators as well as women and minorities to apply. This was made clear in the program notice, “those at early to middle stages of their careers, and women and members of groups underrepresented in biomedical research are especially encouraged to nominate themselves,” and additionally in the instructions to the reviewers when they were told to “watch for women, minorities, investigators at early to middle career stages.” The decision to make the competition open only to self-nominees (rather than self-nominees and those nominated by others) was at least partially driven by community and liaison feedback that women and underrepresented minorities may be disadvantaged in a process where senior mentors are involved in the nomination process.^{ix} Across all years, women have comprised approximately one quarter of the total candidate pool (Figure 3.2).

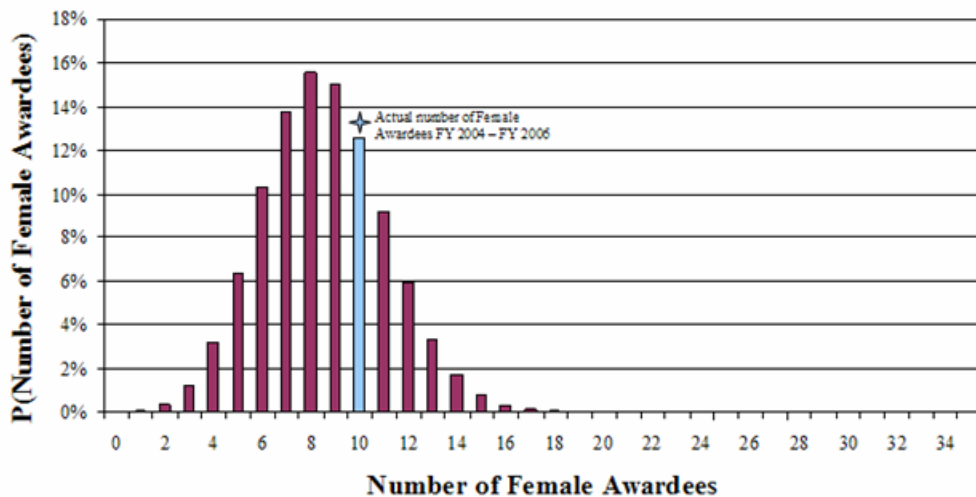
Figure 3.2:
Gender Distribution of NDPA Candidates, FY 2004 - FY 2006



Source: STPI Analysis of FY 2004 - 2006 NDPA Administrative Data

Though there were no female awardees in FY 2004, for all years collectively, there is no significant difference between the total number of female awardees and the expected number based on the total candidate pool (assuming that the selection process was completely random; chi-square test, $\chi^2=0.34$, $df=1$, $p=0.6$). The median expected number of female awardees is eight. The actual number of women awardees is in fact higher than expected given the initial male/female ratio of NDPA candidates (Figure 3.3).

Figure 3.3:
Binomial Distribution: Probability of Resulting Number of Female Awardees Given Total Male/Female Ratio of Candidates, FY 2004 - FY 2006



Source: STPI Analysis of FY 2004 - 2006 NDPA Administrative Data

3.1.2 Seniority

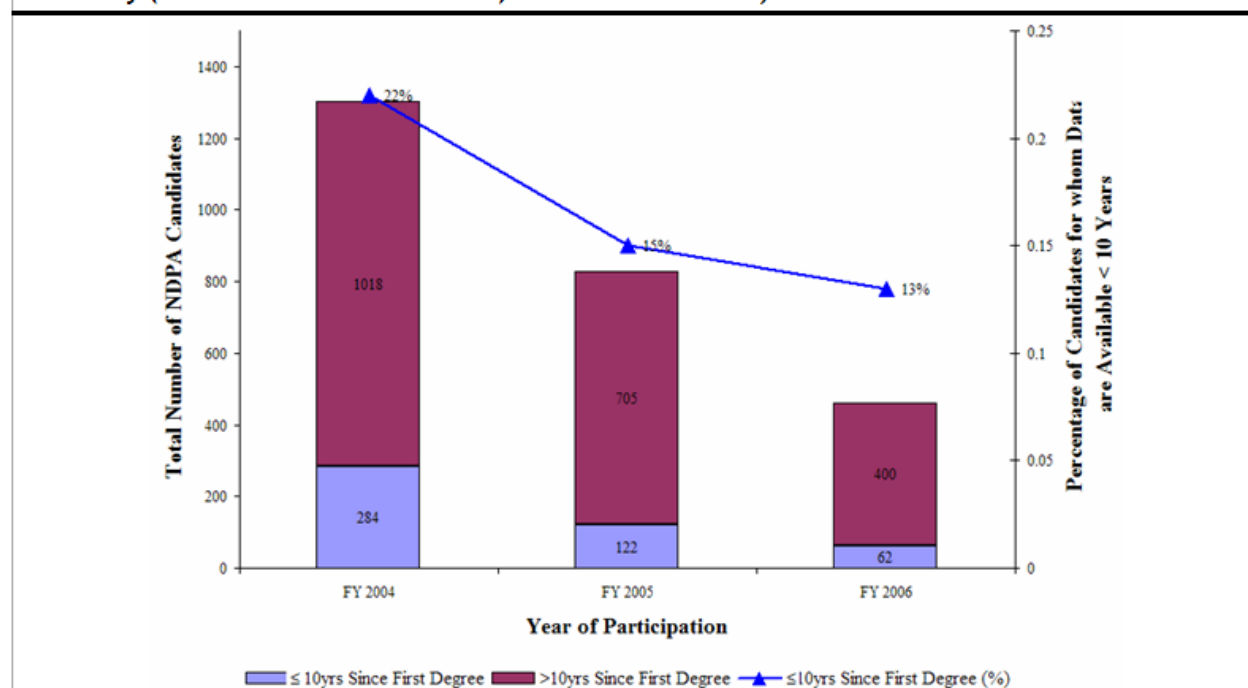
The majority of all FY 2004 - FY 2006 NDPA candidates hold PhDs (68%), though some hold MDs (19%) or joint MD/PhDs (11%).^{x,xi} Recipients were placed into seven categories based on the year their first doctoral-level degree was obtained: < 0 years - no doctoral degree or doctoral

degree in progress, 0-5 years since degree, 6-10 years since degree, 11-15 years since degree, 16-20 years since degree, 21-25 years since degree, and > 25 years since degree.

Over the years, the proportion of candidates with fewer than 10 years of experience has been declining from 284 (22% of total candidate pool for whom seniority data were available) in FY 2004 to 62 (13% of the total candidate pool for whom seniority data were available) in FY 2006 (Figure 3.4). Average seniority of the candidates has gradually increased (19.8 years in FY 2004, 20.8 years in FY 2005, and 21.6 years in FY 2006 (Figure 3.5)

Awardee distribution, however, is somewhat different from that of the candidates. In all years, the awardees were on average less senior than the total candidate pool, with the largest divergence occurring in FY 2005. The FY 2005 awardees were on average four years less senior than their counterparts in FY 2004 or FY 2006, perhaps attributable to the new emphasis on early- to middle-career stage first introduced in the second year of program implementation (Figure 3.5).

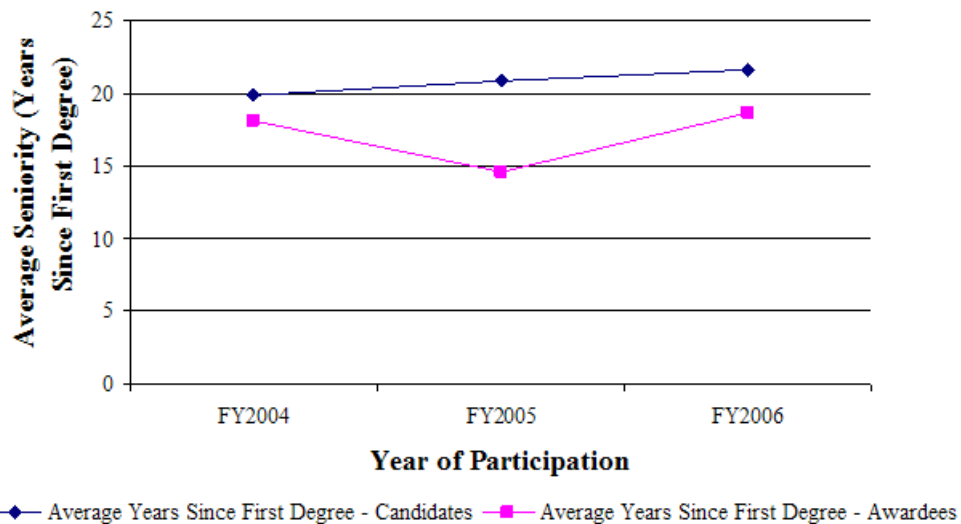
Figure 3.4:
Seniority (Years Since First MD or PhD) of NDPA Candidates, FY 2004 - FY 2006*



**Note: Numbers and Percentages in Figure 3.4 are based on available data only; therefore, the column totals do not sum exactly to the year totals. Also, candidates with no higher degree are not included in the figure (data missing or not relevant for 29 (2%) candidates in FY 2004, 6 (1%) candidates in FY 2005, and 7 (1.5%) candidates in FY 2006).*

Source: STPI Analysis of FY 2004 - 2006 NDPA Administrative Data

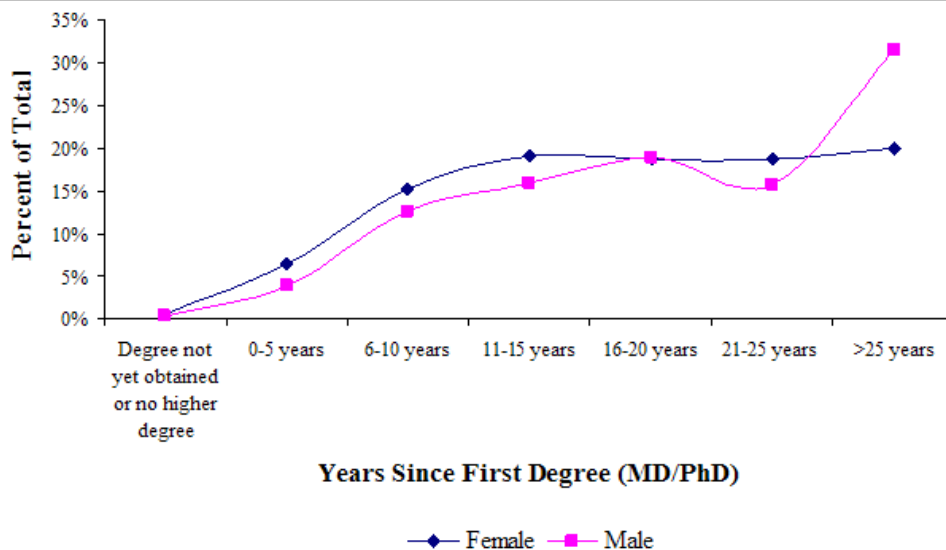
Figure 3.5:
Average Seniority (Years Since First MD or PhD) of NDPA Awardees, FY 2004 - FY 2006



Source: STPI Analysis of FY 2004 - 2006 NDPA Administrative Data

In all years, male candidates were more senior than their female counterparts; there was a significant difference between the seniority of male and female candidates over all three years of implementation ($t(2591) = 6.2, p < 0.0001$), with females averaging 18.2 years and males averaging 21.2 years since first degree. Seniority distribution by gender is displayed in Figure 3.6.

Figure 3.6:
Seniority (Years Since First MD or PhD) of NDPA Candidates by Gender, FY 2004 - FY 2006



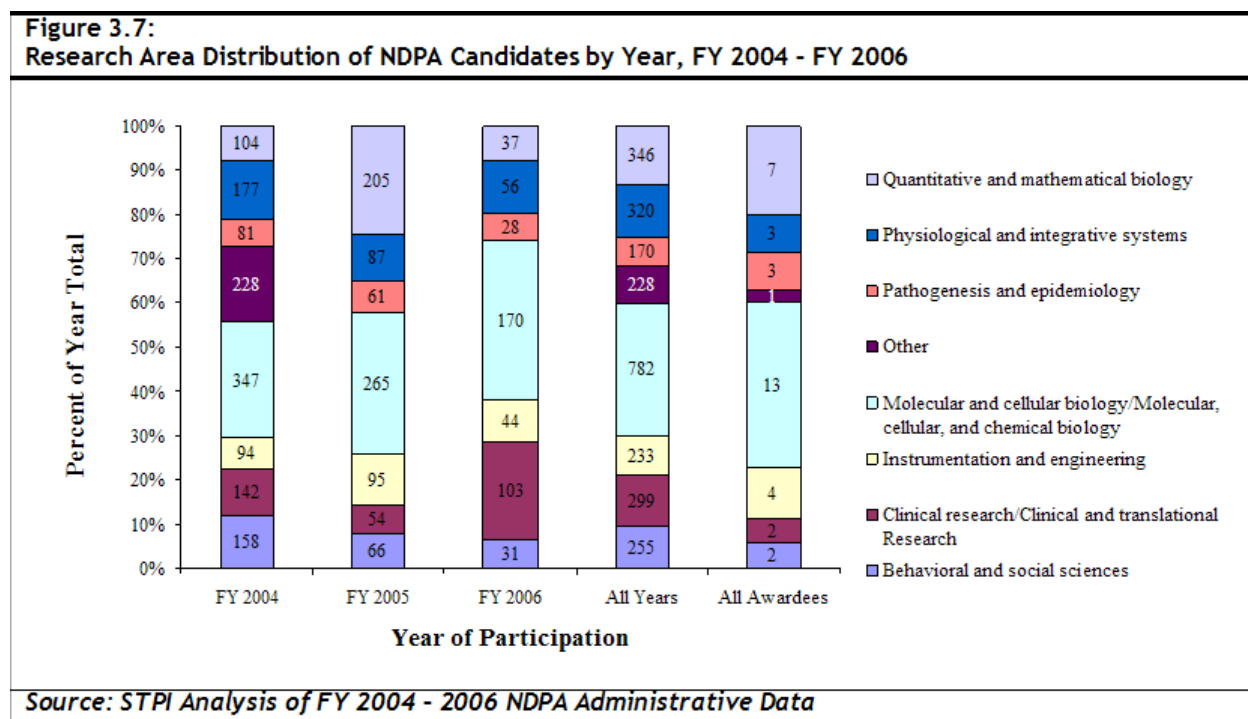
Source: STPI Analysis of FY 2004 - 2006 NDPA Administrative Data

3.1.3 Research Areas

As part of the nomination process, candidates were asked to categorize their research into one of seven categories:

1. Molecular and Cellular Biology (renamed Molecular, Cellular, and Chemical Biology in FY 2006)
2. Quantitative and Mathematical Biology
3. Instrumentation and Engineering
4. Physiological and Integrative Systems
5. Behavioral and Social Sciences
6. Pathogenesis and Epidemiology
7. Clinical Research (renamed Clinical and Translational Research in FY 2006)

All of the nominees chose one of these categories; however in FY 2004 candidates were also able to designate an eighth “other” category instead of one of the seven categories. In FY 2005 all candidates were required to select a primary field of research, but could also designate a secondary “other” category to add more detail. The specific descriptions given in the “other” categories were coded by STPI staff. Most of the research areas specified as “other” could be grouped into one of the original seven given categories. Other common categories specified included Neuroscience, Biophysics/Bioengineering, and Genomics/Bioinformatics, and Biochemistry. Overall, Molecular and cellular biology/Molecular, cellular, and chemical biology was the most common field of research in all years (Figure 3.7). Both male and female nominees were similarly distributed by research area (data not shown).



3.1.4 Affiliation

Overall, the large majority (85%) of NDPA candidates were drawn from universities or university-affiliated medical institutes (data not shown). More than one-quarter of all candidates were drawn from eleven institutions (Table 3.1).

Table 3.1:
Most Common Institutional Affiliations of FY 2004 - FY 2006 NDPA Candidates*

Institutional Affiliation	FY 2004	FY 2005	FY 2006	All Years	% Total
Harvard University	90	64	26	180	7%
Stanford University	40	20	15	75	3%
Johns Hopkins University	31	20	14	65	2%
University of Washington	24	23	7	54	2%
Columbia University	27	14	12	53	2%
University of Pennsylvania	23	22	8	53	2%
University of California-Los Angeles	23	17	9	49	2%
University of Michigan	30	10	5	45	2%
University of California-San Francisco	20	12	9	41	2%
Yale University	26	10	5	41	2%
Duke University	21	10	6	37	1%
Subtotal	355	222	116	693	26%

**Note: Institutional affiliations are generalized and include affiliated hospitals and research centers in addition to the central university. It also should be noted that the affiliations account for the researchers' location at the time of applying for the NDPA. For instance, at the time of the award, one individual was at California Institute of Technology, but moved to Stanford shortly after receiving the award - this individual is listed as affiliated with the California Institute of Technology.*

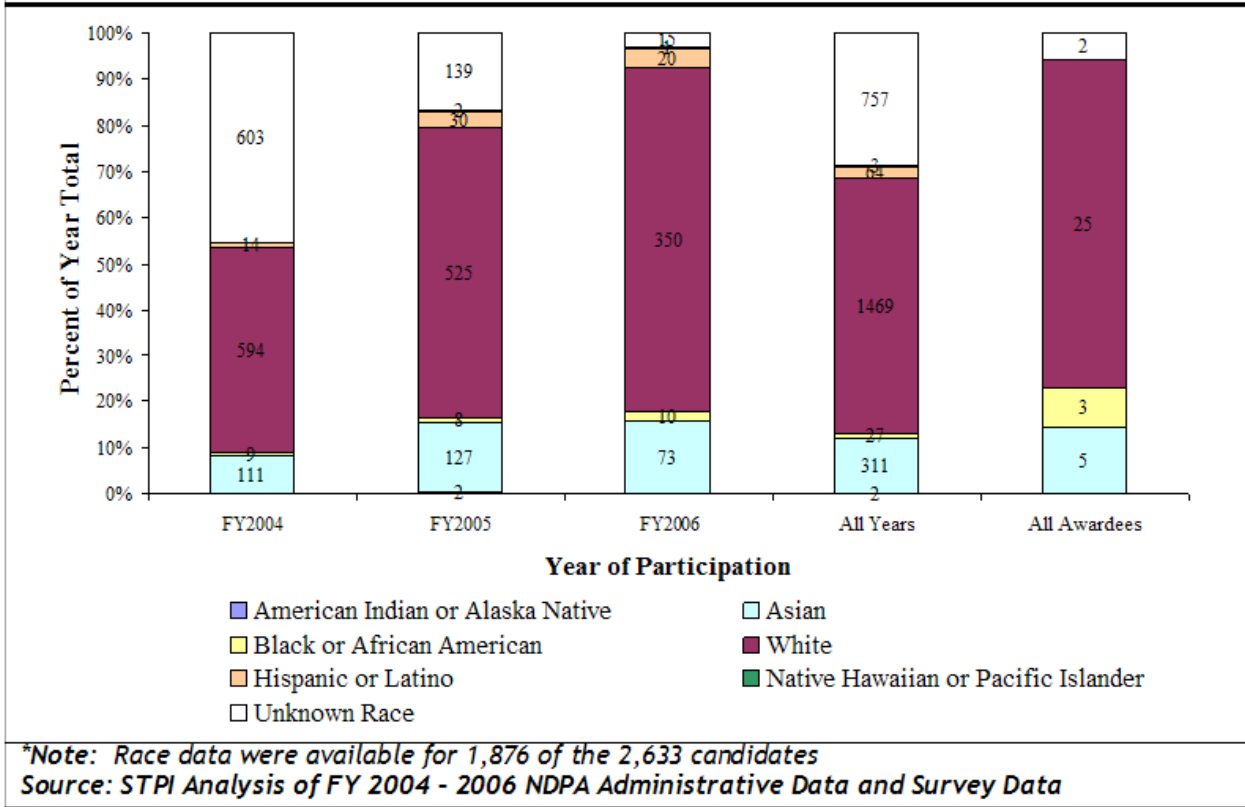
Source: STPI Analysis of FY 2004 - FY 2006 NDPA Administrative Data

3.1.5 Race/Ethnicity

Of the 2,633 total candidates, race information was obtained for 1,876.^{xii} Of the candidates, 56% were White, 12% were Asian, 2% were Hispanic or Latino, 1% were Black or African American, and fewer than 1% were Native Hawaiian or other Pacific Islander or American Indian or Alaska Native (Figure 3.8).^{xiii} No data were available for the remaining 29% of the candidates. Ethnicity data were obtained for 1,692 of the 2,633 total candidates. Overall, 62% were Not-Hispanic or Latino, 36% were of unknown ethnicity, and 3% were Hispanic or Latino (data not shown).

Though under-represented groups only represented approximately 5% of the total candidate pool for which race data were identified, the representation is close to the overall NIH funding distribution of under-represented minorities. Section 3.5 below benchmarks NDPA against other NIH programs and the general biomedical research community.^{xiv}

Figure 3.8:
Race of NDPA Candidates by Year, FY 2004 - FY 2006*



3.2 Characteristics of NDPA Awardees

Though some of the demographic characteristics of NDPA awardees are summarized in Section 3.1 together with all candidates, other characteristics include:

- Most are basic researchers with all but one holding a PhD or MD/PhD (only one has an MD alone).
- More than a third declared that their research idea was primarily in the field of molecular and cellular biology, and more than a fifth in quantitative and mathematical biology.
- Most are at universities or university-affiliated medical institutes in the United States (except one who is in England); most are affiliated with science-oriented departments but five are at engineering/bioengineering departments.
- NDPA awardees are drawn from elite institutions. Awards were given to individuals at 23 different institutions, with the most awardees affiliated with Stanford University (Table 3.2).
- Eleven of the awardees obtained their final degrees from Harvard or Harvard-affiliated institutions. Five have degrees from universities abroad.

Table 3.2:

Institutional Affiliations of FY 2004 - FY 2006 NDPA Awardees

Institutional Affiliation	FY 2004	FY 2005	FY 2006	All Years
Stanford University	0	3	3	6
Harvard University	2	1	1	4
California Institute of Technology*	2	0	0	2
Duke University	1	1	0	2
Massachusetts Institute of Technology	0	0	2	2
University of Texas - Southwest Medical Center	1	0	1	2
Brandeis University	1	0	0	1
Cold Spring Harbor Laboratory	0	1	0	1
Johns Hopkins University	0	1	0	1
New York University	0	0	1	1
Northwestern University	1	0	0	1
Rockefeller University	0	1	0	1
Scripps Research Institute	0	1	0	1
University of Arizona	0	1	0	1
University of California-Berkeley	0	0	1	1
University of California-San Francisco	1	0	0	1
University of California-Santa Barbara	0	1	0	1
University of Cambridge	0	1	0	1
University of Massachusetts-Amherst	0	0	1	1
University of North Carolina	0	0	1	1
University of Texas - Houston	0	0	1	1
University of Washington	0	0	1	1
University of Wisconsin-Madison	0	1	0	1
Total	9	13	13	35

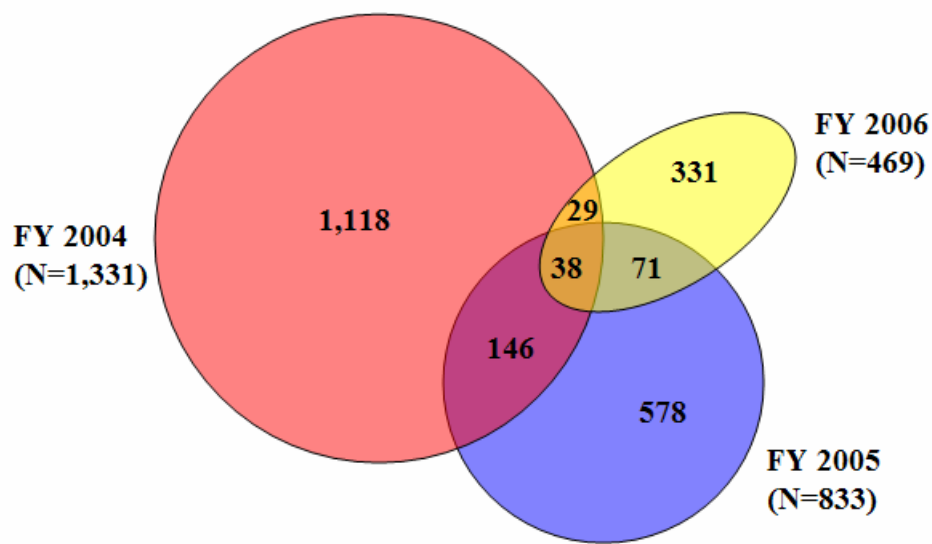
**Note: At the time of the award, one individual was at California Institute of Technology, but moved to Stanford shortly after receiving the award.*

Source: STPI Analysis of FY 2004 - FY 2006 NDPA Administrative Data

3.3 Repeat Candidates

As mentioned previously, many of the candidates reapplied to NDPA after not meeting initial success. In FY 2005, of the 833 candidates, 184 (22%) also applied in FY 2004; in FY 2006, 138 of the 469 (29%) applied in at least one previous year (Figure 3.9). Men and women were almost equally as likely to reapply in both FY 2005 and FY 2006.

**Figure 3.9:
Participation by Candidates in Multiple Years of NDPA**



Source: STPI Analysis of FY 2004 - 2006 NDPA Administrative Data

The FY 2005 and FY 2006 repeat applicants were surveyed about how their submission changed from one year to the next. Overall, 46% of the total respondents who applied multiple times made no changes, or minor changes to the resubmissions, and 40% made substantial changes or submitted a completely new idea. Chapter 5 addresses whether repeat candidates have a greater probability of success in the program.

3.4 Funding Sources of NDPA Candidates

Two methods were employed to determine if the NDPA was attracting new researchers to the NIH (one of the initial goals of the program). First, an IMPAC II search was performed for all FY 2004, FY 2005, and FY 2006 candidates to determine if they had received NIH funding in the five years prior to applying to NDPA. The IMPAC II search revealed that the majority of NDPA candidates in every year have received funding from the NIH in recent history. In FY 2006, three quarters of the candidate pool held NIH grants in the past five years. Similarly, in all years, the majority of the awardees were not new to the NIH system. In every year, there were two awardees who did not have NIH funding in the past five years (Table 3.3).

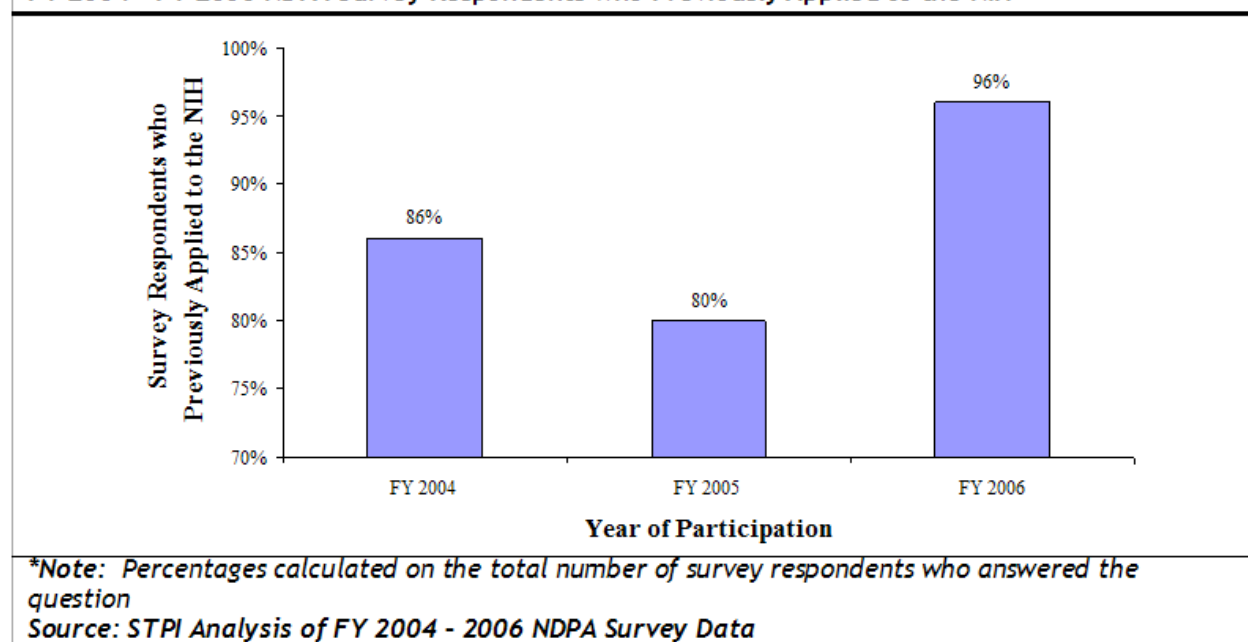
Table 3.3:
Funding Details of FY 2004 - FY 2006 NDPA Candidates and Awardees

<i>Candidate funding history</i>	FY 2004	FY 2005	FY 2006
Number (percent) of <i>candidates</i> who had a minimum of one NIH award as a PI in the last five years	824 (62%)	498 (60%)	354 (75%)
Total value of all NIH awards to NDPA <i>candidates</i> in the five years preceding NDPA application (in billion dollars)	\$2.6	\$1.5	\$1.2
Average number of unique awards per <i>candidate</i> (of the candidates with a minimum of one NIH award in the last five years)	2.8	2.6	3.3
Average total amount of NIH funding for a <i>candidate</i> in the five years preceding NDPA (in million dollars) (of the candidates with a minimum of one NIH award in the last five years)	\$3.1	\$3.0	\$3.3
<i>Awardee funding history</i>	FY 2004	FY 2005	FY 2006
Number of <i>awardees</i> with no prior NIH awards	2 (out of 9)	2 (out of 13)	2 (out of 13)
Total value of all NIH awards to NDPA <i>awardees</i> in the five years preceding NDPA application (excluding NDPA funds) (in million dollars)	\$30.6	\$28.4	\$30.5
Average number of unique awards per <i>awardee</i> (of the awardees with a minimum of one other NIH award in the last five years)	3.0	2.4	3.3
Average total amount of NIH funding for an <i>awardee</i> in the five years preceding NDPA (in million dollars) (of the awardees with a minimum of one other NIH award in the last five years)	\$4.4	\$2.6	\$2.8

Source: STPI Analysis of FY 2004 - FY 2006 NDPA Administrative Data and IMPAC II Data

The ability of the program to attract new researchers not within the “NIH fold” was further explored in the surveys of the candidates. Participants were asked if this was the first time they had ever applied for an NIH award. Overall, the majority of candidates had applied to the NIH at some point in their career. The program was more successful in attracting “new” candidates in FY 2004 and FY 2005 than in FY 2006 where only 4% of the candidates were first time applicants (Figure 3.10).

Figure 3.10:
FY 2004 - FY 2006 NDPA Survey Respondents who Previously Applied to the NIH*



Respondents were also asked about the distribution of their total funding in the past five years to assess other funding sources. Overall more than three-quarters of the respondents have held at least some amount of NIH funding in the past five years, and more than half received the majority of their funding (between 50-100%) from NIH (Table 3.4).

Hospitals, universities, or other non-profit organizations were the second most important source of funding for NDPA candidates, followed by foundations, other U.S. government sources, and for-profit companies. (Table 3.5)

Table 3.4:
Percentage of Funding in the Past Five Years from NIH for FY 2004 - FY 2006 NDPA Candidates

% NIH Funding	FY 2004	FY 2005	FY 2006	All Years
75-100%	34%	29%	30%	31%
50-74%	17%	24%	23%	21%
25-49%	17%	13%	12%	14%
1-24%	10%	10%	11%	10%
<i>Total with some amount of NIH funding</i>	78%	76%	77%	77%
0%	22%	15%	11%	16%
No response		9%	13%	7%
Total	411	420	330	1161

**Note: Since some respondents left the question blank, the total with some amount of NIH funding might be an underestimate. Also note that this question captures investigators who receive funding from NIH not only as principal investigators, which results in a more expansive definition of "receiving funding"*

Source: STPI Analysis of FY 2004 - FY 2006 NDPA Survey Data

Table 3.5:
Percent of FY 2004 - FY 2006 NDPA Candidates who have Received Some Funding from NIH and Other Sources in the Past Five Years

% Funding	FY 2004	FY 2005	FY 2006
National Institutes of Health	78%	76%	77%
Hospitals, Universities, or other Non-Profit Institutions	47%	47%	41%
Foundations	45%	43%	40%
Other US Government Sources	44%	39%	34%
For-profit companies	27%	26%	24%
Other	32%	11%	11%

**Note: Since some respondents left the question blank or responded to multiple possibilities, the columns do not sum to 100%*

Source: STPI Analysis of FY 2004 - FY 2006 NDPA Survey Data

3.5 Demographic Benchmarks

One way to determine whether certain characteristics are being well represented in the NDPA process is to compare NDPA to general funding trends at NIH. The process evaluation team researched NIH funding trends by characteristic in order to establish a baseline to which NDPA could be compared. As a recent National Academies' report^{xv} asserts, NIH trends by characteristic are not consistently tracked; but by using different studies and sources, STPI was able to construct rough estimates:

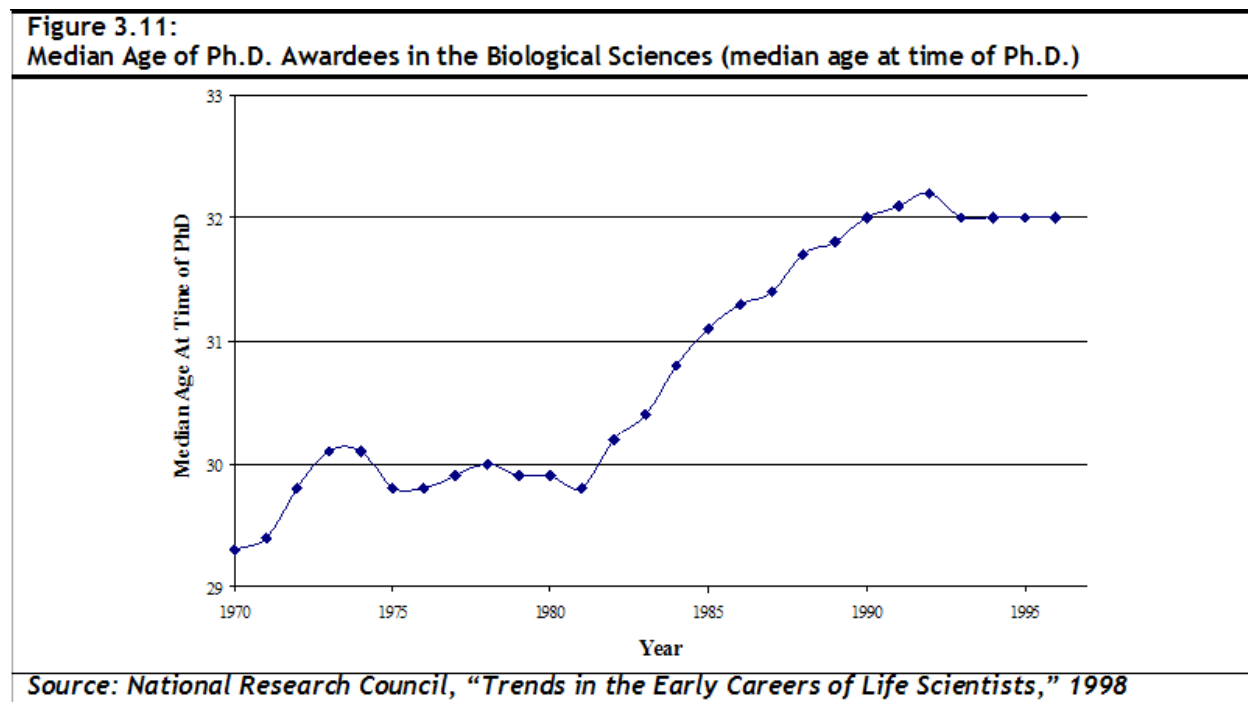
- Gender benchmarks were derived from the 2005 RAND report “*Gender Differences in Major Federal External Grant Programs*”^{xvi}
- Seniority benchmarks came from a study published in a 2002 Issue of *Science* magazine (*NIH Grantees: Where Have All the Young Ones Gone?*)^{xvii}
- Race/ethnicity benchmarks came from NIH data as reported in the *Journal of the National Medical Association* in August 2005^{xviii}

3.5.1 Benchmarking NDPA Gender Distribution

The RAND report “*Gender Differences in Major Federal External Grant Programs*” found that at the NIH, 28% of PIs applying for a grant in 2001-2003 were women and the percentage that received an award was 29%. NDPA trends by gender were very close to that NIH baseline: combined year totals indicated that 24% of all candidates and 29% of awardees were women.

3.5.2 Benchmarking NDPA Seniority Distribution

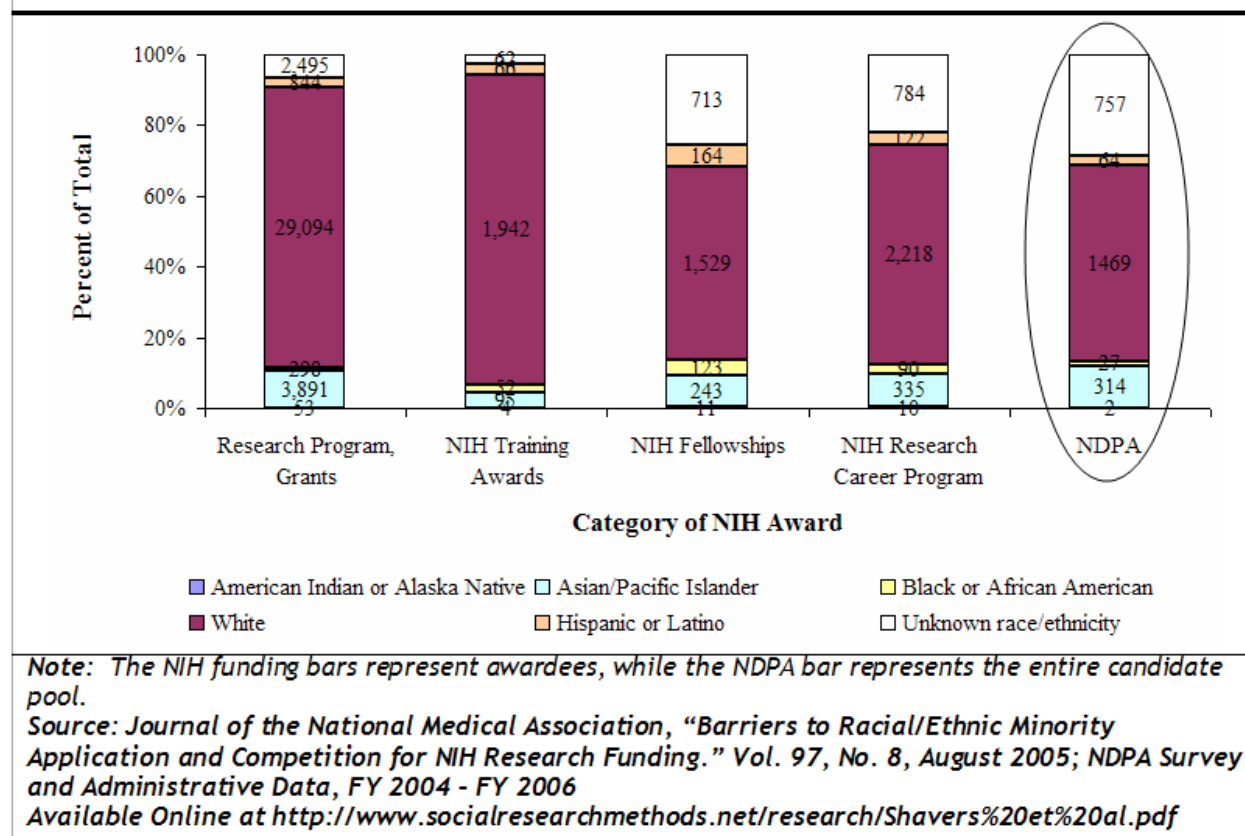
The percentage of NIH grants awarded to 35-and-under investigators has been declining during the past decades, reaching 4%, according to a National Research Council report in 1998 and a subsequent article in *Science* in 2002. It is suspected that the primary reason for the decline may be that the median age of Ph.D. awardees in the biological sciences has been steadily increasing during the past decades (Figure 3.11) and post-doctoral appointments are being held for longer periods of time. This declining trend among younger researchers was highlighted in a FASEB Journal study in 2000 that concluded that “nominations of physician-scientists age 45 or younger to honorary societies such as the American Society for Clinical Investigation have declined by almost 30% over the past decade, suggesting that the pool of talented young investigators is shrinking.”^{xix} NDPA is not insulated from this trend, as observed in the average seniority of candidates for FY 2004-2006 and the low percentage of younger researchers in the candidate and awardee pools as reported earlier in Section 3.1.2.



3.5.3 Benchmarking NDPA Race/Ethnicity Distribution

Funding trends for NIH by race and ethnicity were reported in an article that appeared in the *Journal of the National Medical Association* in August 2005.^{xx} The article presented IMPAC II data for FY 2002, which indicated that investigators from underrepresented minority groups (Black or African American, Hispanic or Latino, and Native American) comprised 3.2% of funded principal investigators on all Research Program Grants, 5.5% on NIH Training Grants and 10.7% on NIH Fellowships.^{xxi} Although complete race data were not readily available for NDPA, an estimate derived from Section 3.1.5 appears to be roughly consistent with these reported NIH trends for all Research Program Grants in 2002. (Figure 3.12)

Figure 3.12:
NIH Funding Awards for Selected Mechanisms, FY 2002 by Race/Ethnicity Matched against NDPA Candidate Pool



4. Characteristics of NDPA Reviewers (FY 2004 - FY 2006)

4.1 NIH Liaisons

A total of 53 individuals (NIH “liaisons”) were involved in the administrative review of candidates in FY 2004-2006 for a total of 73 counts of review (see Appendix table for an explanation of the review phases):

- 28 were involved in FY 2004
- 18 were involved in FY 2005 (Nine were new to the process, and nine had participated in FY 2004)
- 27 were involved in FY 2006 (16 were new to the process and eleven had participated in at least one previous year of review - three in FY 2005, three in FY 2004, and five in FY 2004 and 2005)
- In all years, there were more female than male liaisons (in total 58% of the liaisons were female and 42% were male)

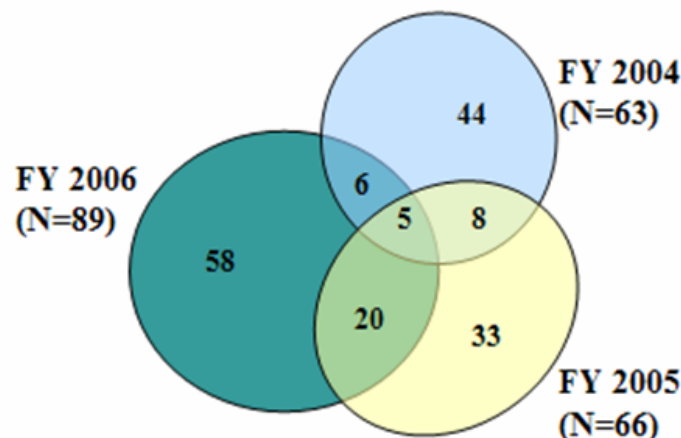
Overall, liaisons were drawn from 21 of the 27 NIH ICs and the Office of the Director. The FY 2004 cohort of liaisons was the most highly representative of the NIH ICs, while the FY 2005 cohort was the least representative.

4.2 External Evaluators

Given the critical role of evaluators in identifying pioneers, NDPA leadership made a concerted effort to attract experts from a diverse set of backgrounds to participate as external evaluators. A total of 174 individuals have participated in at least one year of review, though many have participated in multiple years for a total of 218 individual participant-years. Though the total number of candidates has declined over the years, more evaluators have been recruited in each subsequent year and repeat evaluators have become more common (Figure 4.1):

- 63 individuals were involved in FY 2004
- 66 individuals were involved in FY 2005
- 89 individuals were involved in FY 2006
- Five were involved in all three years of review, and 34 were involved in at least two years of review (8 in FY 2004 and FY 2005; 20 in FY 2005 and FY 2006; 6 in FY 2004 and FY 2006).

Figure 4.1:
Participation by External Evaluators in Multiple Years of Review



Source: STPI Analysis of FY 2004 - 2006 NDPA Evaluator Data

In FY 2004, the recruitment of evaluators was done in a shorter period of time, and the resulting evaluator pool was predominately white, male, and senior. In subsequent years, a targeted effort was made to attract a more diverse pool of evaluators - these efforts were successful in terms of attracting more women.

Overall, across all years, the evaluators were:

- predominately men (67% male, 33% female), though the proportion of women evaluators increased in FY 2005 and FY 2006 (Figure 4.2);
- primarily drawn from universities or university-affiliated medical institutes (58% university, 27% university-affiliated medical institutes, 8% private research institutes, 5% corporate, and 2% from government institutes) (data not shown);
- relatively senior (the majority obtained degrees more than 25 years ago), though the evaluators in FY 2005 and FY 2006 were slightly less senior than in FY 2004 (Figure 4.3);
- generally matched to the research area distribution of the candidate pool (Figure 4.4) and;
- predominantly white (84% White, 9% Asian, 4% Black or African American, and 3% Hispanic or Latino) (Figure 4.5).

Figure 4.2:
Gender Distribution of NDPA External Evaluators, FY 2004 - FY 2006

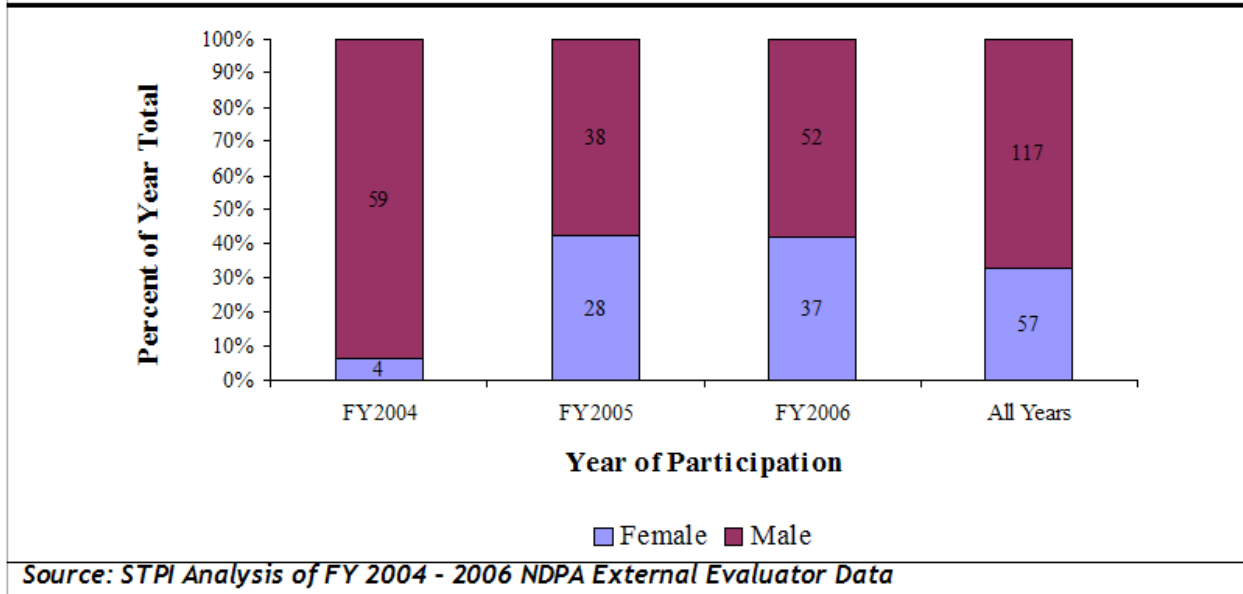


Figure 4.3:
Seniority of NDPA External Evaluators, FY 2004 - FY 2006

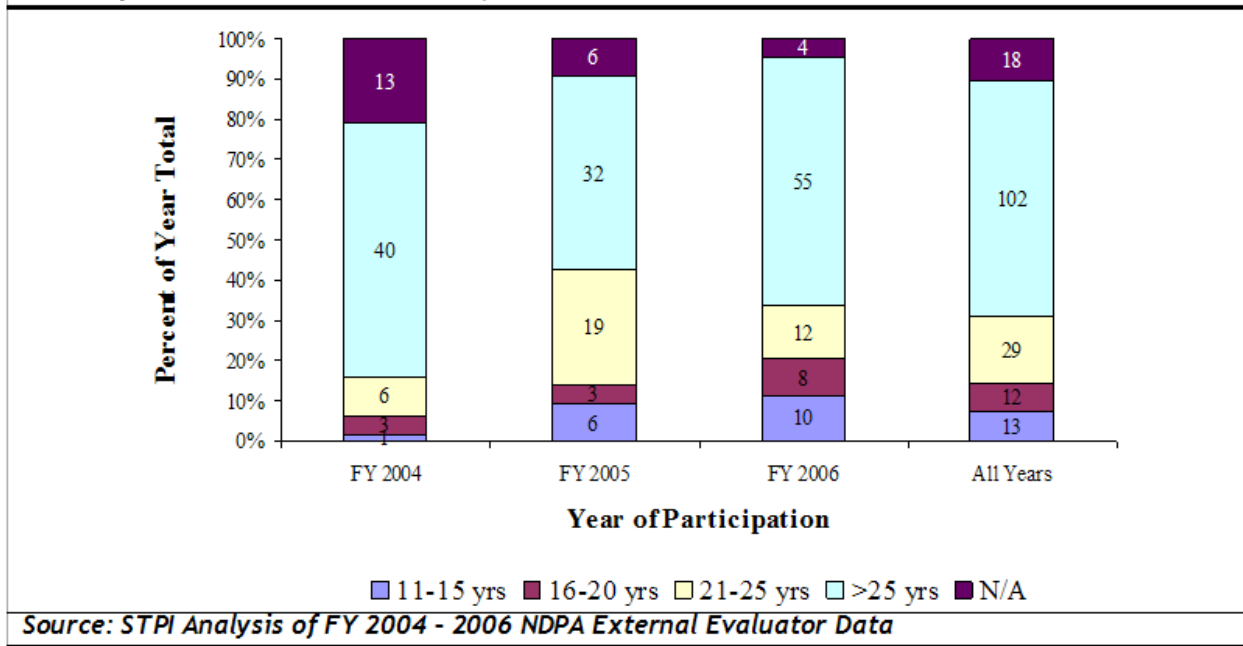
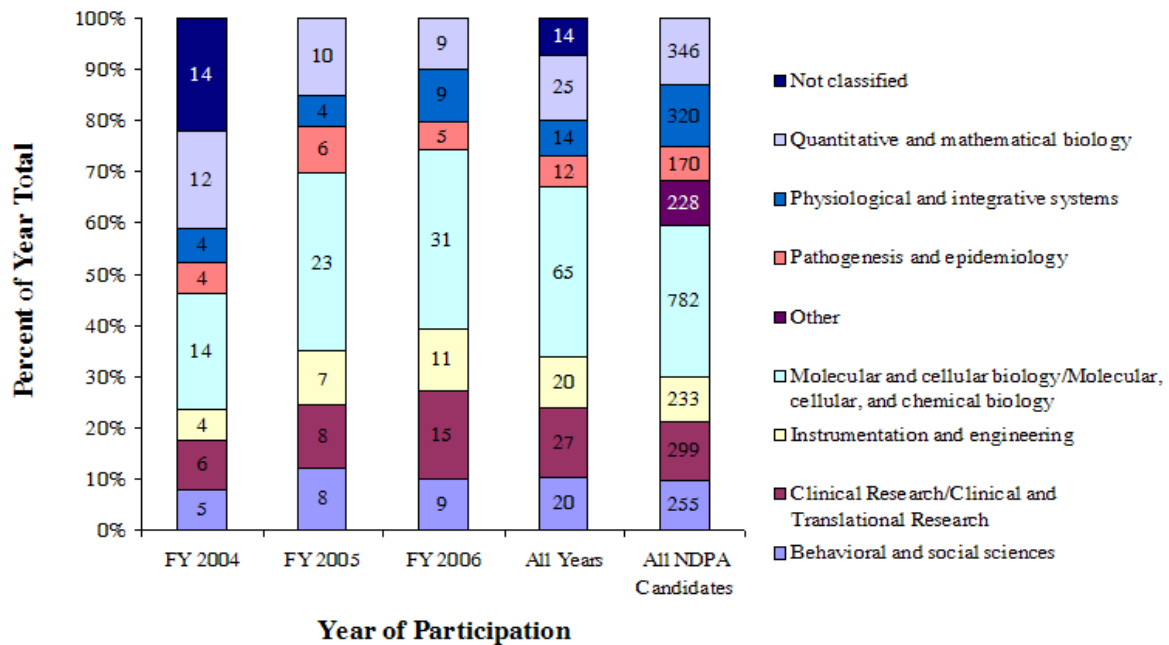


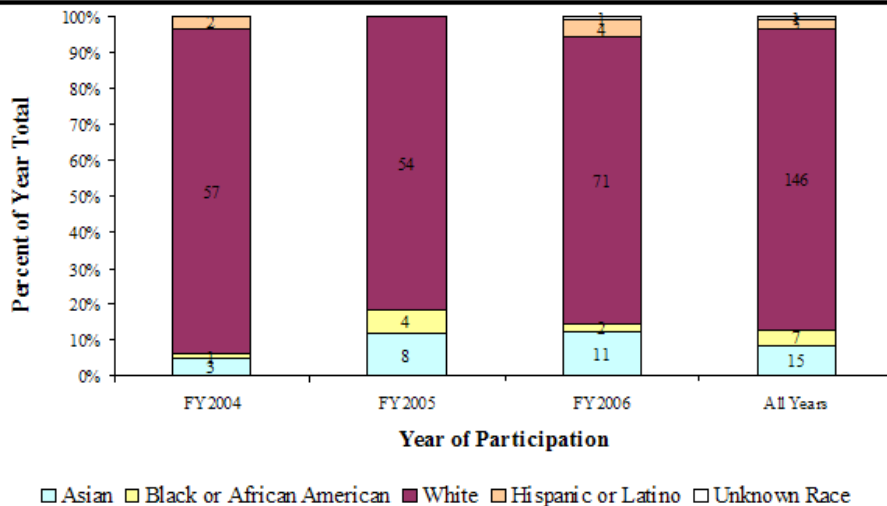
Figure 4.4:
Research Areas of NDPA External Evaluators, FY 2004 - FY 2006



Note: Some repeat evaluators were classified differently by research area in their various years of participation. The "Total" column, therefore, sums to 197 instead of 174 as some of the evaluators are counted multiple times. Additionally, the research categories were altered in FY 2006 - for the sake of simplicity, the "molecular and cellular biology" and the "molecular, cellular, and chemical biology" categories and the "clinical research" and "clinical and translational research" categories are grouped together.

Source: STPI Analysis of FY 2004 - 2006 NDPA External Evaluator Data

Figure 4.5:
Race of NDPA External Evaluators, FY 2004 - FY 2006



Source: STPI Analysis of FY 2004 - 2006 NDPA External Evaluator Data

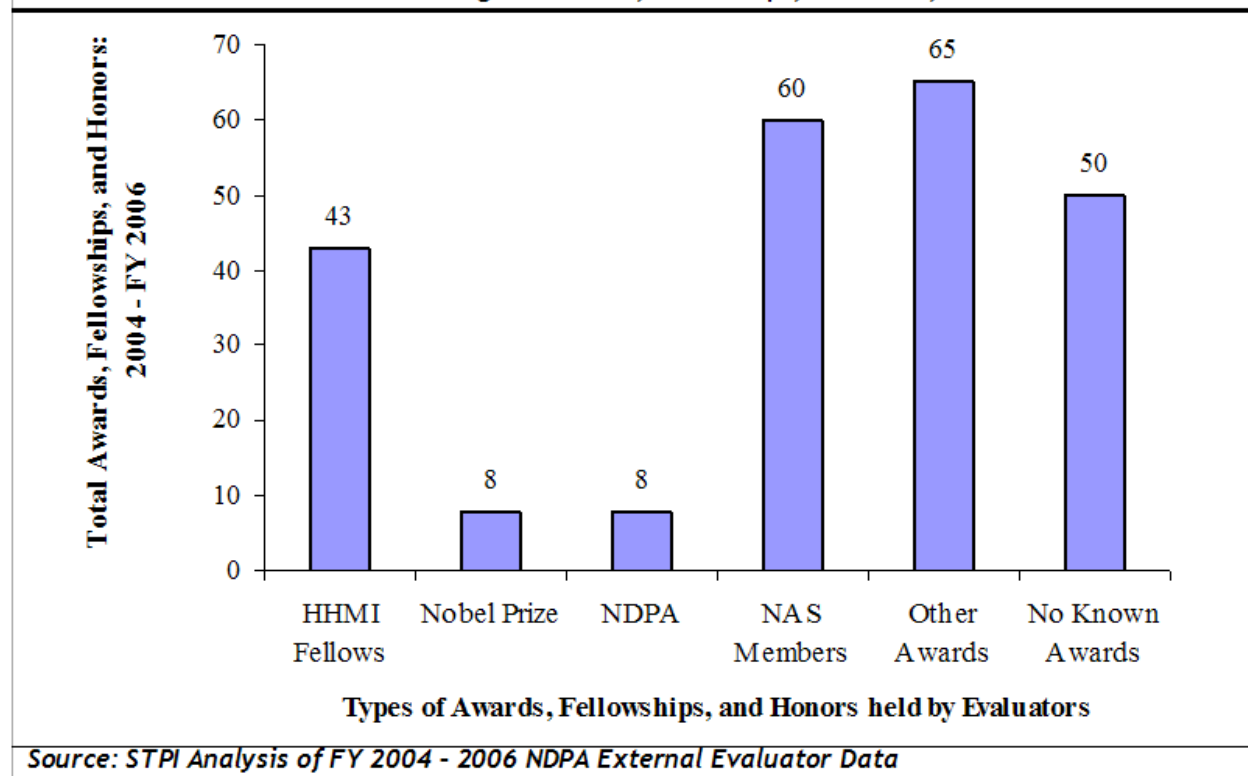
4.3 Accomplishments of the NDPA External Evaluators

By design, the evaluators recruited to participate in the NDPA selection process are well-known researchers and leaders in their fields. Though it is difficult to judge the ability of the evaluators to identify pioneering research, it is clear from researching curricula vitae and personal websites that the evaluators are an accomplished group. Though information could not be found for all of the evaluators, there are accomplishments worth noting (Figure 4.6):

- Overall, at least 127 of the 174 (73%) external evaluators have won major awards in their fields or have received prestigious fellowships or honors
- Sixty of the evaluators are members of the National Academy of Sciences (NAS)
- More than forty of the evaluators are, or have been Howard Hughes Medical Investigators (HHMI) at some point in their career
- Eight of the evaluators are Nobel Laureates, winning awards in Physiology or Medicine, Physics, and Chemistry
- Eight of them are NDPA awardees from previous years

The program has been able to recruit accomplished evaluators in all years. For example, the proportion of evaluators with prestigious awards, fellowships, or scholarships was 75% in FY 2004, 70% in FY 2005, and 72% in FY 2006 (data not shown).

Figure 4.6:
NDPA External Evaluators with Prestigious Awards, Fellowships, or Honors, FY 2004 - FY 2006



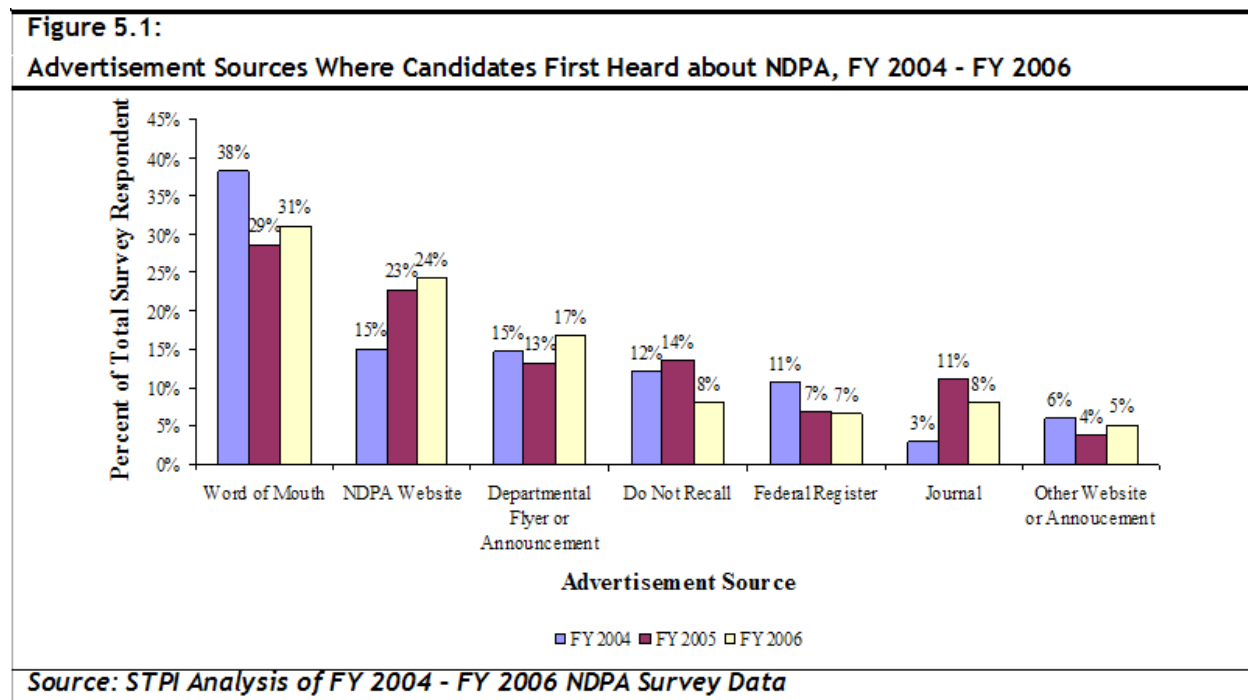
5. NDPA Process Description, FY 2004 - FY 2006

5.1 Phase Overview

The NDPA selection process has evolved since its inception in FY 2004 (as displayed in the Appendix table). The following sections summarize the key differences in this process across the years. More detailed information about the Phases is provided in the annual data reports which can be obtained from NIH upon request. Because not all phases apply in each year, in the discussion below, the external evaluator phases (Phase 2 and 3 in FY 2004 and FY 2005, and Phase 2 in FY 2006) have been grouped together.

5.1.1 Program Advertisement

In all years, information about NDPA was advertised broadly and disseminated via numerous sources, including journal articles, society announcements, newsletters, and the NDPA website. More than half of the candidates who responded to the survey in all years indicated that they *first* learned about the program by word-of-mouth or from the NDPA website. Overall, the Federal Register, journal advertisements, and other websites/announcements were the least relied upon sources of information by candidates; however, the journal advertisements were more relied upon in FY 2005 and FY 2006 than in FY 2004 (Figure 5.1).



As would be expected with a pilot, there were programmatic changes from year to year that may have influenced program advertisement and reception by the biomedical research community. These changes are particularly pronounced between the Program Announcement (PA) in the inception year, FY 2004, and the RFA in FY 2005.^{xxii} The Appendix table displays many of the changes of interest, though it is worth highlighting several key changes:

- In FY 2005 and FY 2006, a greater effort was made to draw in a more diverse pool of candidates. As stated in the FY 2005 and FY 2006 RFAs, “those at early to middle stages of their careers, women, and members of groups underrepresented in biomedical research are especially encouraged to apply.” Guidelines to reviewers in FY 2005 and FY 2006 similarly emphasized this priority.
- In FY 2004, the PA did not specifically define the term “pioneering,” though it was explicitly defined in the RFAs of subsequent years.
- The review criteria shifted between FY 2004 and FY 2005 from purely “people-based” criteria (Innovation/creativity; Intrinsic motivation/enthusiasm/intellectual energy; and Potential for or actual scientific leadership/evidence of, or potential for, effective communication/educator skills) to a mix of “people- and project-based” criteria (Scientific problem to be addressed; Investigator; and Suitability for NDPA mechanism).
- Though awardees in FY 2004 were expected to commit “the major portion of their effort to activities supported by NDPA,” in FY 2005 and FY 2006 the RFA specified that “Awardees are expected to commit the major portion (at least 51%) of their research effort to activities supported by the NDPA.”
- In FY 2004 candidates were given little instruction about what to include in their essays whereas in FY 2005 and FY 2006, candidates were required to specifically address the following questions:
 - “What is the scientific problem that will be addressed, and why is this important?”
 - “How will the new research direction differ from the individual's past or current work?”
 - “Why is the planned research uniquely suited to the stated goal of the NDPA program?”

5.1.2 Administrative Review (Nomination Phase/Phase 1)

In FY 2004, candidates could nominate themselves, or they could be nominated by a colleague. Analysis revealed no systematic differences in terms of demographics or success rates between the two groups of candidates (self vs. nominated by other). As a result of community and liaison feedback that the two-pronged nomination process resulted in disparate information on the candidates and had the potential to favor certain groups, only self-nominations were allowed in subsequent years.

Once nominated, all candidate materials were reviewed to ensure that the nomination packages were complete and that the candidates met program eligibility requirements. In FY 2004, NIH liaisons specifically were instructed not to consider existing grant support or career stage in their reviews and not to perform a scientific review. Being the first year of the pilot, there was much variability in the reasons for eliminating a nomination package - while the majority of the liaisons appeared to have followed the instructions and performed an administrative review, several conducted more of a scientific review than an administrative one. In FY 2005 and FY 2006, NIH liaisons were specifically instructed to conduct a more scientific review and eliminate nomination packages that were not sufficiently different from the candidates’ previous work, or not sufficiently suitable for NDPA (as will be discussed in Chapters 6 and 7, the term “suitability” was left to the interpretation of the reviewer). Phase 1 review in these later years was somewhat more extensive than the Phase 1 review in FY 2004. FY 2005 and FY 2006 liaisons were also more consistent in their reviews, and followed given instructions closely. In all years, NIH liaisons underwent training sessions to ensure that review criteria and purpose were well-understood and uniformly applied.

The role of the administrative review became less relevant in FY 2006 - thirty percent of the candidates were screened out during the administrative review phase in FY 2004, 32 percent in FY 2005 and only 13 percent in FY 2006. It appears that the community has understood the rules and scope of NDPA, and the initial screening may have become less important. As a result, the role of this phase going forward is uncertain.

5.1.3 External Evaluator Review (FY 2004 - FY 2005 Phases 2 and 3; FY 2006 Phase 2)

With the number of candidates consistently declining in each year, a decision was made in FY 2006 to reduce the number of review rounds. In FY 2004 and FY 2005, there were two rounds of external review. The goal of the first round was to further screen out individuals who did not fulfill the spirit of NDPA, albeit with a more scientific review rather than an administrative one - evaluators gave the application “yes/no” votes. Candidates who received at least two “yes” votes in FY 2004 and one “yes” vote in FY 2005 were then reviewed by a second round of external evaluators (Phase 3). In the second round, the evaluators scored the applications on each of the three criteria, gave overall scores, and designated their “top-4” choices. In FY 2006 candidates submitted all of their application materials up front and there was only a single round of external review. All FY 2006 candidates receiving at least one “yes” vote at the administrative review were rated by external evaluators (Phase 2) - Phase 2 in FY 2006 and Phase 3 in FY 2004 and FY 2005 are equivalent and are discussed in tandem.

Evaluators in FY 2004 were generally in agreement about the criteria and their operationalization (criteria listed in section 5.1.1 and in the Appendix table); however, they viewed the leadership criterion as the least relevant and their lowest priority. This criterion was therefore dropped in FY 2005. Evaluators in FY 2005 and FY 2006 were also divided on the importance of the three new criteria - some thought the idea was the most important factor to consider, others were seeking “pioneering” individuals, while others weighed all of the criteria equally. Additionally, FY 2005 and FY 2006 evaluators were split on how to use existing grant support and career stages of the applicants. Nearly half of the evaluators interviewed gave existing grant support and applicant career stage “no consideration,” while half gave them either “significant” or “some” consideration. Several evaluators expressed their view that when looking for high risk, or creative research, it did not really matter whether what was written in the application corresponded to the NDPA criteria - they simply “knew it when they saw it.”

Evaluators also were ambivalent about the requirement that awardees commit to devoting 51% of their time to NDPA. While many evaluators considered the 51% factor, they suggested that this was the least important criterion in the review. Several evaluators indicated that most creative individuals do not have 51% free time to devote to any given project, nor should this requirement necessarily be enforced. This led to the criterion being applied inconsistently throughout the process. In some cases, candidates were excluded as late as the final award phase because NDPA staff perceived them to have existing support that would preclude them from being able to spend the needed time on NDPA.

The spread in number of applications reviewed by FY 2004 external evaluators was much wider (11 to 43 applications per evaluator) than in FY 2005 or FY 2006. This created statistical problems when looking at the total applicant pool - for instance, the individual who reviewed 11 applications still designated his or her “top-4” choices; however, these “top-4” may not have been the same caliber of the “top-4” of another evaluator who reviewed 43 applications (Table 5.1). This discrepancy was eliminated in future years by ensuring that all evaluators reviewed approximately the same number of applications. In FY 2004, a

statistician standardized the scores of all evaluators; this correction was deemed unnecessary in subsequent years.

Table 5.1
Summary of External Evaluator Scoring (Phase 3 in FY 2004 and FY 2005; Phase 2 in FY 2006)

Year of Participation	Reviews per Evaluator	Average Overall Score for all Applicants	Range of Average Overall Scores by Evaluator
FY 2004	11 to 43	4.37 (\pm 0.7) on 7pt scale	2.8 to 6.3 on 7 pt scale
FY 2005	21 to 25	3.06 (\pm 0.44) on 5 pt scale	1.9 to 4.2 on 5 pt scale
FY 2006	14 to 17	3.21 (\pm 0.18) on 5 pt scale	2.3 to 4.5 on 5 pt scale

Source: STPI Analysis of FY 2004 - FY 2006 NDPA External Evaluator Scoring Trends

Though the evaluators in FY 2004 indicated that there was inconsistency in understanding of the process and that the information provided was inadequate to judge the candidates, the FY 2005 and FY 2006 evaluators were more satisfied with the process and the materials provided. They were, however, split on the effectiveness of the 5-point scale and “top-4” vote system. Approximately half of the evaluators thought the system was effective; and the others thought that improvements are still needed. One evaluator indicated that it was a mistake to switch from the 7-point to the 5-point scale in FY 2005. This individual expressed, “I have no idea what difference it made switching from the 7-pt to 5-pt scale system, but I can speculate. Anything looking more like the typical NIH system gets people back into the NIH rut - I think this was a mistake.” Others also suggested that it is very difficult to compare the scores of one evaluator to that of another. As stated by one evaluator, by nature of the distribution, “one person’s 4th choice might be better than another evaluator’s top choice. This could be a little tricky when looking across multiple scores.”

In the first two years of implementation, a larger number of phases resulted in greater path dependency in the process. Though evaluators were trained on the criteria and the purpose of the various phases, not all of them applied the criteria in the same manner. Fewer phases in later years implied fewer counts of review. In FY 2004 there were *six* counts of external review per application (three external evaluators in Phase 2; three external evaluators in Phase 3), in FY 2005 there were *five* counts of external review per application (three external evaluators in Phase 2; two external evaluators in Phase 3), in FY 2006 there were *three* counts of external review per application (three external evaluators in Phase 2).

External evaluators generally were satisfied with the candidate materials although there were some concerns about the letters of reference. One evaluator who had participated in multiple years summarized his views as follows: “The materials were all relevant except for the letters of reference...these were really useless. Basically people had their buddies write them glowing letters - I didn’t end up giving much weight to these recommendations. At this stage in an investigator’s career, they shouldn’t need letters of reference - they should be proven based on their ideas, past performance, and future potential.” On the other hand, some evaluators liked these letters, though they wished for more standardization. One evaluator stated that “the letters of reference became very important at the interview phase - it would have been nice to have had a bit more standardization in the types of information provided.” Some evaluators recommended keeping the letters with more specific instructions, where referees are asked to address each criterion separately and specifically in their letter. The essays were viewed both by the applicants and evaluators as the most useful components to display “pioneering” qualities.

Overall, in terms of the potential “pioneeringness” of ideas attracted to the program, candidates and evaluators held somewhat differing perspectives. As might be expected, the

majority of candidates who responded to the survey in all years indicated that they would be unlikely to receive funding for their proposed NDPA idea through a traditional NIH grant mechanism. The evaluators did not concur on this point and suggested that they found the applicant pool to be somewhat of a “mixed-bag” - there were some very strong candidates, and others who clearly should have been applying for an R01 or other type of award. Evaluators, however, did indicate that the NDPA process itself was very different from a traditional study section, and conducive to allowing investigators to submit more innovative and creative applications.

5.1.4 Selection of Interview Candidates

After the review of applications, applicants deemed to be the most “pioneering” (determined in part based on evaluator scores and “top-4” votes, and in part on NDPA staff judgment) were invited to participate in the interview round. Of the total 928 applicants reviewed over the years, 410 (44%) received at least one “top-4” vote from the external evaluators.^{xxiii} Though not all individuals with “top-4” votes were asked to interview, “top-4” votes were good predictors of success (Table 5.2). This observation will be discussed more extensively in Section 5.2. However, the selection of the interviewees from the applicant pool cannot be determined based on average overall scores and “top-4” designations alone - there was some subjectivity involved in the selection of the interviewees in each year. There were individuals who were not asked to interview who had higher scores and more “top-4” designations than some of the interviewees (and awardees); similarly, there were some individuals asked to interview with low scores and only one “top-4” designation (Figure 5.2).

Over the years, the selection of interviewees has become somewhat more transparent; there was less variability in the scores and “top-4 votes” of those invited to interview in FY 2006 than in previous years. While participants did not have access to this data, some evaluators and candidates alike have indicated that the selection of interviewees is somewhat of a “black box.” Candidates in all years have been especially concerned because they get no feedback on their scores or reasons as to why they were not asked to interview.

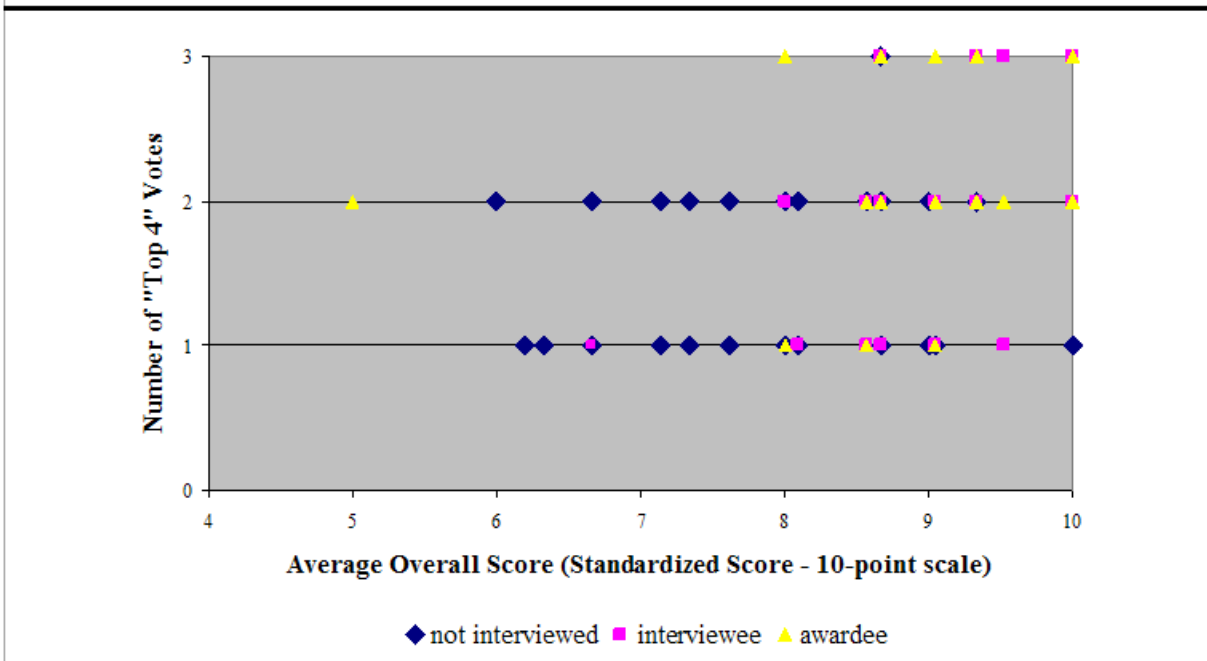
Table 5.2
Summary of External Evaluator Scoring (Phase 3 in FY 2004 and FY 2005; Phase 2 in FY 2006)

Year of Participation	FY 2004	FY 2005	FY 2006
Applicants with one “top-4” vote	61 (8 interviewees; 2 awardees)	101 (4 interviewees; 2 awardees)	120 (0 interviewees; 0 awardees)
Applicants with two “top-4” votes	18 (11 interviewees; 6 awardees)	18 (13 interviewees; 9 awardees)	69 (12 interviewees; 4 awardees)
Applicants with three “top-4” votes	3 (3 interviewees; 1 awardee)	3 (3 interviewees; 2 awardees)	17 (13 interviewees; 9 awardees)
Percent of Applicants with at Least 1 “top-4” Vote*	33%	43%	51%

**Note: The total number of external evaluators increased in each year, therefore there were more “top-4” votes given in FY 2006 than in FY 2004 or FY 2005*

Source: STPI Analysis of FY 2004 - FY 2006 NDPA External Evaluator Scoring Trends

Figure 5.2:
FY 2004 - FY 2006 NDPA Interviewee Selection - Applicants with Standardized Overall Scores Higher than 6 and a Minimum of 1 “Top-4” Vote (n=340)*



**Note: It should be noted that each point on the graph likely represents multiple candidates, as there were several candidates with the same average overall score and number of “top-4” votes. Additionally, because the scoring scale was different in FY 2004 (7-point scale) from in FY 2005 and FY 2006 (5-point scale), overall scores were standardized based on a 10-point scale. Source: STPI Analysis of FY 2004 - 2006 NDPA External Evaluator Scoring Trends*

5.1.5 Panelist Scoring and Selection of Awardees

The interview phase of NDPA is viewed to be the most distinctive aspect of the program and has remained largely unchanged over the years. Since FY 2004 a total of 67 applicants have been invited to the NIH to present their ideas to the external expert review panel; 35 (52%) of these individuals went on to win NDPA.

Panelists listened to the interviewee presentations, then were given time to ask questions and to discuss each candidate. In all years, interviewees were placed into “tiers” based on the recommendations of the panelists - top, middle, and bottom tiers. Candidates in the top tier were absolutely recommended for funding, those in the middle were suggested for funding if money was available, and those in the bottom tier were not recommended for funding. Recommendations were given to the NIH Director and the Advisory Committee to the Director (ACD) for final decisions to be made.

There is very little documentation for the final phase of the NDPA selection process, though it is known that NIH Director and the Advisory Committee to the Director (ACD), with input from the NDPA leadership made final decisions on the award winners in all years. In FY 2004, all funding was provided through the NIH Roadmap initiative. In FY 2005 and FY 2006, additional funds were secured through other NIH Institutes and Centers (ICs) in order to increase the total number of awards given. Before final decisions were made, the co-chairs of the NDPA Oversight Committee discussed all candidates with IC Directors who were identified earlier to be interested in supporting NDPA awardees. Funding for additional awardees was secured.

Because some of the ICs specified research they wanted to see funded, several individuals from the bottom tier were selected as awardees (particularly in FY 2005). As highlighted in Table 5.3, overall, 20 of 26 (77%) candidates in the top tier received an NDPA award, 12 of the 18 (67%) middle tier candidates were selected, and three of the 23 (13%) bottom tier candidates were given awards. Some FY 2005 and FY 2006 panelists interviewed expressed disappointment that “some of the people they strongly recommended should not get funding were pulled in by other institutes. This defeats the purpose of the award.”

Table 5.3
NDPA Awardees by Panel Designation, FY 2004 - FY 2006

Panel Designation	FY 2004		FY 2005		FY 2006		All Years	
	Funded	Not Funded	Funded	Not Funded	Funded	Not Funded	Funded	Not Funded
Top Tier	6	4	6	1	8	1	20	6
Middle Tier	3	2	5	1	4	3	12	6
Bottom Tier	0	7	2	5	1	8	3	20

Source: Internal NDPA Documents, FY 2004 - FY 2006

Of the panelists interviewed, all indicated that they had enough time and materials to prepare for the panel session. The panelists stated that the system to rank interviewees was efficient and successful. All believed that the interview duration was appropriate - it was enough time to get a feel for the candidates, to ask questions and to reach a consensus. Panelists enjoyed the interview process, though some gave specific suggestions for future years:

- Several panelists requested that briefing books be provided in future years - these books should have extensive material on each of the interviewees.
- Though the interviewees were all given the same information before the interview, panelists indicated that there was a tremendous amount of heterogeneity in the presentations. One panelist recommended that in the future there should be some sort of template for the presentation so that the panelists get the same level of detail from each individual.

Despite some dissatisfaction with the final selection of awardees in FY 2005 and FY 2006, panelists in general were very happy and enthusiastic about the final round of the NDPA review - they expressed that the process is exciting and highly effective and that in general, interviewees and awardees were qualified candidates.

5.2 Candidate Perception of the NDPA Process

Over the years, candidates have been mixed in their views about the effectiveness of the NDPA process. In particular, they have expressed concerns that NDPA criteria are applied inconsistently and the program might be biased. The surveys and interviews in FY 2004 - FY 2006 revealed some candidates suspect that reviewers were favoring certain characteristics over others, and that the process was “controlled” in some way. The absence of feedback to candidates heightened this perception and added to the suspicion that reviewers had something to hide, such as bias for or against specific groups or disciplines. Survey responses from candidates who did not plan to reapply to NDPA indicated that there was a perception of bias in the process.

In FY 2004, based on the absence of women or minority awardees, there was a perception of bias in the program (although the FY 2004 process evaluation report found that the results were within the range of statistical possibility and that there was no explicit evidence of

“bias” that emerged in the interviews with liaisons and evaluators). In FY 2005 and FY 2006, some candidates still suggested that the process was inherently biased against minorities despite the explicit call for them to apply in the program solicitation and the directions to reviewers to consider gender, ethnicity, and seniority in scoring decisions; others believed it was biased in favor of women and minorities. For example, one applicant indicated, “Had the criteria been applied, they would have been good. But judging by the list of awardees, a completely different set of criteria was in fact applied, with political correctness high on the list.” Similarly, another stated, “The selection process is obviously highly controlled by ‘political correctness’. Being minority or female is statistically highly favored in the selection process. This of course has nothing to do with the stated selection process, but rather reflects the affirmative action mentality.”

Some candidates expressed that they were too junior and did not have the desired record to reapply, while others felt that they were too senior and were automatically eliminated from the competition. Along these lines, many felt that the same, well-funded people were still winning, reinforcing the “old boys network.” As expressed by one respondent, “A pioneer is someone who settles a new land; sometimes these new territories have pre-existing indigenous peoples already living there. The Pioneer Award seems to favor the indigenous types.”

With respect to the interview phase, in general, finalist opinions were mixed about whether interviewers adequately understood their ideas and whether they had a fair chance (Table 5.4). The combined survey data indicate that awardees and interviewees who did not receive an NDPA held very different opinions about the interview process. As would be expected, non-awardee interviewees’ opinions contrast sharply with those of the awardees, most of who reported satisfaction with the interviewer comprehension of the proposed research. Non-awardees who were interviewed had mixed opinions on the clarity of the invitation instructions, indicated that the panelists only somewhat understood their ideas or did not understand them at all, and the majority suggested that the interview duration was too short.

Table 5.4:

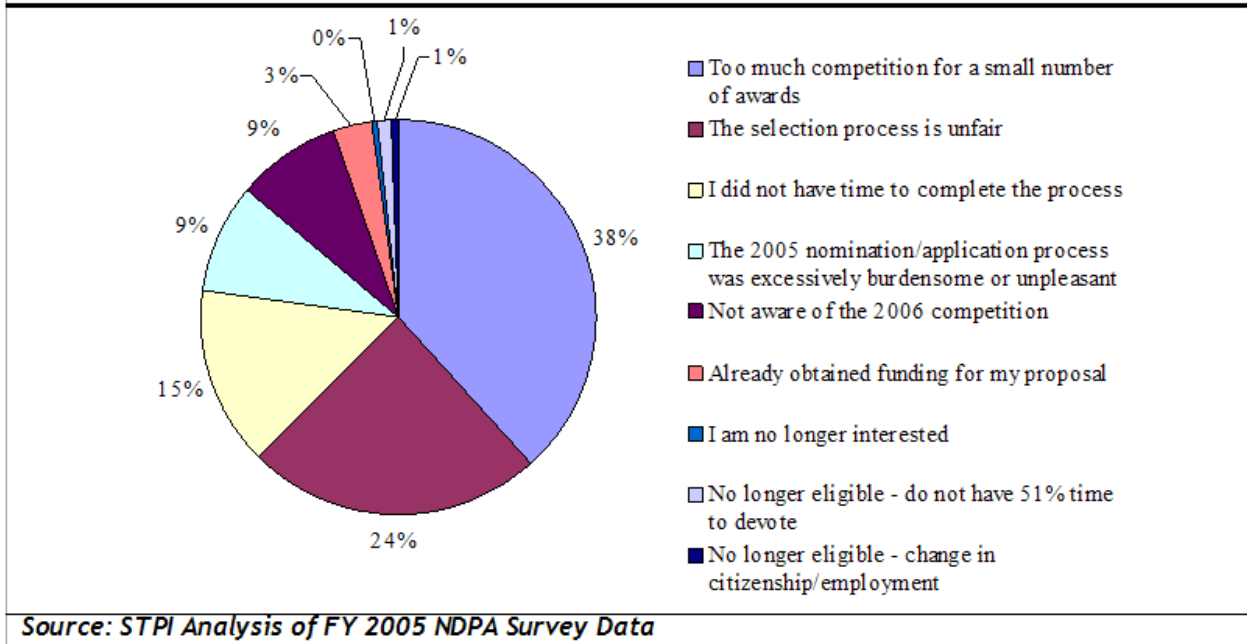
FY 2004 - FY 2006 Combined Survey Data: *“Do you agree that the interviewers adequately understood your ideas and gave you a fair chance to convey your ideas?”*

Response	Interviewees/Non-awardees				Awardees			
	FY2004	FY2005	FY2006	Total	FY2004	FY2005	FY2006	Total
Agree	4	0	2	6	9	8	11	28
Disagree	6	2	8	16	0	3	1	4
Total	10	2	10	22	9	11	12	32

Source: STPI Analysis of FY 2004-2006 NDPA Survey Data

Because the FY 2005 survey was fielded after the FY 2006 administrative data were obtained, a unique opportunity was presented to ask individuals why they chose not to reapply in FY 2006. Candidates who responded to the survey who did not reapply for NDPA in FY 2006 primarily indicated that they did not reapply because there is too much competition for a small number of awards (43%), or because they thought the process was somehow unfair (27%) (Figure 5.3).

Figure 5.3:
Reasons FY 2005 NDPA Candidates did not Reapply in FY 2006 (n=409 respondents)



The primary complaint about the program evident from coded free responses was that without feedback given to the candidates, they had no guidance on what was lacking in their application, and how to improve in future years. As explained by one respondent, “In the absence of feedback, it is impossible (or an ineffective use of my efforts) to know how to improve upon my prior application. As such, I can only assume that I and this project were not of interest to the program.” Another said, “I was discouraged by the complete lack of feedback for the process. Some evaluative information, even if it had been cursory, would have encouraged me to try again. Was the idea too ambitious? Was the idea uninteresting to the reviewers? Did they feel that the system or techniques I was proposing were inappropriate? Without this feedback, I felt it was unlikely that I would succeed in creating a better proposal. I was also hesitant to trouble my colleagues to write letters of recommendation on my behalf for an application that was unlikely to be successful.”

Others claimed that the chances of success were so low that reapplication was not worth their time. As mentioned by one respondent, “The Pioneer Award is a wonderful idea in principle. However, these decisions are made behind closed doors from the perspective of the submitting scientist. Therefore, your odds of being funded are slim to none, and when you are not funded, you have no idea of why. The process left me with the feeling that the NDPA is the science equivalent of buying a lottery ticket...somebody will win, but it’s purely based on the luck of the draw.”

Some believed that the stated review criteria were not adequately operationalized, resulting in a mis-match between stated program goals and actual outcomes. This led to a perceived lack of transparency in the overall process and discouraged candidates from reapplying. One respondent wrote that “it appears to me that the criteria for awarding the NIH Pioneer award contradict the criteria stated. This unfair evaluation process leaves a bad taste in one’s mouth. I am really concerned that the NIH Pioneer award selection process is not consistent or true to its objectives but rather driven by political considerations that prevail in a study section environment.”

A substantial number of respondents did not reapply simply because they did not want to trouble their colleagues for yet another letter of reference. Others believed that the program was not suited for their field or research (clinical researchers were especially concerned about the lack of representation in the clinical and translational fields).

The candidates' perceptions about bias prompted the STPI evaluation team to consider a statistical analysis of available data to explore the possibility of bias in the program. This analysis is discussed below.

5.3 Modeling Determinants of Candidate Success for Combined Years

The objective of the statistical analysis was to examine whether a candidate's characteristics and scores had an effect on the selection of interviewees. The interview phase became the focus of the analysis, as opposed to the awardee final selection, because (1) nominal and ordinal data were readily available for this key selection phase, (2) the final award phase itself does not have empirical data, and (3) the odds of winning an award were much higher for candidates once he/she advance to the interview round (being selected for the interview is a decisive factor). With three years of NDPA competitions, the combined data allowed for the building of statistical tests to understand the relationships between variables and their importance to the outcome of the selection for the interview.

The primary methodology used to examine the relationships between the variables was to create an explanatory model using multivariate regression. Because the outcome variable was dichotomous, logistic regression was selected as the preferred procedure for the analysis. In addition to the full model, a series of secondary models were built to understand specific relationships. The model was also applied to the individual years in an effort to test robustness; due to data limitations, in some instances, univariate statistical tests were used to confirm inclusion of the variables. Finally, predictive versus actual outcomes for each applicant were compared to test the full model given various thresholds. Table 5.5 summarizes the various statistical analyses that were performed; detailed descriptions of the methodology, statistical tests, and select outputs are available upon request.

Table 5.5

Summary of statistical analyses performed to measure the effect candidate characteristics and scores had on the selection of interviewees for FY2004 - FY2006

Statistical Test	Purpose	Key Variables	Result
Chi-square Tests for combined years	Tests of independence of explanatory variables vs. proceed to interview	Proceed to interview and explanatory variables	Seniority was significant
Logistic Regression using demographic variables only	Analyze relationship between proceed to interview with demographic characteristics	Response: Proceed to Interview Explanatory Variables: All Demographic Characteristics	Seniority was significant
Logistic Regression using all variables	Analyze relationship between proceed to interview with demographic and scoring characteristics	Response: Proceed to Interview Explanatory Variables: All Demographic Characteristics and scores	Result: Seniority, Overall Score and Top-4 were significant
Logistic Regression and Chi-square Tests for individual years	To validate the selected model variables on the individual years.	Response: Proceed to Interview Explanatory Variables: All demographic characteristics and scores	Result: Model variables were significant for each individual year
Logistic Regression using scores to predict Top-4	Understand relationship between scores and Top-4	Scores, Overall score, Top-4	Result: Overall score was significant
Chi-square Tests for individual years	Validation of inclusion of variables in model to check for robustness of the multi-year regression modeling	Proceed to interview, Overall score, Top-4, seniority	Result: For each year, the explanatory variables from the model were significant on all individual years
Classification Analysis	Compare the predicted output of the model vs. the actual output of the process	Predicted proceed to interview vs. actual proceed to interview	Result: Model is effective at weeding out applicants who are not qualified, and moderately effective at identifying qualified applicants

Source: STPI Analysis of FY 2004 - FY 2006 NDPA Data

The statistical analyses resulted in three important findings:

- *“Top-4” and “average overall score” were strong predictors of whether or not a candidate proceeded to the interview phase.* The results of a series of stepwise logistic regressions indicated that top-4 and average overall score were predictive of the outcome; additionally, the average overall score was significant for each individual year. When the scoring was examined and the model was tested further, it was found that the average overall score and top-4 were strongly related.
- *With one important exception (seniority), statistical differences by gender, race, research area, affiliation, degree, and NIH funding were not found.* When testing

for all characteristics, the analyses indicated that seniority did have a relationship with a candidate's ability to proceed to the interview round (specifically - more experienced applicants fared slightly worse than their less experienced counterparts), and the effect of that relationship appears to have increased in significance for later years. It is difficult to conclude from the statistics alone why this is significant, but it is known that in FY 2005 and FY 2006, NDPA has placed a special emphasis on younger biomedical researchers. Apart from seniority, all other characteristics were not significant in predicting whether or not a candidate was selected for an interview.

- *Being a repeat applicant was not a predictor of whether that candidate would be selected for the interview phase.* Several tests were performed to examine the relationship of repeat applicants with the outcome; these tests indicated that this characteristic was not significant in FY 2005 or in FY 2006.

6. Overall Assessment

6.1 Program Design

6.1.1 Program Structure and Evolution

In its first three years of existence, the NDPA program adapted to incorporate many of the lessons learned from previous years of implementation. Review criteria and program structure (application materials and outreach) were adjusted to promote consistency among the reviewers' assessment and scoring.² The Appendix table highlights the changes across the years in terms of program emphasis, selection process, selection criteria, application materials, and the research areas of candidates.

The FY 2005 RFA marked a departure from the original program announcement. In addition to the criteria changes (discussed in Section 6.1.2), the RFA specifically used the term “pioneering” to “describe highly innovative approaches that have the potential to produce an unusually high impact.” Such a distinction was not made in FY 2004. Though subtle, the FY 2005 RFA (and the FY 2006 RFA in the subsequent year) began to emphasize the desire to identify “pioneering” people *and* projects as opposed to simply identifying promising individuals as was the emphasis in FY 2004. The FY 2005 RFA outlined a more explicit structure for the application materials (e.g., the 51% requirement, specific points to address in the application) in effort to make the submissions more comparable for the reviewers. Additionally, only self-nominees were invited to apply.

In FY 2006, the review and selection process was even more streamlined than in FY 2004 or FY 2005. Changes included the removal of the second round of external review, the requirement that all application materials be submitted up-front, and a more structured format for applicants to follow. Additionally, about half of the evaluators in FY 2006 had participated in previous years and therefore were familiar with the process. Overall, evaluators and liaisons alike were pleased with the changes made in FY 2005 and FY 2006. In FY 2005 the switch to self-nominations only was viewed as advantageous - this adjustment served to level the playing field and also to increase the reviewers' ability to compare candidates. Evaluators and liaisons were also pleased with the more structured application materials.

6.1.2 Selection Criteria

FY 2005 marked a shift in the NDPA review criteria from all “people-based” criteria, to a mix of people- and “project-based” criteria. The shift was intended to make the criteria easier to operationalize - some of the criteria from year one (e.g., leadership) were more subjective than others, difficult to apply consistently, and not viewed by all reviewers as a marker of a “pioneer.” All reviewers were trained to implement the new criteria and care was taken to ensure that everyone understood the meaning and proper application of the criteria.

² As mentioned in Chapter 1, the term “reviewers” refers to all individuals who participated in some phase of the review - both liaisons and external evaluators; “liaison” refers to the NIH representatives who participated in the administrative portion of the review; and “evaluators” refers to external experts who participated in one or more rounds of the scientific review. “Panelists” refers to evaluators involved in the interview process.

Despite the effort to clarify the selection criteria for candidates, the reviewers in FY 2005 and FY 2006 were clearly divided in their methods for deciding whether or not an application was competitive. Specifically, reviewers were unsure if they were to look primarily for a creative, “pioneering” individual or an innovative project. Several evaluators, NIH liaisons, and candidates noted as well that the program appears to be shifting towards a “project-based” award as opposed to the “people-based” award that was originally conceived. As stated by one evaluator, “the program had flaws in FY 2004, but I think the program is creeping away from what it was originally intentioned to do.” The change in selection criteria likely resulted in this split perception of the program.

The majority of the FY 2005 and FY 2006 reviewers were satisfied with the general effectiveness and relevance of the three new criteria specified for identifying a pioneer. Overall, evaluators tended to rely on a mix of the criteria, though they indicated that the 51% time requirement was difficult to judge. Some were not even sure that it was appropriate - they suggested that many pioneers are well-funded researchers, and a strict application of the constraint may exclude potential awardees. While many evaluators considered whether or not the investigator was over-committed, they suggested this was the least important criterion in the review. As a result, in FY 2005 and FY 2006, this criterion was applied inconsistently and at different phases of the process. Evaluators also expressed that the suitability for the NDPA mechanism criterion was a “Catch-22.” They indicated that “truly great” people with ideas suitable for NDPA are likely funded already; however, they were asked to seek individuals who would not be funded under traditional NIH mechanisms.

Generally, candidates who responded to the survey agreed that the criteria used for selecting awardees were adequate and appropriate to choose scientists of exceptional creativity; however, there was suspicion about the application of the criteria.

6.2 Program Implementation

6.2.1 Adequacy of Information

As the program has matured over the past three years, evaluators and candidates alike have been increasingly satisfied with the *application format and materials submitted*. The survey data for combined years indicated that, on-the-whole, candidates also increasingly agree that they were given adequate opportunity to display their qualifications.

The essays were viewed both by the applicants and evaluators as the most useful components to display “pioneering” qualities. Some applicants indicated that the reference letters were unnecessary and burdensome. Having to request letters from colleagues and mentors was cited as one reason they would not apply in future years. Reviewers questioned their utility as well, and some claim to have ignored them altogether during the review.

Panelist feedback on the interview process has been very positive in all years. They were generally very happy and enthusiastic about the interview round - they expressed that the process is exciting and highly effective, the materials and interview duration were appropriate, and the interviewees were of high quality. Candidate feedback on the interview process, however, has been mixed for all three competitions. As might be expected, awardees reported satisfaction with the process, while non-awardees indicated that the panelists only somewhat understood their ideas, or did not understand them at all; the majority suggested that the interview duration was too short.

6.2.2 Scoring Trends

NDPA's use of "top-4" picks and scores to select finalists and awardees has increased in consistency during these past three years. Whereas in FY 2004 and FY 2005, many of the awardees received "mixed" scores (e.g., one "yes" and one "no" score at the first phase of external review) or at least one low score in the second round of external review - in FY 2006, all 13 awardees received at least two "top-4" votes and an average overall score above 4.5 (out of a maximum of 5). However, in all years, top scores alone did not fully determine the selection of interviewees. The discretion of the Oversight Committee co-chairs allowed them to select some interviewees whose average overall scores were below those of non-interviewees.

6.2.3 Transparency of Process

Overall, the NDPA selection process may not have been entirely transparent in any given year. The most common criticism for the FY 2004 - 2006 NDPA competitions was the lack of feedback. It is standard for NIH programs to provide feedback, and NDPA candidates stated that it is difficult to improve their proposal without reviewer comments. This lack of transparency contributes to the perception by some candidates that reviewers have something to hide, such as bias for or against certain groups. As discussed in Chapter 5, candidates who do not plan to reapply to NDPA believe that certain characteristics that they do not possess were favored. Statistical analyses showed that most characteristics (with the exception of seniority) do not have a significant relationship with the selection of finalists.

This lack of transparency may contribute to claims of some reviewers and candidates that the NDPA process is too "controlled." As one candidate said, "Had the criteria been applied, they would have been good. But judging by the list of awardees, a completely different set of criteria was in fact applied, with political correctness high on the list." The data analyses performed for this process evaluation did not support these claims.

A concern was raised by the reviewers interviewed about the selection of awardees in FY 2005 and FY 2006. After the interview phase and panelist rankings, the finalists were presented to IC directors in order to secure funding for additional awards. While this resulted in nearly twice as many awards in FY 2005 and FY 2006 than would have been possible strictly through the NIH Roadmap funds, it also resulted in several individuals being funded from the "middle" and "bottom tiers" of the interviewees. Several panelists were concerned that individuals from the bottom tier were funded, as, in the words of one panelist, this "defeated the purpose of the panel review."

Despite their concerns about lack of feedback and transparency, in FY 2006 over half of the candidates who responded to the survey stated that they planned to reapply in future years - a general indicator of satisfaction with NDPA.

6.3 Characteristics of Candidates

During the first three years, the total number of NDPA candidates declined steeply - by FY 2006, there were only about half of the total number of candidates as in FY 2005 (469 vs. 833 total candidates) and about two-thirds down from FY 2004 (469 vs. 1,331 total candidates). Although the number of candidates declined, the overall demographic distributions of candidates did not change much in terms of gender, race/ethnicity, or institutional affiliation. Over the years, there were some changes in the distribution of candidates by research area - but the extent to which this change reflects differences in definition (e.g.,

the change from “clinical research” as a category descriptor in FY 2005 to “clinical and translational research” in FY 2006) rather than true changes in the candidate pool could not be determined. In FY 2006 there was an apparent increase in the percentage of candidates in clinical and translational research and a corresponding decrease in quantitative and mathematical biology. Despite efforts to increase the number of junior candidates (those with fewer than 10 years of experience), the number and percent of junior investigators has decreased in each year of the program.

In part, NDPA grew out of the desire to attract new researchers (i.e., those not already receiving NIH funds) to the NIH who were proposing “pioneering approaches to major challenges in biomedical and behavioral research.” However, the majority of the candidates were part of the existing NIH-supported external investigator community. In FY 2004, fourteen percent of candidates who responded to the survey never before had applied to NIH for funding. By FY 2006, this fraction was down to just four percent.

With respect to attracting new ideas or approaches (as distinct from attracting new individuals), candidates themselves believed their ideas to be at odds with prevailing wisdom, and indicated that they would be unlikely to receive funding for their proposed NDPA idea through a traditional mechanism. Evaluators differ with each other as to whether the applications they reviewed were “pioneering.” As might be expected, evaluators and liaisons interviewed believed that the initial pool of candidates attracted to the program was a “mixed-bag” - there were some very strong candidates, and others who clearly should have been applying for an R01 or other type of award. Evaluators indicated that the NDPA process itself was very different from a traditional study section, and was conducive to allowing investigators to submit more innovative and creative ideas.

6.3.1 Characteristics of Evaluators

In FY 2005 and FY 2006, an effort was made to attract a larger external evaluator pool, even as the number of applications dropped. Though many of the evaluators were repeat participants, new individuals were invited to participate (including more women and minority evaluators). This helped to ensure that individual evaluators did not have to review too many applications. In all years, the research area distribution of evaluators was comparable to the research areas of the candidates; however, some candidates expressed concern that the evaluators did not have the appropriate expertise to understand their ideas.

Collectively, the external evaluators were an accomplished group of researchers in their respective fields. More than two-thirds of the evaluators have won major awards or have been recognized with prestigious honors and fellowships.

6.3.2 Characteristics of NDPA Awardees

NDPA was created in the spirit of attracting unique researchers not typically supported through traditional NIH mechanisms. While six of the FY 2004 - FY 2006 awardees had received no prior support from the NIH as a PI (each of the three years of the program included two awardees who had not received support from NIH in the five years preceding their application), the majority are well-funded NIH investigators from elite institutions; collectively, the awardees in each of the cohorts held about thirty million dollars in NIH funding over a 5-year period. Most are basic researchers at universities or university-affiliated medical centers in the United States. Approximately thirty percent are women, and fewer than ten percent are underrepresented minorities. Awardees are also on average less senior than the candidates.

Despite holding other NIH awards and being part of the mainstream research community, nearly all awardees believe that their submitted ideas were sufficiently different that they could not be funded through other mechanisms or sources. Additionally, most evaluators, panelists and liaisons considered the awardees to be “pioneers” as well. However, when asked specifically if NDPA is meeting its goal of bringing in unique ideas, approaches, and/or people that are not being funded through NIH traditional peer review system, approximately half of the reviewers thought it was still too early to make that determination.

7. Recommendations for Future Years

In its first three years of implementation, NDPA has adapted and evolved to effectively improve and streamline the awardee selection process. Generally, the annual process evaluations found that the program is being implemented as designed; however, there are a few additional adjustments that should be made to further improve the process. Based on the analysis of interviews with liaisons and reviewers, surveys of candidates, and administrative data, STPI has compiled a list of recommendations for future years of the NIH Director's Pioneer Award. The recommendations are primarily aimed to increase program transparency and consistency while attempting to minimize the time-burden tradeoff of candidates and evaluators.

7.1 Program Design

Selection Criteria

- 7.1.1 *NDPA program leadership should clearly articulate whether NDPA is primarily intended to identify the best people, the best ideas, or a mixture of both. This would make the program purpose clear to all participants (candidates and evaluators alike).* Criteria changes initiated in FY 2005 caused some confusion as to the purpose of the program, as it appears to have shifted away from the original “people-based” model under which it was conceived. Some reviewers are informally “weighing” the first two criteria; some emphasize the “person” component while others emphasize the “idea” component introducing unnecessary variation into the program.
- 7.1.2 *Program leadership should revisit the “suitability” criterion and consider operationalizing it so it can be applied more consistently by reviewers.* The “suitability” criterion has not been explained clearly either in the RFA or the guidelines to the reviewers. If the term implies high-risk research, it should be stated clearly. As such, the criterion has been interpreted differently, and applied variably by external evaluators.
- 7.1.3 *If NDPA wants to continue the practice of requiring that awardees be able to devote 51% of their time to their NDPA research, then the constraint needs to be applied consistently and at the same stage.* Candidates should be asked to indicate at the application stage whether or not they can devote 51% of their time to NDPA if they win the award. After this initial consent, the candidates' commitment should not be second-guessed by various levels of reviewers. The requirement that applicants be available 51% of their time is being applied at different stages of the program by different reviewers, including at the end for final selection of awardees. It is also being applied inconsistently; some reviewers believe that most researchers qualified to be pioneers are prolific, and excluding those who are busy (but able to give up other research efforts were they to receive NDPA) is to the detriment of the program.

Program Structure and Evolution

- 7.1.4 *The number of qualified junior investigators (as well as their fraction) in the candidate pool has been shrinking; if the emphasis on young investigators is to remain, new strategies should be developed to attract them in future years.* NDPA program guidelines emphasize junior- to mid-level researchers, and as the statistical analysis shows, the program has been selecting younger investigators. Yet over the years, there has been a decline in the total number and percentage of junior candidates. In the last three years, the number of total candidates obtaining a degree within 10 years of the time of application decreased from 284 (22% of total candidate pool) in FY 2004 to 62 (13% of the total candidate pool) in FY 2006. If the emphasis on junior investigators remains, new strategies may need to be developed to attract qualified candidates, especially because the program is competing with other programs such as the NIH Director's New Innovator Award.^{xxiv}

7.2 Program Implementation

Adequacy of Information

- 7.2.1 *NDPA should eliminate the letters of reference since they are not used by all reviewers and might be unduly burdensome for some candidates. However, if letters of references continue to be used, providing a formal structure might ensure that comparable information is available for all candidates.* Candidates have indicated that letters of reference are difficult and burdensome to obtain, especially if they want to apply in multiple years. Evaluators have indicated that letters of reference are non-standardized, making it difficult to compare candidates. Generally evaluators believe that candidates ask their friends to write letters - all letters tend to be favorable but in different ways, and therefore do not provide them with objective information they need.
- 7.2.2 *Interviewees should get more detailed instructions on expectations for the presentations, though a structured format is not recommended.* Some interviewees reported that they did not have adequate instructions about what to present at the interviews. Similarly, several panelists stated that the presentations were largely heterogeneous and difficult to compare. More detailed instructions to the candidates would help to ensure that the presentations are analogous.
- 7.2.3 *NDPA should provide the panelists with more extensive briefing materials about each of the interviewees (e.g., top five publications of each interviewee, or other information as NIH deems appropriate).* Panelists were not uniformly prepared for the interviews; some dedicated time to extensively research each interviewee before the presentations, while others did not do any additional research. Some panelists indicated that they would like more information about the finalists before participating in the interviews. For example, NIH should consider having interviewees submit more than one example of their most significant accomplishments.

Transparency of Process

- 7.2.4 *NDPA should provide more feedback to candidates who are not invited to interview, and promote consistency across evaluators by requiring them to justify their scores (perhaps through a standard summary score template, and free-form comments).*

- In the first three years of implementation, NDPA candidates expressed concerns about the lack of feedback (and thus transparency) in the process. Some candidates would like to apply again; however, without knowing why their application was rejected before, they do not know how to improve. More specific feedback should be given to candidates if they are not invited to interview. To reduce burden on the external evaluators, they may be given a series of check box options to allow them to articulate why an application received the score that it did. Some synthesized version of the check box outcomes could be shared with the applicants.
- Some reviewers are not only introducing new criteria but also interpreting existing criteria differently from each other. Requiring score justification, perhaps through a series of check boxes and free form comments, would allow the NIH and the process evaluators to examine whether or not reviewers are using common criteria and applying them consistently. Such a step will promote both internal and external transparency of the Program.

7.2.5 *In future years, if funding is sought from NIH Institutes and Centers (ICs) for additional awardees, then that selection process should seek to support and reward the remaining top-tier and middle-tier finalists before funding bottom-tier finalists.* In FY 2005 and FY 2006 the list of finalists was presented to NIH ICs in order to secure funding for additional awardees. While this allowed NDPA to fund a larger number of qualified researchers, it also resulted in the funding of some researchers from the bottom “do-not-fund” tier of the panelist rankings. This led to unease on the part of the panelists who believed their advice was not heeded, and a sense of unfairness in the system (i.e., those who gain support from an Institute that has a specific research area of interest, even as other finalists with higher ranks do not).

7.3 Program Participation

7.3.1 *Attracting a diverse pool of investigators to the NDPA program should remain a top priority for the leadership team. As long as a diverse pool of candidates has been attracted, guidelines that encourage evaluators to pay special attention in their selection to underrepresented groups, levels of seniority, etc. may be redundant.* NDPA was created to be a merit-based program to specifically support exceptionally creative researchers. The FY 2004 competition strove to be purely merit-based with no specific focus on demographic distribution. Given NIH’s particular interest in promoting underrepresented groups, in FY 2005 and 2006 evaluators were specifically instructed to “watch for women, minorities, investigators at early to middle career stages” and to “eliminate very senior, well-funded investigators who are doing related work.” Some reviewers did in fact systematically eliminate or select candidates based on these guidelines. Others ignored the guidelines entirely, deeming them inappropriate for the program. Statistical analysis indicated that there may not be a relationship between demographics (other than seniority) and finalist selection. Finalist status is determined primarily by scores and top-4 status. The presence alone of the guidelines is creating a perception of bias in the program exacerbating other accusations of lack of transparency in the program.

Appendix: Evolution of the NDPA Program FY 2004 - FY 2006: Key Changes

Aspect of NDPA	Year of NDPA Process		
	FY 2004	FY 2005	FY 2006
<i>Candidate Recruitment Emphasis</i>			
<i>Emphasis given in PA or RFA</i>	“Investigators at early stages of their career as well as those who are established will be eligible”	“Investigators at all career levels are eligible. Those at early to middle stages of their careers, women, and members of groups underrepresented in biomedical research are especially encouraged to apply.”	“Investigators at all career levels are eligible. Those at early to middle stages of their careers, women, and members of groups underrepresented in biomedical research are especially encouraged to apply.”
<i>Definition of “pioneering” and “award” given in RFA</i>	<i>Not specifically defined</i>	The term “pioneering” is used to describe highly innovative approaches that have the potential to produce an unusually high impact, and the term “award” is used to mean a grant for conducting research, rather than a reward for past achievements	The term “pioneering” is used to describe highly innovative approaches that have the potential to produce an unusually high impact, and the term “award” is used to mean a grant for conducting research, rather than a reward for past achievements
<i>Selection Process - Phase Mechanics and Candidate/Evaluator Participation</i>			
<i>Number of Phases</i>	5	5	4
<i>Phase 1 (Nominees)</i>	Phase 1: 1,331 nomination packages (self-nominees and individuals nominated by someone else) - screened for responsiveness by 28 NIH liaisons.	Phase 1: 833 nominations (all self-nominees) were submitted and screened for responsiveness by 18 NIH liaisons.	Phase 1: 469 nominations (all self-nominees) were submitted and screened for responsiveness by 27 NIH liaisons.
<i>Phase 2 (Responsive Nominees)</i>	Phase 2: 936 responsive nomination packages were reviewed by 49 external evaluators (yes/no vote)	Phase 2: 567 nominees were deemed responsive and were reviewed by 47 external evaluators (yes/no vote)	<i>no initial (yes/no) screening by external evaluators in FY 2006</i>
<i>Phase 3 (Applicants)</i>	Phase 3: 239 individuals invited to submit a full application package to be reviewed by a second group of 29 external evaluators - scored on an inverted 7-point scale; “top-4” votes assigned	Phase 3: 283 individuals invited to submit a full application package to be reviewed by a second group of 37 external evaluators - scored on an inverted 5-point scale; “top-4” votes assigned	Phase 2: 406 responsive individuals were reviewed by a group of 80 external evaluators - scored on an inverted 5-point scale; “top-4” votes assigned
<i>Phase 4 (Interviewees)</i>	Phase 4: 22 of the applicants were invited to the NIH for an interview with a panel of 8 experts	Phase 4: 20 of the applicants were invited to the NIH for an interview with a panel of 13 experts	Phase 3: 25 of the applicants were invited to the NIH for an interview with a panel of 14 experts
<i>Phase 5 (Awardees)</i>	Phase 5: 9 awards were made on September 29, 2004	Phase 5: 13 awards were made on September 29, 2005	Phase 4: 13 awards were made on September 19, 2006
<i>Counts of Review</i>			
<i>Total Counts of External Review per Candidate</i>	6 total counts of review per application (3 external evaluators - Phase 2; 3 external evaluators - Phase 3)	5 total counts of review per application (2 external evaluators - Phase 2; 3 external evaluators - Phase 3)	3 total counts of review per application (3 external evaluators - Phase 2)

Aspect of NDPA	Year of NDPA Process		
	FY 2004	FY 2005	FY 2006
<i>Selection Criteria</i>			
<i>Criterion 1</i>	Innovation/creativity: Does the applicant display evidence of scientific creativity? Does she/he initiate new areas of, approaches to, scientific research? Is the applicant truly visionary in his/her thinking? Does the applicant think in complex, multi-, or interdisciplinary ways?	Scientific problem to be addressed: Biomedical significance/importance; if successful, likelihood of high impact on biomedical problem; Creativity/innovativeness	The scientific problem to be addressed: The biomedical significance/importance of the problem, the likelihood that, if successful, the project will have a significant impact on a biomedical problem, and the innovativeness of the project.
<i>Criterion 2</i>	Intrinsic motivation/enthusiasm/intellectual energy: Is the applicant willing to take scientific risks and show persistence in the face of adversity? Is the applicant comfortable with uncertainty? Is the applicant able to move into new areas that present an opportunity to solve a problem or expand knowledge base? Is the applicant intellectually independent and tenacious? Is the applicant able to make scientific leaps and change the current paradigms of medical research?	Investigator: Evidence for claim of innovativeness/creativity (innovation density); demonstrated ability to devote 51% or more effort on NDPA project	The investigator: Evidence for the investigator's claim of innovativeness/creativity (innovation density), and the demonstrated ability of the investigator to devote 51% or more effort on NDPA project.
<i>Criterion 3</i>	Potential for or actual scientific leadership; evidence of, or potential for, effective communication/educator skills: Does the applicant have the ability to communicate the impact of her/his work? Has the applicant shown the ability (or potential) to bring together diverse teams of scientists; to inspire with his or her scientific vision and lead others; to serve as a mentor or role model?	Suitability for NDPA mechanism: Evidence that proposed project is of sufficient risk/impact to make it more suitable for NDPA than for traditional NIH grant mechanism; distinct from other research by investigator	The suitability for NDPA mechanism: Evidence that the proposed project is of sufficient risk/impact to make it more suitable for the NDPA than for the traditional NIH grant mechanism and that it is distinct from other research previously or currently conducted by the investigator.
<i>Application Materials - Raw Materials and Time of Submission</i>			
<i>Application Material</i>	<i>Submitted in Phases</i>	<i>Submitted in Phases</i>	<i>Submitted up front</i>
<i>3-5 page essay</i>	Submitted at Application Phase	Submitted at Nomination Phase	Submitted at Nomination Phase
<i>2 page biosketch</i>	Submitted at Nomination Phase	Submitted at Nomination Phase	Submitted at Nomination Phase
<i>Current support</i>	<i>Not submitted in FY 2004</i>	Submitted at Nomination Phase	Submitted at Nomination Phase
<i>3 reference letters</i>	Submitted at Application Phase	Submitted at Application Phase	Submitted at Nomination Phase
<i>Most significant accomplishment</i>	Submitted at Application Phase	Submitted at Application Phase	Submitted at Nomination Phase
<i>300 word abstract</i>	<i>Not submitted in FY 2004</i>	<i>Not submitted in FY 2005</i>	Submitted at Nomination Phase
<i>Nomination letter</i>	Submitted at Nomination Phase	<i>Not submitted in FY 2005 - Self-nominees only</i>	<i>Not submitted in FY 2006 - Self-nominees only</i>

Aspect of NDPA	Year of NDPA Process		FY 2006
	FY 2004	FY 2005	
<i>Research Areas of Candidates</i>			
<i>Option 1</i>	Behavioral and Social Science	Behavioral and Social Science	Behavioral and Social Science
<i>Option 2</i>	Clinical Research	Clinical Research	Clinical and Translational Research
<i>Option 3</i>	Instrumentation and Engineering	Instrumentation and Engineering	Instrumentation and Engineering
<i>Option 4</i>	Molecular and Cellular Biology	Molecular and Cellular Biology	Molecular, Cellular, and Chemical Biology
<i>Option 5</i>	Pathogenesis and Epidemiology	Pathogenesis and Epidemiology	Pathogenesis and Epidemiology
<i>Option 6</i>	Physiological and Integrative Systems	Physiological and Integrative Systems	Physiological and Integrative Systems
<i>Option 7</i>	Quantitative and Mathematical Biology	Quantitative and Mathematical Biology	Quantitative and Mathematical Biology

Source: STPI Analysis of FY 2004 - FY 2006 NDPA Programmatic Information

Endnotes

ⁱ Interviews with NIH staff.

ⁱⁱ “Proceedings of the 88th Meeting of the Advisory Committee to the Director.” National Institutes of Health. May 6, 2004. Available online at: <http://www.nih.gov/about/director/acd/minutes/050604acd.htm>

ⁱⁱⁱ “Risky Business: Can the U.S. government do a better job of betting on long shots in science? NSF and NIH hope the answer is yes.” *Science*. October 8, 2004. 220-221.

^{iv} Brenner, Sydney. “The Impact of Society on Science.” *Science*, 1998, 282, 1411.

^v The NIH Director’s Pioneer Award Program press release. Jan. 20, 2004. Available online at: <http://www.nih.gov/news/pr/jan2004/od-20.htm>.

^{vi} “Working Group on Review of Bioengineering and Technology and Instrumentation Development Research.” Center for Scientific Review. May 1999.

^{vii} The term “candidates” refers to all investigators who applied to the NDPA program in FY 2004 - FY 2006. Candidates include “nominees,” who were either nominated by others (in FY 2004 only) or who self-nominated themselves (in FY 2004 - 2006); “applicants,” who were invited to submit full applications; “interviewees,” who were invited to present their ideas in-person; as well as awardees.

^{viii} The term “reviewers” refers to all individuals who participated in some phase of the review - both liaisons and external evaluators; “liaison” refers to the NIH representatives who participated in the administrative portion of the review; and “evaluators” refers to the body of external evaluators who participated in some round of the scientific review.

^{ix} Although this is a commonly held belief, the Year 1 process evaluation did not find such a bias. For example, men and women were similarly distributed in the self- and other-nominated pools.

^x Degrees of candidates and seniority were coded using information from the biographical sketches and personal websites. STPI considered the number of years elapsed since the nominee obtained his or her MD or PhD (and for candidates with both MD and PhD or multiple PhD degrees, the earlier degree obtained was used to calculate seniority). Information was not available for some of the nominees - they were coded as “N/A” (Seniority data are most incomplete for the FY 2006 cohort, as the application materials were not available and all seniority had to be coded based on personal websites).

^{xi} The remaining 4% hold one of the following degrees: BA/BS, DMD, DVM, Ed.D., JD, or PharmD.

^{xii} Because race/ethnicity data of NDPA candidates from NIH sources was not 100% complete, an optional survey question also was asked of candidates in order to get additional information on the racial/ethnic distribution.

^{xiii} Race and Ethnicity categories correspond to those stored in the NIH IMPAC II database.

^{xiv} Shavers, V.L. *et al.* 2005. Barriers to Racial/Ethnic Minority Application and Competition for NIH Research Funding. *Journal of the National Medical Association*, 97(8): 1063-1077

^{xv} National Research Council, “Assessment of NIH Minority Research and Training Programs, Phase 3,” 2005.

^{xvi} Susan D. Hosek, et al. “Gender Differences in Major Federal External Grant Programs.” RAND Technical Report, RAND Corporation, 2005. ISBN 0-8330-3854-0.

^{xvii} National Research Council, Commission on Life Sciences. “Trends in the Early Careers of Life Scientists.” National Academy Press, 1998; Erica Goldman and Eliot Marshal. “NIH Grantees: Where Have All the Young Ones Gone?” *Science*, Vol. 298, 4 October 2002, 40-41.

^{xviii} Vickie L. Shavers, et al. “Barriers to Racial/Ethnic Minority Application and Competition for NIH Research Funding.” *Journal of the National Medical Association*, Vol. 97, No. 8, August 2005, 1063-1077.

^{xix} Tamara Z. Zemlo, et al. “The Physician-Scientist: Career Issues and Challenges at the Year 2000.” *The FASEB Journal*, 2000,14:221-230. <http://www.fasebj.org/cgi/content/full/14/2/221>

^{xx} Vickie L. Shavers, et al. “Barriers to Racial/Ethnic Minority Application and Competition for NIH Research Funding.” *Journal of the National Medical Association*, Vol. 97, No. 8, August 2005, 1063-1077

^{xxi} Shavers, *ibid.* IMPACII database. Courtesy of Belinda Seto, National Institute of Biomedical Imaging and Bioengineering (NIBIB).

^{xxii} FY 2004 PA and FY2005 and FY 2006 NDPA RFAs. Available online at:
<http://nihroadmap.nih.gov/pioneer/>

^{xxiii} It should be noted that because the total number of external evaluators increased in each year, there were more “top-4” votes given in FY 2006 than in FY 2004 of FY 2005.

^{xxiv} NIH Director’s New Innovator Award:

http://grants.nih.gov/grants/new_investigators/innovator_award/