Transformative Research – An Exploration of Six Propositions

By Bhavya Lal

S.B. in Nuclear Engineering, June 1990, Massachusetts Institute of Technology
S.M. in Nuclear Engineering, June 1990, Massachusetts Institute of Technology
S.M. in Technology and Policy, June 1992, Massachusetts Institute of Technology

A Dissertation Submitted to

The Faculty of
The Columbian College of Arts and Sciences
of The George Washington University
in partial fulfillment of the requirements
for the degree of Doctor of Philosophy

August 31, 2012

Dissertation directed by

Nicholas Vonortas
Professor of Economics and International Affairs

The Columbian College of Arts and Sciences of The George Washington University certifies that Bhavya Lal has passed the Final Examination for the degree of Doctor of Philosophy as of July 3, 2012. This is the final and approved form of the dissertation.

Transformative Research – An Exploration of Six Propositions

Bhavya Lal

Dissertation Research Committee:

Nicholas Vonortas, Professor of Economics and International Affairs, Dissertation Director

Kathryn Newcomer, Professor of Public Policy and Public Administration Committee Member

Scott Pace, Professor of the Practice of International Affairs, Committee Member

To Kinjal

Acknowledgements

My greatest gratitude is reserved for my advisor and dissertation chair Nicholas Vonortas for seeing the potential of this topic, and giving me the opportunity to work on it. I am grateful to him not only for the confidence he showed in me, and for the incalculable ways he stretched me intellectually, but also for the clear direction I got whenever I drifted too far off course.

I thank my committee members Scott Pace and Kathryn Newcomer, for pushing me to defend every statement, and in so doing, sharpening my work and preparing me for any audience. Thanks also go to my readers David Bray and Valerie Schneider for their guidance and advice, and thorough readings of many versions of this document, even the early ghastly ones.

I would like to acknowledge the contributions of my colleagues at the Science and Technology Policy Institute (STPI), especially Alyson Wilson, Stephanie Shipp, Mary Beth Hughes, Mario Nunez, Seth Jonas, Amy Marshall, Elizabeth Lee, Asha Balakrishnan and Vanessa Pena, who were alongside me in our many high-risk high-reward adventures, and who guided me in the completion of the dissertation. They were my thought partners in every way. Above all, they provided me with an intellectual home at STPI.

Last but not least, I would like to thank President David Chu and the leadership of the Institute for Defense Analyses (IDA) and the Science and Technology Policy Institute (STPI), for not letting me slow the momentum of this work, and the IDA Scholar program, that allowed me to start and complete the doctoral program while still working full time and raising a family

Abstract of Dissertation

Transformative Research – An Exploration of Six Propositions

U.S. federal programs that fund transformative research have proliferated in recent years, with both Congress and the Administration urging science agencies to fund more of it. However, there appears to be no firm understanding of the concept. Its definitions tend to be inspirational but vague. More importantly, there is no operational agreement on how to identify, fund, and evaluate such research. This dissertation attempted to bridge this knowledge gap, and examined transformative research using qualitative case study, quantitative analysis, and text mining methods.

Building on two reviews, the literature in the interdisciplinary field of the science of science policy, and federal programs that support transformative research, I developed six propositions - three about those who conduct transformative research, and three about transformative research itself. The propositions were then explored using data from a transformative research program created at the National Institutes of Health.

Findings from the exploratory analysis showed that projects supported by transformative research is neither more interdisciplinary, nor are its performers younger, more productive, or more distinctive in their track records. Moreover, while transformative research is perceived to be risky, riskiness of proposals does not appear strongly associated with transformative outcomes. The only distinguishing characteristic of transformative research is the level of disagreement among peers reviewing it.

The findings make intuitive sense, and may not appear new or unexpected. However what makes them different is that they are supported by data, a feature that the current literature on transformative research lacks. The study has limitations, the principal ones being that the findings were based on basic research in biomedical sciences, and no attempt was made to generalize to other research fields. Two potential policy implications emerge assuming that the propositions are tested more broadly: to achieve transformative outcomes, research programs may be better served by funding high-quality researchers to conduct high-quality research, rather than trying to identify potentially transformative ideas. In addition, extra attention could be paid to research proposals with divergent peer rankings, as this divergence may suggest the kernel of a paradigm-shifting idea.

Table of Contents

Dedication iii

Acknowledgements iv

Abstract of Dissertation v

Table of Contents vi

List of Figures ix

List of Tables xi

1.0 Introduction 1

1.1 Background 1

1.2 What is "Transformative Research"? 2

1.3 Origin of the Concept in S&T Policy 4

1.4 The Policy Challenge 9

1.5 Research Questions and Overall Approach 10

2.0 Defining Transformative Research and Related Concepts 12

2.1 Defining Transformative Research 12

2.2 Key Characteristics 20

2.3 Operationalizing Transformative Research 26

2.4 Operationalizing High-Risk Research 39

2.5 Conclusion 41

3.0 Review of Programs Funding Potentially Transformative Research 43

3.1 Methodology 43

3.2 Overview of Programs 46

3.3 Definitions of Terms of Interest 51

3.4 Program Origins 53

3.5 Areas of Funding 54
3.6 Competitiveness of Programs – Funding Rates 55
3.7 Timeline and Renewals 58
3.8 Selection Criteria 60
3.9 Post-Award Support 63
3.10 Types of Programs Found – By Program Goals 63
3.11 Types of Programs Found –By Management Strategies 66
3.12 Types of Programs Found – by Approach to Research 71
3.13 Conclusion 72
4.0 Introducing Propositions 74
4.1 Proposition 1: Track Record 74
4.2 Proposition 2: Youth 75
4.3 Proposition 3: Productivity 77
4.4 Proposition 4: Risk 77
4.5 Proposition 5: Interdisciplinarity 78
4.6 Proposition 6: Skepticism 79
4.7 Conclusion 80
5.0 Methodology 82
5.1 Data Source - The NIH Director's Pioneer Award (NDPA) 82
5.2 Overall Approach 88
5.3 Operationalizing Concepts of Interest 91
5.4 Developing a Comparison Group 95
5.5 Data Sources 101
5.6 Potential Limitations 102

5.7 Summary 106

6.0 Exploring the Propositions 111

6.1 Proposition 1: Track Record 111

6.2 Proposition 2: Youth 115

6.3 Proposition 3: Productivity 117

6.4 Proposition 4: Risk 118

6.5 Proposition 5: Interdisciplinarity 137

6.6 Proposition 6: Skepticism/Peer Disagreement 141

6.7 Summary 153

7.0 Overall Summary and Conclusion 157

7.1 Summary 157

7.2 Recommendations for Future Research 159

7.3 Recommendations for R&D Managers 161

7.4 Conclusion 168

8.0 References 171

Appendix A: Questionnaire for Program Managers 197

Appendix B: Evolution of the Pioneer Program 198

Appendix C: Pioneer Research Characteristics 202

List of Figures

Figure 1.1. Location of estimated worldwide R&D expenditures: 1996 and 2009. 5

Figure 1.2. Average annual growth rates in number of researchers, by country/economy: 1995–2002 and 2002–09. 5

Figure 1.3. Engineering articles, by selected region/country: 1995–2009 5

Figure 3.1. A model to illustrate a traditional research funding model 67

Figure 3.2. A model to illustrate the approach used by transformative research programs 68

Figure 3.3. A model to illustrate various ways in which transformative research is funded 69

Figure 3.4. NSF's EFRI program – A synergy program 70

Figure 3.5. ARPA agencies – End game programs 70

Figure 4.1. Propositions expressed in the proposed framework 81

Figure 4.2. Potential policy recommendations if propositions were supported by data 81

Figure 5.1. NDPA program logic model 84

Figure 5.2 NDPA's approach to supporting transformative research 85

Figure 5.3. Direct costs for 35 R01s in matched comparison group 97

Figure 5.4. Institutional prestige for NDPA and R01 comparison group 98

Figure 5.5. Years since degree for NDPA and R01 comparison group 98

Figure 5.6. Prior NIH funding for NDPA (left) and R01 comparison group 99

Figure 6.1. Productivity of Pioneers and comparison group researchers 113

Figure 6.2. H-Index of Pioneers and comparison group researchers 113

Figure 6.1. Citations of Pioneers and comparison group researchers 110

Figure 6.4. Journal impact factors of Pioneers and comparison group researchers 114

Figure 6.5. Distribution of years since Ph.D. of researchers in groups of interest 116

Figure 6.6. Distribution of years since Ph.D. for researchers by group 116

Figure 6.7. Researcher-level publications, 1980 start of grant 118

Figure 6.8. Distribution of risk scores for awardees and nonawardees 126

Figure 6.9. Relationship between risk and whether a proposal was an ideal candidate 128

Figure 6.10. Expert scores on the Pioneeringness of NDPA research 130

Figure 6.11. Averaged ranking of pioneers by experts 130

Figure 6.12. Distribution of the number of researchers in similar fields 135

Figure 6.13. Mapping NDPA Citations with Crowdedness 136

Figure 6.14. Mapping H-Index of Pioneers with Crowdedness 136

Figure 6.15. Integration Score of NDPA and R01 Programs 138

Figure 6.16. Specialization of grant-funded publications 139

Figure 6.17. Integration scores of publications citing grant-funded research 140

Figure 6.18. PIs' integration scores and specialization scores 141

Figure 6.19. Grant attributed citations through 2011 147

Figure 6.20. Citations to NDPA-funded publications, broken down by PI 148

Figure 6.21. Citations to Matched R01-funded publications, broken down by PI 148

Figure 6.22. Time to publication from time of grant 149

Figure 6.23. Citation latency of the test and comparison group publications 150

Figure 6.24. Notional agreement "distance" among reviewers 151

Figure 6.25. Evaluator assessment of the uniqueness/value of the Pioneer mechanism 153

Figure 7.1. Policy recommendations emerging from research and analysis 168

List of Tables

Table 2.1 *Typology of the Outcomes of Creative Research* 14

Table 2.2 *Innovation Definitions* 18

Table 2.3 *Measures Used in Recent Evaluations of Transformative Programs* 31

Table 3.1 *U.S. Federal Government Research Programs Examined* 45

Table 3.2 *Terms Used by Programs* 52

Table 3.3 *Program Funding Rates* 57

Table 3.4 *Timeframe for Each Type of Grant* 60

Table 3.5 *Length of Application Form and Prescreening of Proposals* 62

Table 3.6 *A Proposed Typology of Transformative Funding Programs* 64

Table 3.7 *Major Designs Seen in Transformative Research Programs* 65

Table 4.1 *Propositions About Transformative Research* 80

Table 5.1 *Relevant Comparison for Each Proposition Explored* *98*

Table 5.2 *Proposed Approach to Testing Propositions* 108

Table 6.1 *Descriptive Age-Related Statistics for NDPA and All R01 Awardees* 116

Table 6.2 *Awardees' Assessment of the Nature of the Risks of Their Research* 121

Table 6.3 *Experts' Assessment of the Nature of the Risks of the Pioneer's Research* 123

Table 6.4 *Conversion of 2004 Scores to 2005 Equivalents* 125

Table 6.5 *Pioneer Pre-Award Risk and Post-Award Pioneeringness Rankings* 132

Table 6.6 *Mapping Perceived Risk with Pioneering Outcomes* 132

Table 6.7 *Success Rates According to Accomplishment of Stated Goals* 134

Table 6.8 *Average Specialization, Integration, and Diffusion Scores of Pioneers and Matched R01 Grantees* 138

Table 6.9 *A Typology of Interdisciplinary Researchers* 140

Table 6.10 *Inter-rater Reliability Scores for Different Groups of Proposals* 144

Table 6.11 *Proposals Recommended for Funding and their Postaward Rating* 146

Table 7.1 *Summary of Findings from the Analysis* *162*

Table 7.2 *Overcoming Threats to Measurement Validity for Transformative Research Evaluations* 166

Table B1.1 *NDPA Process Changes in Detail: Candidate Recruitment Emphasis* 198

Table B1.2 *NDPA Process Changes in Detail: Selection Process* 199

Table B1.3 *NDPA Process Changes in Detail: Selection Criteria* 200

Table C1.1 *Summary of Pioneer Research* 202

1.0 Introduction

In this dissertation, I define transformative research, develop six propositions that distinguish it and those who conduct it, and explore these propositions using data from a program at the National Institutes of Health (NIH). This chapter provides a brief introduction to transformative research and its origin and evolution in the policy community, introduces my research questions, and outlines the approach taken to address them.

1.1 Background

Over the last decade, the scientific community has increasingly come to believe that the current science and technology (S&T) research portfolio has become too conservative, and encourages only incremental advances (National Science Foundation [NSF], 2004; Mervis, 2004). For example, in a survey of its grantees and reviewers, NSF found that almost two-thirds of the reviewers consider less than ten percent of what they review as transformative (NSF, 2007i). Similar perceptions exist in the biomedical sciences community as well (Brainard, 2007). Other than the observations of the community, there is no evidence, however, whether research funding has become more conservative; the perception seems connected to falling funding levels of research. The NSF report summarizes the connection (NSF, 2007i):

> With respect to impacts on transformative research, a widely-held concern is that as funding rates drop, reviewers become more conservative and less receptive to revolutionary ideas that challenge existing paradigms. This in turn discourages PIs from submitting proposals containing potentially transformative research ideas; as a consequence, support of transformative research decreases. (p. 13-14)

As a result of this concern, a number of government agencies and privately-funded organizations have initiated new efforts to remedy the situation, and begun to support what is generally referred to as transformative or potentially transformative research. Despite the proliferation of such programs, there appears to be no theory of change or motivation that accompanies the design and implementation of these programs, and it remains unclear if such set-aside programs produce more transformative research than traditional ones.

1.2 What is "Transformative Research"?

Transformative research is an amorphous concept, and defined differently based on the context in which the term is used. Chapter 2 explores the definition in detail; for the purpose of this chapter, suffice it to say that it refers to a type of research that has "the potential to radically change our understanding of an important existing … concept or leading to the creation of a new paradigm" (National Science Board [NSB], 2007, p. v). In his book *Pioneering Research*, British researcher Donald Braben contends that transformative research is largely responsible for dramatic increases in global gross domestic product during the past three centuries, with a recent slowing in growth attributable to the additional constraints placed on scientists interested in doing high risk research (Braben, 2004). It carries with it the connotation of being something beyond business-as-usual, and a shifting of paradigms. Classic examples include Darwin's theory of natural selection that replaced Lamarckism as the mechanism for evolution, development of quantum mechanics, which supplanted classical mechanics, and plate tectonics that replaced the static geosynclinals theory in describing continental drift.

Despite the preponderance of anecdotes in the basic sciences, it is important to remember that transformative research is not the province of bench science alone. Methodological or instrumental advances can be transformative as well. The invention of the

scanning tunneling microscope, for example, was fundamental to the development of research areas such as high-energy physics and nanotechnology.

The concept is just as applicable to activities outside scientific research. In the realm of education, for example, the development of the *Force Concept Inventory in Physics*, a framework to measure students' understanding of basic concepts in Newtonian physics, was considered an unexpected insight, and continues to lead to the development of new educational methodologies.

Going beyond the domain of science altogether, transformativeness can also come from application-oriented activities. Examples include the Green Revolution – a series of initiatives that were transformative in that they saved the lives of a billion human beings worldwide (Avery, 2006).

As the section below illustrates, the nation's interest in transformative research has increased in recent years. The number of programs that explicitly aim to fund transformative research has exploded over the past 5 years, appearing in eight agencies and with programs levels ranging from $30-$300 million (Hughes et al., 2010). The trend isn't confined just to the United States. The High Innovation/Gain/Expectation Program survey identified forty European funding agencies that claimed to have specific programs supporting transformative research projects (Pendergast, 2007). In 2012, European research agencies, through programs like EU's NEST (New and Emerging Science and Technology), are funding about $1.9 billion in transformative research (Froderman & Holbrook, 2012, p. 42). China has begun to go beyond peer-reviewed research with the *pinyin* concept, with the Natural Science Foundation of China adopting a series of special policies on transformative research in its 12th 2011-2015 Five-year Program (Wang, 2012).

1.3 Origin of the Concept in S&T Policy

The earliest reference to transformative research in the S&T policy community can be found in deliberations of the National Science Foundation (NSF). In 1980, the NSF Advisory Council established a Task Force on "the problem of funding those research proposals which are highly creative or innovative, but which there is a high risk of failure of accomplishing the proposed goals" (NSF 1980, p. 1). The recommendations of the task force got little traction, and the concept found its profile raised only at the start of this millennium, when the U.S. research portfolio, once regarded as the world's best, began to be characterized as being under strain, with federal funding of R&D (Figure 1.1) and researchers (Figure 1.2) in relative decline, and increasingly being challenged by other nations in various fields of research (Figure 1.3).

It was against a backdrop of uncertainty and anxiety, that in 2005, the National Academy of Science's Committee on Science, Engineering, and Public Policy (COSEPUP) was charged by the US Congress to identify the top ten actions to ensure that the United States can "successfully compete, prosper, and be secure in the global community of the 21st century." (NAS, 2007, p. x-xi). The study, led by former Lockheed Martin CEO Norman Augustine, was published two years later. Entitled *Rising Above the Gathering Storm: Energizing and Employing America for a Brighter Economic Future* (RAGS), the report declared that,

> Having reviewed trends in the United States and abroad, the committee is deeply concerned that the scientific and technological building blocks critical to our economic leadership are eroding at a time when many other nations are gathering strength. …We are worried about the future prosperity of the United States. Although many people assume that the United States will always be a world leader in science and technology, this may not continue to be the case inasmuch as great minds and ideas exist throughout the world. We fear the abruptness with which a lead in science and technology can be lost—and the difficulty of recovering a lead once lost, if indeed it can be regained at all. (NAS 2007, p. 3)

The authors of RAGS identified the recent decline in support of "high-risk or transformative research," particularly in the physical sciences, engineering, mathematics, and information sciences as one major factor that contributes to the United States' eroding competitiveness in the global economy (NAS 2007, p. 149).

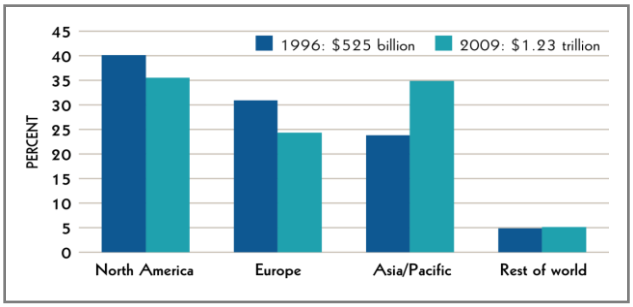


Figure 1.1. Location of estimated worldwide R&D expenditures: 1996 and 2009.
Source: Science and Engineering Indicators (2012). Arlington, VA (NSB 12-01)

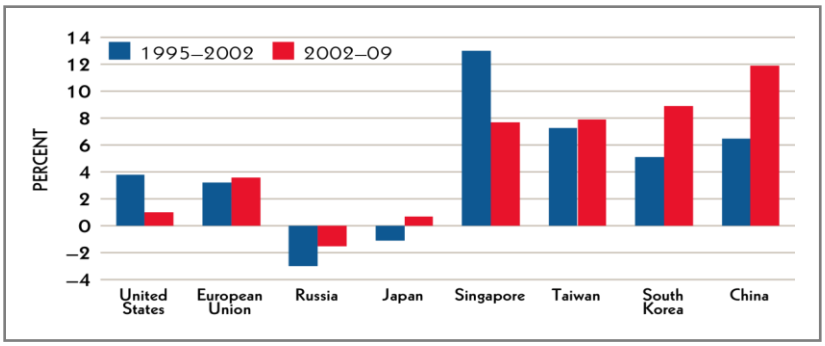


Figure 1.2. Average annual growth rates in number of researchers, by country/economy: 1995–2002 and 2002–09.
Source: Science and Engineering Indicators (2012). Arlington, VA (NSB 12-01)

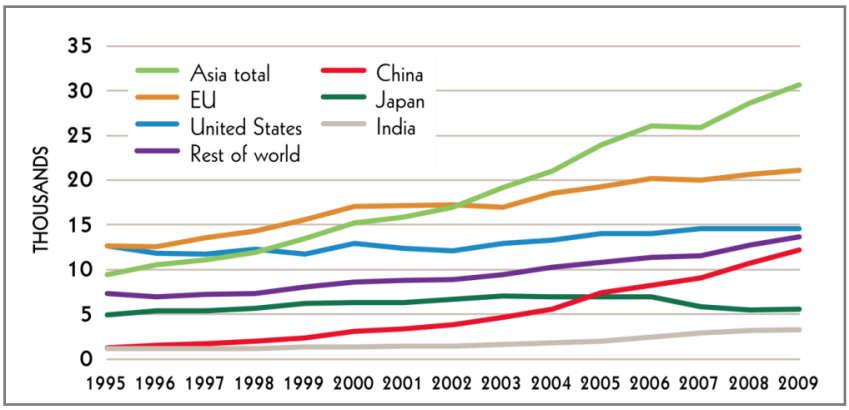


Figure 1.3. Engineering articles, by selected region/country: 1995–2009.
Source: Science and Engineering Indicators (2012). Arlington, VA (NSB 12-01)

The report listed the factors that inhibit funding of transformative research. These factors included (a) flat or declining funding in many disciplines that makes it harder to justify risky or unorthodox projects; (b) the peer review system that tends to favor established investigators who use well-known methods; (c) pressure to produce short-term results;(d) increased public scrutiny of government R&D spending that makes it harder to justify nonpeer-reviewed awards, (e) peer reviewers' tendency to place confidence in older, established researchers; and (f) the fact that high-risk, high-potential projects are prone to failure, and government oversight and media and public scrutiny make those projects increasingly untenable to those responsible for the work.

The report further noted that "reducing the risk for individual research projects *increases* the likelihood that breakthrough, 'disruptive' technologies will not be found—the kinds of discoveries that yield huge returns" (p. 33). One of the recommendations labeled as "most urgent" therefore was to "sustain and strengthen the nation's traditional commitment to long-term basic research that has the potential to be transformational, to maintain the flow of new ideas that fuel the economy, provide security, and enhance the quality of life" (p. 7). One of the major recommendations of RAGS was that at least 8% of the budgets of federal research agencies should be set aside for discretionary funding to catalyze high-risk, high-payoff research.

The National Science Board (NSB), which oversees the NSF, also organized a Task Force on Transformative Research that met during 2004-06. The Task Force's 2007 report (also referred to as the "butterfly" report because of the presence of a stylized butterfly on its cover), entitled *Enhancing Support of Transformative Research at the National Science Foundation* (NSB, 2007), noted that "the underlying concern … is that failure to encourage and to support revolutionary ideas will jeopardize not only our Nation's ability to compete in

today's and tomorrow's global economy, but also the progress of science as a whole" (p. 2). In fact, it is the Board that could be credited with creating the term "transformative research" in anticipation of a conference in September 2004. [1]

There were similar concerns bubbling in the defense establishment. In a 2005 assessment of the Department of Defense basic research portfolio, the National Research Council (NRC, 2005) had found a general decline in support of basic research over the past decade, and commented on the de-emphasis on "unfettered exploration, which historically has been a critical enabler of the most important breakthroughs in military capabilities" (p. 17)

In 2008, the academic community published its own report, *Advancing Research in Science and Engineering. Investing in Early-Career Scientists and High-Risk, High-Reward Research* (The American Academy of Arts and Sciences, 2008). The report reinforced the findings of the RAGS and NSB reports:

> Leadership in science and technology is necessary to compete effectively in the global economy. Today the dominant position of the United States in the international research and education community is being challenged as never before. … We strongly believe that, regardless of overall federal research funding levels, America must invest in young scientists and transformative research in order to sustain its ability to compete in the new global environment. (ARISE, 2008, p. 1)

These reports, especially RAGS, galvanized Congressional action. Congress held hearings on investing in high-risk, high-reward research, getting input from representatives from academia, private foundations, and the R&D agencies themselves on the need for

[1] Author was in attendance at the Santa Fe conference entitled "Identifying, reviewing, and funding transformative research," and wrote the White Paper preceding the conference under contract with the National Science Board Office.

novel funding mechanisms (Congressional Hearings, 2009). Congress also considered providing funding for new high-risk, high-reward programs that previously had been authorized but not funded, such as the Advanced Research Projects Agency-Energy (ARPA-E). Activities culminated in the form of the *America COMPETES Act* of 2007 (America COMPETES, 2007). The Act, with broad bipartisan support, expressed a sense of Congress that "each federal research agency should support and promote innovation through funding for high-risk, high-reward research" (America COMPETES, Section 1008).

Agencies responded rapidly to the challenges. Just months before the final RAGS report, based in part on concerns that the traditional peer review process had become overly conservative (Langfeldt, 2006), the NIH had unveiled the *Roadmap for Medical Research* (NIH 2004; Mervis, 2004). The *Roadmap* aimed to promote high-impact, cutting-edge research, which often did not fall into the interest of a single NIH Institutes or Centers. NSF wasn't far behind. In addition to creating transformative research programs (discussed in Chapter 3), in 2007, NSF issued Important Notice No. 130: Transformative Research, which announced a change to NSF's Intellectual Merit Review Criterion effective January 5, 2008 (Bement, 2007). Reviewers would now be asked: "To what extent does the proposed activity suggest and explore creative, original, or potentially transformative concepts"?

The Executive branch also responded to the calls to action by directing agencies through the annual *Office of Science and Technology Policy – Office of Management and Budget R&D Priorities Memo* to explain in agency budget submissions how they will support this type of research (OSTP 2010): "Agencies should pursue transformational solutions to the Nation's practical challenges, and budget submissions should therefore explain how agencies will support long-term, visionary thinkers proposing high-risk, high-return (or "potentially transformative") research" (p. 2). President Obama, in a speech at the National Academies

of Science Annual Meeting announced the creation of the Advanced Research Projects Agency for Energy (ARPA-E), which sought to support high-risk, high-reward research related to energy (NAS, 2010).

Needless to say, the country's research enterprise has been viewed by policymakers from Vannevar Bush (1945) to the 111th Congress (Public Law, 111-358) as undergirding researchers' ability to innovate and provide leadership in all of these areas. Transformative research is the new silver bullet, and policymakers' interest in it continues unabated.

1.4 The Policy Challenge

As the Section 1.2 illustrates, programs that fund transformative research have proliferated, with both Congress and the Administration urging the science agencies to fund more of it. The concept of transformative research, however, remains mired in mystique. At the same time, there appears no firm understanding of what transformative research is. As Chapter 2 shows, its definitions tend to be sublime - inspirational for certain, but also vague. More importantly, as the diversity of program implementation in Chapter 3 shows, there is no operational understanding of what transformative research is. So what should a transformative research program do to identify and fund such research?

There are other conundrums as well: To what extent can transformative research be deliberately created? In the past, it was often created either serendipitously – note the experiments of Alexander Fleming who accidentally discovered penicillin – or simply through the perseverance of scientists conducting incremental research – note the discovery of graphene, the allotrope of carbon expected to transform materials. Is it possible to *orchestrate* the creation of transformative research? If so, are there theories that could reasonably assure the production of transformative research? Indeed, is there a "science of

transformative science?" These are important questions to contemplate, and to answer them, transformative research must be better characterized than it is today.

1.5 Research Questions and Overall Approach

In this dissertation, I endeavored to initiate such an effort, posing three primary questions. First, what is transformative research, and how is it defined and operationalized in the literature and the S&T policy community? Second, to what extent do attributes such as riskiness of research, interdisciplinarity, and peer skepticism describe transformative research, and can be supported by data? Third, to what extent do attributes such as youth, track record, and prior productivity describe those who conduct such research, and can be supported by data?

The first step to address these questions was to conduct a brief review of the relevant literatures in five areas: history and philosophy of science, the sociology of science, psychology, organizational theory, and finance. This review is summarized in Chapter 2. The next step was to compile a list of federal programs, and collect and synthesize information about them to see if there were common threads in the way transformative research programs were being designed and implemented. The review was conducted not to assess or evaluate these programs, but to explore what "theories of change" they were based on, in an attempt to discover heuristics that were missing in the academic literature. The findings from the program review are summarized in Chapter 3.

Both these reviews led to the discovery and development of six "propositions" about transformative research and those who conduct it. Based on a review of the literature of the science of science policy , and a review of federal programs funding transformative research, some heuristics about transformative research came to the fore. These are more perceptions

than they are proven principles. They are stated as propositions, and form the core of this dissertation. These proposition are summarized in Chapter 4.

The qualitative and quantitative methods to explore the propositions are laid out in Chapter 5. The Chapter begins with an introduction of the case study used in the dissertation – the NIH Director's Pioneer Program (NDPA) program, a transformative research program funded by NIH. The NDPA program was selected primarily because of the availability of data about the selection process, and about the outcomes of the program. The chapter also describes the analytic approach for each proposition, including data sources, the comparison groups developed to explore the propositions, and how all concepts of interest were operationalized. The chapter also discusses the potential limitations of the findings, the most important of which is that the propositions are explored using data from a single basic research oriented program in biomedical research, and cannot be generalized to other fields of research. The propositions must be tested in other domains of research, in particular the physical sciences and engineering.

Chapter 6 is the heart of the dissertation where each of the propositions is explored using comparative data when feasible. All the findings are then synthesized into conclusions and lessons learned, and these insights are summarized in Chapter 7, both in the form of recommendations for future research, and lessons for managers of R&D programs.

2.0 Defining Transformative Research and Related Concepts

In this chapter, I review several streams of literature under the rubric of the "science of science policy." This includes literature in the history and philosophy of science, the sociology of science, psychology, organizational theory, and finance, to explore the definition of the term "transformative research." Section 2.1 explores the definition of transformative research and related terms, Section 2.2 focuses on characterizing transformative research, and Section 2.3 summarizes the various ways this type of research has been measured. Together these sections help discover the role of "risk" in transformative research. Sections 2.4 and 2.5 examine the business and finance literatures (in addition to the literature in the sociology of science) to define and operationalize the term "high risk." The final section summarizes the findings.

2.1 Defining Transformative Research

While the term "transformative research," as used in the context of scientific research, is relatively new, the study of revolutions in science itself is not. Indeed, the concept of scientific revolution, while clarified by Kuhn (and discussed below), had its predecessor ideas (especially in the domain of economics and society) in the works of Joseph Schumpeter. In *Capitalism, Socialism and Democracy*, Schumpeter (1942) introduced the term "creative destruction" to describe innovative entry by entrepreneurs as the force that sustained long-term economic growth, even as it destroyed the value of established companies that enjoyed some degree of monopoly power. According to Schumpeter, because of the significant barriers to entry that monopolies enjoyed, new entrants would have to be radically different, ensuring fundamental improvement was achieved, not a mere difference of packaging. The threat of market entry, Schumpeter claimed, would keep monopolists and oligopolists' disciplined and competitive, ensuring they invest their profits

in new products and ideas. Schumpeter believed that it was this innovative quality that made capitalism the best economic system.

While the concept of "creative destruction" was articulated in the context of a capitalist free market society, it has just as apt a use in the domain of science. The words have changed from "creative destruction" to many of its synonyms (e.g., "transformative," "pioneering,"), but the concept endures. Nonetheless, there exist no commonly accepted definitions, and no coherent typology. The situation is complicated by differences in opinion about what constitutes progress of science. A distinction is made, for example, between radical and incremental innovation. Incremental innovation refers to cumulative advanced based on current S&T, while radical innovation is about completely new ideas that upend prevailing wisdom. However, these boundaries are fluid, and terms such as "transformative" and "potentially transformative", "paradigm shifting, "breakthrough", "pioneering," "high-risk high-reward," "creative," and "innovative," have been used interchangeably to describe the former. This section discusses definitions of some of the most commonly used terms in the literature, starting with the best-studied, creative research.

2.1.1 Creative Research

Of the synonyms identified above, creativity is the most commonly used, though no single authoritative definition or description of creativity exists. Simonton (1997) defined creativity as "the output of ideas that are both original and adaptive." Alternatively, Ochse (1990) incorporated the concepts of utility and originality into his definition of creativity, and included the production of an object or idea in his definition. Finally, Amabile et al. (1996) expanded upon both of these definitions by asserting that creativity involves heuristic (encouraging discovery of solutions) rather than algorithmic tasks or thinking. It would be a fair to say that creativity is generally defined as the capability of human beings to do things

that are novel, original, and valuable (Amabile, 1996; Sternberg, 2003). The literature on creativity also points to certain behavioral traits that distinguish creative individuals from their peers, such as a high level of curiosity, willingness to learn from experience, preparedness to take risks, persistence in situations of failure, high levels of energy, and distinctive goal-orientation. As both a result and a precondition of these traits, creative people typically tolerate contradictions, ambiguities, and uncertainties in their work (Sternberg 1997; Weinert, 2000).

Sternberg (1990) had more of an institutional approach to the concept in that he stated that creativity is community-specific, and acknowledges that the characteristics of creativity in the humanities may differ from the sciences. Simonton (2003) concurred, asserting that input from the respective communities are necessary to develop the definition of creativity in the particular community. In the context of scientific research, Heinze has developed a typology of creative research outcomes (Heinze, 2008). Table 2.1 reproduces this typology.

Table 2.1

Typology of the Outcomes of Creative Research

Type of scientific research creativity	Example
Formulation of new ideas (or set of new ideas) that open up a new cognitive frame that brings theoretical claims to a new level of sophistication	Theory of specific relativity (Einstein, 1905)
Discovery of new empirical phenomena that stimulated new theorizing	Biodiversity, led to the development of the theory of evolution (Darwin, 1859)
Development of a new methodology, by means of which theoretical problems could be empirically tested	Factor analysis, led to advancements in theory on mental abilities (Spearman, 1904 and 1927)
Invention of novel instruments that opened up new search perspectives and research domains	Scanning tunneling microscopy, led to advancements in nanotechnology (Binnig and Rohrer, 1982)
New synthesis of formerly dispersed existing ideas into general theoretical laws enabling analyses of diverse phenomena within a common cognitive frame	General systems theory (Bertalanffy, 1949; Ashby, 1956; Luhmann, 1984)

Source. Heinze (2008)

2.1.2 Transformative Research

The sociology of science literature does not define the term "transformative research." This is not surprising given that the term may have been introduced by the policy leaders of the S&T community. As discussed in Chapter 1, the National Science Board (NSB) "butterfly" report introduced the term "transformative research," defining it as follows:

> Transformative research is defined as research driven by ideas that have the potential to radically change our understanding of an important existing scientific or engineering concept or leading to the creation of a new paradigm or field of science or engineering. Such research is also characterized by its challenge to current understanding or its pathway to new frontiers. (p. 10)

The report also stated that transformative research "has the capacity to revolutionize existing fields, create new subfields, cause paradigm shifts, support discovery, and lead to radically new technologies" (p. 1). The NSB report went further and linked the concepts of transformative research and risk. It asserted that high-risk, high-impact projects do not "fare well wherever a review system is dominated by experts highly invested in current paradigms" and "during times of especially limited budgets that promote aversion to risk" (p. 4). The report stated further that "by its very nature, transformative research often is challenging to and frequently crosses disciplines" (p. 4). Last but not least, the report introduced the concept of "potentially" transformative research, because more often than not, federal agencies have to fund such research before it is established as transformative. The term "potentially" qualifies the term "transformative."

In a follow-on report entitled *Facilitating Transformative and Interdisciplinary Research (FacTIR)*, the National Science Foundation (NSF) built on the NSB definition (NSF, 2009):

> Transformative research involves ideas, discoveries, or tools that radically change our understanding of an important existing scientific or engineering concept or educational practice or leads to the creation of a new paradigm or field of science, engineering, or education. Such research challenges current understanding or provides pathways to new frontiers. Transformative research results often do not fit within established models or theories and may initially be unexpected or difficult to interpret; their transformative nature and utility might not be recognized until years later. Characteristics of transformative research are that it: Challenges conventional wisdom, Leads to unexpected insights that enable new techniques or methodologies, or Redefines the boundaries of science, engineering, or education. (p. 3)

In a notice to university presidents informing them of the requirement to incorporate the criterion for potentially transformative research in all NSF solicitations, then NSF Director Ardent Bement clarified the term further:

> The term "transformative research" is being used to describe a range of endeavors which promise extraordinary outcomes, such as: revolutionizing entire disciplines; creating entirely new fields; or disrupting accepted theories and perspectives -- in other words, those endeavors which have the potential to change the way we address challenges in science, engineering, and innovation. (NSF, 2007, p. 2)

2.1.3 High-Impact Research

Typically, studies that discuss the impact of science focus on economic impact (Godin & Doré, 2005). While economic impact is an important and non-negligible factor, impact has other dimensions, namely scientific impact, as well as social, organizational, and cultural impact. Godin and Dore operationalized scientific impact as

- The diffusion and appropriation of new knowledge, theories, methodologies, models, and facts;
- The formation and development of specialties and disciplines; and
- The development of interdisciplinary, intersectoral, and international research.

Many of these impacts can be quantified through both simple and sophisticated analysis of publication, citation, and patent counts (Cozzens 1996; NRC 1999, 2011; Van Raan 2003). These are discussed in Section 2.2.

2.1.4 Innovative Research

The term "innovation" has many different definitions, and Rose et al. (2010) summarize the best-known ones (Table 2.2). Innovation is generally seen as referring to research-based technological developments and their commercial applications. Scientific innovations refer specifically to those that advance the progress of science, and provide new insights that reshape the foundations of research. In the literature on innovation in science, "innovative" has traditionally been defined as being related to, but distinct from, creativity (described above). Amabile et al. (1996) defined "innovation as the successful implementation of creative ideas within an organization" (p. 1154). According to the author, researchers often see innovation as the usage or diffusion of creative ideas, and innovation and creativity are related concepts:

> All innovation begins with creative ideas… we define innovation as the successful implementation of creative ideas within an organization. In this view, creativity by individuals and teams is a starting point for innovation; the first is necessary but not sufficient condition for the second (p. 1180)

Table 2.2

Innovation Definitions

Innovation is "the commercial or industrial application of something new—a new product, process or method of production; a new market or sources of supply; a new form of commercial business or financial organization." **Schumpeter, *Theory of Economic Development***
Innovation is the intersection of invention and insight, leading to the creation of social and economic value. **Council on Competitiveness, *Innovate America, National Innovation Initiative Report*, 2004**
Innovation covers a wide range of activities to improve firm performance, including the implementation of a new or significantly improved product, service, distribution process, manufacturing process, marketing method or organizational method. **European Commission, *Innobarometer 2004***
Innovation—the blend of invention, insight and entrepreneurship that launches growth industries, generates new value and creates high value jobs. **The Business Council of New York State, Inc., *Ahead of the Curve,* 2006**
The design, invention, development and/or implementation of new or altered products, services, processes, systems, organizational models for the purpose of creating new value for customers and financial returns for the firm. **Committee, Department of Commerce, Federal Register Notice, *Measuring Innovation in the 21st Century Economy Advisory,* 2007**
An innovation is the implementation of a new or significantly improved product (good or service), or process, a new marketing method, or a new organizational method in business practices, workplace organization or external relations. Innovation activities are all scientific, technological, organizational, financial and commercial steps which actually, or are intended to, lead to the implementation of innovations. **OECD, *Oslo Manual*, 3rd Edition, 2005**
Innovation success is the degree to which value is created for customers through enterprises that transform new knowledge and technologies into profitable products and services for national and global markets. A high rate of innovation in turn contributes to more market creation, economic growth, job creation, wealth and a higher standard of living. **21st Century Working Group, *National Innovation Initiative*, 2004**

Source. Rose et al. (2010)

2.1.5 Paradigm-Shifting Research

The term paradigm has multiple definitions (Giere, 2006; Masterman, 1970), but there seems to be general agreement that a paradigm is "the set of beliefs, norms, and values

shared by members of a group of scientists engaged in studying specific problems in a research area" (Crane, 1980, p. 23). These beliefs, norms, and values can refer to the laws of nature, explanatory models, theories, and scientific predictions that need to be answered and the technical problem solutions that guide the research of scientists (Crane, 1980). In the literature that examines Kuhn's (1962) paradigm theory by means of examples from the history of science, a paradigm is usually a dominant theory in a particular field. Harman and Dietrich (2008) defined a paradigm as a scientific icon— "the central, explicitly stated organizing assumption in a given discipline, an assumption . . . held and practiced by a substantial majority of researchers." (Harman &Dietrich, 2008, p 9)

According to Kuhn's (1962) paradigm theory, the development of science builds on normal science and scientific revolutions. During normal science, research takes place within a given paradigm, with scientists making rather small changes to the dominant theory (law, model, etc.) in their domain. When too many problems and deviations from the theory (falsifiers or anomalies) are identified, a scientific crisis ensues, and eventually, the older paradigm will be replaced by a new paradigm (Werner, 2010). During this crisis, new ideas, perhaps ones previously discarded, are tried. Eventually a new paradigm is formed, which gains its own new followers, and an intellectual "battle" takes place between the followers of the new paradigm and the hold-outs of the old paradigm. After a given discipline has changed from one paradigm to another, this is called, in Kuhn's terminology, a scientific revolution or a paradigm shift.

A well-known example of a paradigm shift is the acceptance of Darwin's theory of natural selection as a replacement of Lamarckism as the mechanism for evolution. Another is the development of quantum mechanics, which supplanted classical mechanics.

2.1.6 Breakthrough Research

An Academy of Finland report (Häyrynen, 2007) used the term "breakthrough research," which it defined as "scientific research that promises significant results and that may be regarded as exceptionally creative, but that at the same time involves pronounced uncertainties" (p. 16). The definition appears to be an amalgam of many of the ones in the previous subsections.

2.2 Key Characteristics

There is no standardized way to characterize transformative research. Dietz & Rogers introduced a set of metaphors to inform key aspects of transformative research. Their first metaphor is that of a stock portfolio. This metaphor purports to illustrate "the decision to pursue certain projects in science by comparing them with the decision to invest in stocks in the stock market" (Dietz & Rogers, 2012, p. 6). The second metaphor is the process of evolution in which the "creative side of science is likened to biological variation and the support of such is likened to natural selection" (p. 6). Their third metaphor relates to the emergence of trends in popular culture, "which cover the whole range of social activities from clothing fashion to music and art" (p. 7). In the authors' vision, the tension between similarity and difference in this process is compared to the degree of newness that can be achieved in transformative research. The fourth metaphor is that of the exploration of a frontier, "where venturing into untouched parts of the world is compared to scientific exploration of new phenomena. The authors believe that each of these metaphors has its strengths and weaknesses in illustrating the emergence of new science, and do not expect any single one to be a complete portrayal of the sum total of what transformative research is.

Indeed based both on the preceding sections, and the Dietz & Rogers metaphor, all descriptions of transformative research point to three common features: it provokes greater

skepticism and disagreement as compared with normal (in a Kuhnian sense) science, it is easier to be identified in retrospect rather than ahead of time, and has a higher risk of failure. Each is discussed in turn below.

2.2.1 Elicit Skepticism

Transformative research is typically confronted by greater skepticism, which is an appropriate response to a jarring shift in paradigms. Extending Bayes Theorem, the more dramatically a new finding contradicts prevailing wisdom, the more likely that it is incorrect (Heinze, 2008). Some of this skepticism is rooted in the community's attachments to the prevailing paradigm, which impairs the ability of a peer to understand and use new ideas (e.g., the work of Avery and colleagues that DNA, not protein, carried biological information; Froderman & Holbrook, 2012).

A well-known example is quantum theory, in which despite many striking confirmations, so strange was Planck's idea that it took 11 years for quantum theory to gain final acceptance by leading physicists' (Polanyi, 1966). Similarly, the path-breaking economic theory of asymmetric information and adverse selection (which ultimately garnered the proposer George Akerlof a Nobel Prize) was initially rejected by three major economics journals (Akerlof, 1994). In a 2011 editorial in the British newspaper The Guardian, science columnist Phillip Ball speculates:

> The kind of idle pastime that might amuse physicists is to imagine drafting Einstein's grant applications in 1905. "I propose to investigate the idea that light travels in little bits," one might say. "I will explore the possibility that time slows down as things speed up," goes another. Imagine what comments these would have elicited from reviewers for the German Science Funding Agency, had such a thing existed. Instead, Einstein just did the work anyway while drawing his wages as a technical expert third-class at the Bern patent office. And that is how he invented quantum physics and relativity. (Ball, 2011)

Another example was showcased in a report published by NSB (2007). The report described Charles Townes, who in the 1940s was asked by Bell Labs to help develop radar bomb-aiming systems as part of the U.S. war effort. Similar war work had been conducted at the nearby Columbia University in New York, so when Bell Labs suggested that Townes focus his work on interests that were more relevant to Bell Labs but those that he did not share, he decided to accept an appointment at Columbia. In his work, he conceived a possible method to generate photon 'avalanches' using excited ammonium molecules. But he hit a roadblock, as reported by NSB:

> (After) we had been at it for two years, Rabi and Kusch, the former and current chairman of the department—both of them Nobel laureates for work with atomic and molecular beams, and both with a lot of weight behind their opinions—came into my office and sat down. They were worried. Their research depended on the support from the same source as did mine. 'Look,' they said, 'you should stop the work you are doing. It isn't going to work. You know it's not going to work. We know it's not going to work. You're just wasting money. Just stop! (p. 5)

But Townes, showing extraordinary nerve, stood his ground and "respectfully told his exalted colleagues that he would continue" (p. 5). Two months later, his experiment worked, and the maser (microwave amplification by stimulated emission of radiation) was born. Three years after that, Arthur Schawlow, Townes' postdoctorate fellow at Columbia University, developed the optical version of the maser –the laser. Townes was awarded the Nobel Prize for Physics in 1964 (NSB, 2007). Lasers are now the basis of multibillion dollar industries in all sectors of society including consumer electronics, information technology, medicine, law enforcement, entertainment, and the military.

A more recent example of the skepticism construct is the 2011 Chemistry Nobelist Dan Schectman, whose discovery of quasicrystals revealed a new principle for packing of atoms and molecules. He was conducting routine research at the U.S. National Institute of Standards and Technology when he made his initial discovery:

> Scientists greeted Shechtman's discovery with resistance, even ridicule... The head of his laboratory [at NIST] suggested he read a textbook on crystallography. When Shechtman persisted in his experiments, he was asked to leave the research group. Shechtman persevered in the face of doubt and ridicule in describing a form of crystal whose patterns are regular but never repeat, a notion that shattered scientists' belief that all crystals consist of recurring patterns. (Gerlin, 2011)

At other times paradigm shifting result is resisted because they are "premature," their "plausibility foundering on expert knowledge that renders them unlikely or unproductive, their acceptance awaiting ideas or findings that fill in details, propose workable mechanisms, shed doubt on competing explanations" (such as Wegener's theory of continental drift; Froderman & Holbrook, 2012).

Dietz & Rogers (2012) summarized this skepticism by noting: "What makes [transformative research] conceptually distinct is both its epistemic and phenomenological nature as well as its practical consequences. It is transformative research precisely *because* it cannot "get along" within the existing system" (p. 20).

Even if there isn't skepticism, there is often little agreement, even after-the-fact, let alone before, regarding what transformative research is. Some of the disagreement is simply around the definition of transformative research. For some in the community, like the NSB and NSF definitions imply, transformation refers to stunning intuitive leaps into new, uncharted scientific territory. For others, less creative application-oriented research such as the development of the polio vaccine is more transformative. For the latter group, it was disappointing that the researchers that discovered tissue culture methods used to grow poliovirus were honored with the Nobel prize, but not those (like Salk) who translated this method into successful vaccination strategies (Johnston, 2008).

Similarly, an editorial in the Annals of Neurology (Johnston et al., 2008) described an experiment the journal conducted to get its editorial board to select ten papers in neuroscience that were transformative. The editors commented on how difficult it was to get the editorial staff to agree on the ten papers, and even when they finally did, they felt that they had missed fifty others that were equally transformative.

2.2.2 Identified in Retrospect

It is often believed that transformative research cannot be predicted, is a result of serendipity, and can only be identified in retrospect (Stoskopf, 2005). Even for areas where there is eventually agreement, it may sometimes take decades for transformative research to be recognized as such. An example is that of Barbara McClintock, whose research in cytogenetics, conducted in the late 1940s, was not recognized as groundbreaking till the 1960s (with the Nobel Prize coming only in the 1980s).

It is not only skepticism that delays this recognition – some ideas are simply ahead of their time. Indeed, some publications are considered "sleeping beauties," implying that it is a publication that goes unnoticed (sleeps) for a long time and then, almost suddenly, attracts a lot of attention, and is awakened by a prince (Van Raan, 2004). This is also called the Mendel syndrome, named after Gregor Mendel whose discoveries in plant genetics were so unprecedented that it took 34 years for the scientific community to catch up to them.

2.2.3 Risky

The NSB (2007) butterfly report explicitly linked the concepts of transformative research and risk. In its definition of transformative research, the Board emphasized that to achieve scientific breakthroughs, risk is inevitable. And risk, it claimed, always entails the distinct possibility of failure, which can be measured at many levels: "At the individual level, a scientist might fail to prove a potentially transformative hypothesis at the laboratory bench;

at the enterprise level, taxpayers might fail to get their money's worth." The Academy of Finland report (Hayrynen, 2007) concurred: "breakthrough research is characterised not only by exceptional innovation, but also by the conscious taking of risks in the choice of its research subjects and methods and by its ambitious goal-setting" (p. 11).

Given the central role of risk in transformative research, it may be worthwhile defining the concept of risk, and examining how it has been operationalized in the domain of science policy. The definition of risk has been fraught with confusion and controversy, and communities define it differently (Holton, 2004). Classical economic theory has typically defined risk as the variance in the probability distribution of possible gains and losses from an investment (Arrow, 1965; Pratt, 1964). A risky alternative is one for which the variance is large; and risk is one of the attributes which, along with the expected value of the alternative, are used in evaluating alternative gambles. The International Organization for Standardization (ISO, 2009) has defined risk simply as the effect of uncertainty on objectives. Holton (2004) similarly defined risk as "exposure to a proposition of which one is uncertain" (p. 19). Indeed the presence of uncertainty happens to be the one common thread across all definitions of risk in the risk-management literature.

Hansson (2011) summarized five technical definitions of risk. The first refers to "an *unwanted event* which may or may not occur." The example provided is that of lung cancer, which is considered a risk taken by smokers. The second refers to the "*cause* of an unwanted event which may or may not occur." The example is provided is that of smoking itself being a health risk. Both are qualitative definitions.

Hansson also listed quantitative definitions, one of which is the "*probability* of an unwanted event which may or may not occur." Example of this usage is as follows: "The risk that a smoker's life is shortened by a smoking-related disease is about 50%." Another

expression of risk is the "statistical *expectation value* of an unwanted event which may or may not occur." The last definition of risk is within the context where a decision is made under conditions of *known probabilities* ("decision under risk" as opposed to "decision under uncertainty").

Definitions of risk are typically customized to the domains in which they are being applied. In the literature on finance, for example, risk is typically defined as the probability that an investment's actual return will be different than expected (Artzner, 1999; Case, 2007; March, 1987; Felli, 2011; Fischhoff, 1984; LeRoy 1987; Petsco, 2008; Renn, 1998; Wageman, 2004). In scientific circles, risky research is defined as new ideas 'where one is unsure whether it will work." (Laudel, 2006, p. 9). Partially because it seeks to fund high-risk research, the term has been defined by at least one federal funding agency. The high risk high reward (HRHR) Demonstration Oversight Group at the NIH defined high-risk high-reward research as "research with an inherent high degree of uncertainty and the capability to produce a major impact on important problems in biomedical/behavioral research" (Austin, 2009, p. 9).

2.3 Operationalizing Transformative Research

Despite multiple definitions of transformative research, there has been little effort to operationalize the definitions provided above. *"I know it when I see it"* is the dominant paradigm most experts choose to identify transformative research. Going beyond the definitions, which tend to be sublime - inspirational for certain, but also vague and flowery, there is no operational understanding of what transformative research is. In a 2008 editorial in the Annals of Neurology, the editors wrote:

> We have almost no information about what predicts transformation. Who are these people who go on to produce transformative studies and win prizes like the Nobel and the Lasker? Are they particularly ambitious, hard-working, smart, creative, or just

> lucky? Are they triple threats, or do they focus tightly on the mission at hand? Similarly, do we have any hope of identifying transformative projects in advance or do they really arise from good fortune, hard work, and resourcefulness? How important is environment? Do these discoveries come from working in isolation or from applying advances in other areas to a whole new problem? These are all key questions if we want to accelerate discovery, and should be the focus of university and department administrators, as well as funders. It seems particularly odd that the predictors of transformative research are completely unstudied. (Johnston & Hauser, 2008, p. A11)

There are no known studies that relate the production of transformative research to independent variables, whether retrospective or prospective. In this section, I explore several *proxies* of transformative research.

2.3.1 Publications-Based Indicators

Productivity. Van Raan (2003) noted that "communication, i.e., the exchange of research results, is the driving force in science" and explained that publications as direct research outputs are critical to understanding the diffusion of knowledge (p. 1). The number of publications per year is an important indicator of the level of activity or productivity of a researcher in a particular field. Simonton (2004) argued that creative scientists are more productive than their peers, but that they also publish a higher number of ignored works. He posited that scientific creativity is a "probabilistic consequence" of research quantity, and the likelihood of a researcher's peers finding his or her work creative is a probabilistic consequence of quantity (p. 24). Under this model, the total number of publications could be used as an indicator for transformative research (Simonton, 2003). Heinze & Bauer (2007), in their research on measuring creativity, tested this hypothesis - that the total number of publications could be used as an indicator for creative research - and found that the total number of publications did appear to be a significant predictor of creativity. They acknowledged, however, that this operationalization is overly simplistic, and suggested that

other indicators, such as measures of impact (see below for discussion), be preferred (Heinze & Bauer, 2007).

Impact. *P*eer recognition is a core aspect of producing breakthrough research. In addition to productivity, experts have found impact (definition discussed in Section 2.1.3 above) on the community to be an indicator of creativity. Most commonly, impact is operationalized using citation rates (Azoulay 2009; Heinze & Bauer, 2007). The number of citations to an article indicates a measure of association between the cited and citing articles and that the number of citations can be indicative of an article's scientific impact.

There are many concerns about the use of citations for understanding scientific impact, The unifying factor in these limitations is the lack of discrimination for the nature of the citation. An article receives one citation count regardless of whether the citation was positive, negative, scientific, nonscientific, or whether it was an author's self-citation. Nevertheless, it might be counter-argued that any citation – positive or negative – is indicative of an impact on the research community. However, for lack of better measures, citation counts are generally accepted in the scientific community (Bornmann & Daniel, 2008).

An alternative to citation counts is the h-index, a bibliometric indicator that has risen in popularity since its introduction in 2005 (Hirsch, 2005). The h-index is a combined measure of the productivity and impact of a scientist. A scientist has an h-index of h if his/her published papers Np have at least h citations each, and the other papers (Np – h) have no more than h citations each. The h-index captures the career-long achievements of a researcher in the sense that it is insensitive to un-cited or lowly cited papers as well as to one or several highly cited papers. The h-index draws upon the idea that knowledge diffusion is dependent upon and integrates both publication and citation counts into one metric.

The h-index was validated as a metric for research quality and impact when van Raan (2005) concluded that the h-index is positively correlated with the peer judgments for articles published by 147 chemistry research groups in the Netherlands. The h-index has been calculated for publication sets affiliated with a single researcher's career, the careers of a group of researchers, a specific research field, a single journal or group of journals, and a country or group of countries (Alonso et al., 2009; Hirsch, 2005). Variations on the h-index attempt to render impact more comparable across researchers by accounting for differences in the length of a researcher's career (Burrell, 2007) and differences in citation patterns from diverse scientific fields (Iglesias & Pecharroman, 2007).

Other metrics such as mean normalized citation score (MNCS) and burst index are derived from citations, and attempt to improve upon limitations of simple citation counts for researcher-level analyses. The MNCS normalizes citation counts for differences among research fields by using a ratio of average international citation rates of all publications in a particular field of interest (Lundberg, 2007). The MCNS has been used in the U.S., Canada, and Spain, and is useful for making comparisons of scientific impact across fields. Most recently, the Center for Science and Technology Studies at Leiden University is moving to adopt the MNCS as an alternative to their own crown indicator, another well-known indicator of citation-related research performance (van Raan 2004; Waltman, van Eck, van Leeuwen, Visser, & van Raan 2010).

Researchers at Thomson Reuters have developed the Breakthrough Paper (BP) indicator, in which a publication is identified as breakthrough based on whether its citations are in the top 0.1% of highly cited publications in the particular field of interest (Ponomarev, Williams, Schnell, & Haak, 2011). The BP indicator is dependent on the number of publications in the research field and should only be used to compare publications in similar

research fields. Ponomarev et al. (2011) showed that a 1-2 year time frame after an article is published provides sufficient time to identify key citation patterns and a majority of breakthrough papers.

Journal Impact Factor. The concept of the journal impact factor is derived from citation analysis. When citations for articles within a journal over a period of time are compiled, the resulting citation count provides an objective measure of the quality of the journal to communicate research results (Garfield, 1979). The number of citations a journal receives per articles published provides an indication of the journal's impact on the scientific community (Garfield, 1972). An article published in a journal that is prestigious, and thus more often cited, is more likely to have a greater impact than an article published in a less prestigious journal. The common assumption is that great researchers tend to continually deliver great work and, along a similar vein, an article published in a high impact journal is also likely to be of high quality itself. Bibliometric researchers and companies have developed several journal impact indicators, including the Journal Impact Factor (JIF) developed by Thomson Reuters, and the SCImago Journal Rank (SJR) developed by the SCImago research group (Garfield 2005; Moed 2010). The JIF is calculated by dividing the number of citations in the current year to all items published in a journal the previous two years with the number of publications in the journal during the previous two years. The SJR not only considers the citation counts but also the quality of the citation measured by the prestige of the journal from which the citation is coming. The importance of the journal is thereby measured by the importance of the citations they receive.

The indicator has been criticized, although most criticisms center more on the lack of usefulness of citation metrics in general, rather than journal impact factors in particular. There are also concerns that impact factors conceal the difference in article citation rates, are

unrelated to the quality of the articles, and vary depending on the research field (Seglen 1997). However, other studies have demonstrated a correlation between high impact factors and journal quality (Saha, Saint, & Christakis, 2003).

2.3.2 Indicators Used in Recent Evaluations of Transformative Research

To further explore the presence of indicators of transformative research, a review was conducted of evaluations of transformative research. Given the novelty of the concept, not many transformative research programs exist (this is discussed in Chapter 3 below), and there are even fewer evaluations. However, three studies were found, and Table 2.3 lists the indicators used.

Table 2.3

Measures Used in Recent Evaluations of Transformative Programs

Authors	Year	Program/Group Evaluated	Examples of Measures Used
Pion and Cordray	2008	Burroughs Wellcome Career Award in the Biomedical Sciences (CABS)	Faculty in top-25 ranked institutions in NIH funds PI on NIH R01 or other NIH grant Age at first R01 Total number of Publications Publications in top-ranked journals Average citations per article
Heinze and Bauer	2007	Nominated Creative Scientists in Nanotechnology	Overall productivity Citation rates Degree Centrality Integration Score
Azoulay, Zivin, and Manso	2009	Howard Hughes Medical Investigators	Publications in top percentiles of citations Nobel Prizes won Elected to NAS/IOM Trained an early career award winner Novel keywords tagging publication

Awards. Some evaluative research has used awards to supplement publication-based measures as indicators of innovativeness. For instance, Azoulay et al. (2009) studied elections to prestigious scientific societies as an indicator of creativity. Although not often used in applied work, Simonton (2003) suggested a comparison of honors and awards listed on

researchers' curricula vitae to measure creative impacts. Simonton identified four categories of awards: (a) international recognitions, such as the Nobel prize, (b) national recognitions, such as election to the National Academy of the Sciences, (c) discipline-specific honors, such as the Distinguished Scientific Contribution Award of the American Psychological Association, and (d) society-level recognitions, such as 'fellow' status within a society (Simonton 2003).

Lab-level indicators. Another non-publication-based indicator of innovativeness is the number of students and fellows trained at a researcher's lab that goes on to win prestigious scholarships (Azoulay et al., 2009). This indicator is hypothesized to be an indicator of innovativeness, since more innovative labs can attract higher quality students. This indicator, however, has questionable measurement validity because the trainees' achievement may be caused by both the lab head's creativity and/or the students' own creativity. This caveat is acknowledged by Azoulay et al. (2009).

Patents. The last indicator of creativity used in the evaluations above is patent activity. The two most common measures used are patent counts and patent citations. Previous studies have found that patents are granted to the most productive individuals in research (Stephan, Gurmu et al., 2007) and that patents are preceded by a flurry of scientific publications (Azoulay et al., 2006). Patents, however, are typically only pursued in situations in which the research may have commercial value, and thus may not have the same interpretation across all fields of research.

2.3.3 Emerging Indicators

The prediction of the potential value, or the impact, of an idea can be made computationally in terms of the degree of "structural change" introduced by the idea (Chen, 2011). The foundational principle here is that if a new idea connects previously disparate

patches of knowledge, then its transformative potential is higher than the potential of ideas that are limited to well-trodden paths over the existing structure. The intellectual structure can be represented by networks of ideas. "The theoretical underpinning of the structural variation is that scientific discoveries, at least a subset of them, can be explained in terms of boundary spanning, brokerage, and synthesis mechanisms in an intellectual space" (Chen et al., 2009, p. 438).

Burt (2004) argued that people who live at the intersection of networks are more likely to be familiar with selecting and synthesizing alternatives into novel ideas. This hypothesis implies that scientists who connect homogeneous groups, such as disciplinary communities or research fields, have a higher probability of exposure to alternative ways of thinking and behaving. There are several emerging ways to measure this integration.

Brokerage. Within structural theory, brokerage is an indicator of whether research could be transformative in its degree of brokerage. Brokerage is defined as a measure of an individual's connections to other scientists. The theory is that people with more connections to distinct social networks are considered brokers, and hypothesized to have more innovative outputs since they are exposed to more diverse ideas (Burt, 2004). Heinze and Bauer (2007) considered this theory by examining the association between the number of disparate authors and groups brokered by a researcher and his or her citation rate. Their longitudinal study of highly creative scientists in Nano science and technology found that "it is not only the sheer quantity of publications that enables scientists to produce creative work but also their ability to effectively communicate with otherwise disconnected peers and to address a broader work spectrum" (p. 114). This finding suggests that brokerage promotes creation of transformational discovery.

Degree Centrality. Degree centrality refers to the size of a researcher's network. This variable has been operationalized by Heinze & Bauer (2007) as the size of a scientist's coauthor ship network in any 3-year time period. They found that degree centrality does not predict a creative event, but does correlate with the number of author-level citations (Heinze & Bauer, 2007). Chen (2012) claimed

> a paper with a high betweenness centrality is potentially a transformative discovery. In addition, it would be possible to use this metric to identify potential future discoveries by calculating the would-be betweenness centrality of a hypothetical connection between two disparate areas of existing knowledge networks....Thus, betweenness centrality can be translated into interestingness, which can be in turn translated into actionability. (p. 12)

Multidisciplinarity/Interdisciplinarity. If a new idea connects previously disparate patches of knowledge, then its transformative potential is higher than the potential of ideas that are limited to well-trodden paths over the existing structure. A transformative discovery is therefore more likely to be made when a novel connection is established between two or more previously disparate units of scientific knowledge. One expression of this connection is through interdisciplinarity[2] (Heinze et al., 2007; Wagner, 2010).

While interdisciplinarity has many meanings in scholarship, it can be defined as the integration of traditional disciplines of knowledge into newly synthesized fields or niche areas of research within an existing field.

---

[2] While there is little clarification in the difference between multidisciplinarity and interdisciplinarity, in general, they both refer to the number of disparate bodies of specialized knowledge (Porter, Cohen et al., 2007; Wagner, Roessner et al., 2010).

The National Research Council has weighed in, with the report *Facilitating Interdisciplinary Research* (NRC 2004). Federal research agencies are increasing their support of interdisciplinary support. The NIH Roadmap includes initiatives that recognize the importance of interdisciplinary research for future advances in biomedical research. NSF's Nanoscale Interdisciplinary Research Teams program requires each funded project to have at least three co-PIs. National Academies' Keck Futures Initiative is a 15-year, $40 M project to foster interdisciplinary research.

Operationalizing interdisciplinarity has been complex and there are no methods with widespread support. Some experts argue that collaboration between theoretically inclined and applied researchers in the same field should be considered interdisciplinary because of difficulty of integration engendered by cultural differences. Differences in conceptual foundations and methods within a "field" may indeed be extreme. In a presentation to the National Academies, interdisciplinarity expert David Roessner (2007) asked the audience to "consider the challenges in having two psychologists --a Freudian clinical psychologist working with a neuroscientist –collaborate."

Researchers have attempted to operationalize the concept, and developed metrics such as the integration score, specialization score, and diffusion score, to assess how scientific disciplines utilize knowledge from other disciplines. The integration score measures the integration of knowledge within a body of research based on the publication's cited references and is a backward-looking metric (Porter, Cohen, Roessner, & Perreault 2007; van Raan 2002). The specialization score measures the degree of diversity of topics within a body of research (Leahey, 2007; Porter et al., 2007). The diffusion score measures the influence of a body of research based on the citations to a publication and is a forward-looking metric (Carley & Porter, 2011). While there are other measures of diversity, Rafols and Meyer

(2010) noted that the integration score is the only measure that integrates the three aspects of diversity - variety, balance, and similarity - into one index.

Researchers also point out (Porter, 2007) that integration and specialization are related concepts: integration (I) measures the extent of diversity among a paper's cited subject categories, and specialization (S) measures the spread of SCs in which the body of research is published. A true Pioneer would not only integrate knowledge from many fields, but also publish in many fields.

Niche. In a 1996 paper, Podolny et al. (1996) introduced the conception of an organization-specific niche in a technological network. According to the authors, a niche is defined by two mutually opposing properties – *status* (which refers to characterization of an actor's position in a social network), and *crowding* (which refers to high similarity in technological competencies with others). For any individual or firm, crowding depresses growth rates and status elevates them. Furthermore, according to the authors, mapping status and crowding in two-dimensional matrix leads to the recognition of four types of roles:

- Organizations in the top left quadrant can be considered **brokers** of new technologies. They build on previously unexploited technology and provide a distinctive foundation for others.
- Organizations in the top right quadrant are **leaders** in well-established technologies.
- Organizations in the bottom right quadrant can be considered **followers**; they are engaging in innovative activity in congested regions of the technological space.
- Organizations in the lower left quadrant were referred to as **isolates**, developers of technologies that are not endorsed by other organizations.

The concepts can be transplanted into the domain of firm networks (as was done by Okamura & Vonortas 2009), or science policy, replacing status indicators with indicators of researcher quality and crowding indicators with those of the number of researchers

publishing in the same field of research. In the case of science, high *status* allows the researchers to undertake high-risk activities, and *crowding* can guide the observer to the degree of emergence of a field of research.

2.3.4 Consensual Assessment

Consensual assessment technique is a method that relies on the subjective judgments of appropriate observers, often experts, to determine whether a research product is creative (Amabile, 1982). The Committee on Science, Engineering, and Public Policy, a joint committee across branches of the National Academies,[3] corroborated Amabile's conclusions in identifying expert review as the most effective means of evaluating federally-funded research programs in comparison to economic-impact studies, and bibliometric analyses on publication and patent data (NAS, 1999). The committee discussed three forms of expert review which could be valuable: (a) quality review, which judges the quality of the scientific research; (b) relevance review, which judges the relevance of the research to the agency's mission; and (c) benchmarking review, which judges the international leadership status of the United States in the context of a program (Sternberg, 1990).

Grant and Allen (1999) used a novel expert review approach to compare Wellcome Trust Showcase awards with a sample of standard project grants to assess the whether the grants were risky, novel, speculative, adventurous, and innovative on a 5-point scale. The evaluators selected 10 research summaries—five from each group from 40 summaries. These 10 summaries were then sent to the 48 members of the expert panels, yielding 12 reviews per

3 The National Academies Committee on Science, Engineering, and Public Policy is a joint unit of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine. See http://sites.nationalacademies.org/pga/cosepup/index.htm.

project. In fact, each summary was reviewed by an average of 7.7 panel members. The authors noted that this method eliminated much of the systematic error that might have occurred using standardized expert review, making the results more robust.

In 2008, a Working Group on Peer Review of the Advisory Committee to the Director of NIH recommended that NIH shorten the length of the application and "engage more persons to review each application…optimally 4 or more" (NIH, 2008, p. 13). The Working Group left the actual number of expert reviewers ambiguous, so Kaplan, Lacerta, and Kaplan (2008) conducted a statistical analysis to provide guidance on the optimal number of expert reviewers. They had 10 short proposals scored by an average of 48 reviewers. They then conducted a sensitivity analysis and found that "funding decisions will vary widely with the number of reviewers in considering proposals that are closely scored" (p. 6). They noted that the length of the application affects how many reviewers can be used for scoring; the shorter the application, the more reviewers that can be used. They conclude that the NIH peer review process should be designed to meet statistically significant criteria.

The recently completed NIH Director's Pioneer Award (NDPA) Outcome Evaluation successfully used expert review to determine whether and how the awardees' research was pioneering (Lal et al., 2011). The goals of the expert review were to assess the impact of the work conducted under the Pioneer Award, and to provide insight regarding the effects of the NDPA program. Awardees were asked to suggest 3 to 5 potential experts, with a mix of supporters and critics. They were also asked to review a one-page summary of their research conducted under the NDPA, which could be sent to the expert reviewers. Finally, the awardees were asked to suggest three publications for the expert review. The expert reviewers were sent these materials as well as the NDPA Program Notice from the year of the award, and a feedback form to record their assessment of the Pioneer project.

Following submission of the completed feedback form, selected experts were invited to share additional feedback during a phone interview, and to clarify answers where necessary. This expert review process was used successfully to identify whether or not the awardees' research was pioneering and why (or why not). The drawbacks were the time it took to prepare the review materials, code the assessment forms, and to conduct follow-up phone meetings.

2.4 Operationalizing High-Risk Research

Quantifying risk is important in business finance, and there is a wealth of literature on the topic, in particular portfolio theory. Early thinkers in the area proposed having standard deviation as a measure of riskiness (Markowitz 1959; Tobin, 1958). In a comparison between two investment alternatives, for example, the one whose outcome is calculated to have the largest standard deviation is regarded as the riskier one. Rothschild and Stiglitz (1970) introduced greater sophistication in this measure – moving "probability mass" from the center of a probability distribution to the tails, while keeping its mean unchanged, increasing the risk associated with the distribution.

The measure of risk has continued to become more sophisticated over time, and in the current incarnation, there are many versions of the concept, including:

- *Sharpe ratio*, which converts total returns to excess returns by subtracting the risk-free rate, and then divides that result by a common measure of dispersion, the standard deviation or sigma to get a measure of "reward per unit of risk." It is a dimensionless unit, and difficult to interpret (Dowd 2000; Sharpe 1966).

- *Jensen-Treynor formulation*, which measures risk-adjusted performance by comparing the excess returns of a "portfolio i with those of a market portfolio that is matched to the risk of portfolio i" (Jenssen 1968; Treynor 1966).
- *Risk Adjusted Performance (RAP)* measures the returns of the portfolio, adjusted for the deviation of the portfolio (which is what the authors took as a proxy for the portfolio's risk), relative to that of the market. The indicator is derived from the Sharpe Ratio above, but it has the advantage of being in units of percent return, which makes it dramatically more intuitive to interpret (Modigliani & Modigliani, 1997).

Using the "stock portfolio" metaphor for transformative research, Dietz and Rogers (2012) have brought further insights from the finance world into the science and technology (S&T) world. In comparing the decision to buy an individual stock to the decision of whether to fund a scientific project, the metaphor appears apt. Risk management is a well-studied concept in the economics of R&D (Dasgupta et al., 1987; Katz, 1986) and private firm R&D strategies (Floricel et al., 2008). It could be used in actively managing risk of Federal R&D – just like in the finance sector, research projects could be identified, considered, selected, and managed relative to one another.

Because science is seen as an inherently risky enterprise, the term "high risk" has never been operationalized as formally in federally-funded R&D circles. In a speech in 2003, Colwell (2003) described a classification of risk:

- First there are projects that are "conceptually" risky – those whose fundamental ideas are at odds with the prevailing wisdom. They may elicit review comments such as "implausible hypothesis."

- Then there are proposals that are "technically" risky. That is, they require equipment or techniques or approaches that have either not been tried or are assumed to be extraordinarily difficult.

- Third, there are ideas that are risky because the investigators are proposing to work outside their previously demonstrated areas of expertise.

- And finally, there is risk that derives from proposals that entail unprecedented combinations of disciplines, or have criteria for success that involve viewing the results from an unfamiliar multidisciplinary perspective. (p. 2-5)

When making funding decisions, Tekes, the Finnish Funding Agency for Technology and Innovation distinguishes between seven risk categories (Hayrynen 2007):

- Risk related to the *research objectives*, such as whether the objectives are realistic and attainable in the first place or whether failure is very likely.

- Risk related to the *research methods*, such as the use of an untried method, a dataset that is poorly fitted with the method or the wrong kind of research tools.

- Risk related to the *field of research*, such as the sense that the subject is too marginal or (in Finland) in an orphan situation, and on the other hand that the field is too crowded.

- Risk related to *personnel*, such as the lack of scientific merits or the anticipated weakness of the manager's role.

- *Ethical* risks related to the research, such as data protection issues.

- The risk connected with *interdisciplinarity*, i.e., weak links between researchers or participating projects representing different fields of science in interdisciplinary or multidisciplinary programmes.

- Risk related to *resources*, i.e., the research cannot be completed with the resources projected in the research plan or on timetable. (p.34-35)

2.5 Conclusion

In a 1964 Supreme Court case related to pornography in the motion picture industry, Justice Potter Stewart said:

> I shall not today attempt further to define the kinds of material I understand to be embraced within that shorthand description; and perhaps I could never succeed in intelligibly doing so. But I know it when I see it. (Supreme Court, 1964)

Transformative research is to the scientific community as pornography was to the U.S. Supreme Court—undefinable but nonetheless easily recognized; and yet, like the Justices of the Court, no two scientists share a common definition. In this chapter, I endeavored to go beyond the "I know it when I see it definition" and attempted to define transformative research. I found that the term is used interchangeably with many others, and while there is loftiness to these definitions, the terms have not been operationalized. In other words, no indicators of transformative research exist. So I examined proxies of transformative research, and sought to operationalize them. Based on a review of the literature in sociology of science, bibliometrics, consensual assessment techniques, and network theory, the operationalizations came in four flavors – publication-based indicators, non-publication-based indicators, emerging indicators, and expert reviews.

The next chapter presents findings from further inquiry into indicators of transformative research, this time by examining programs that support transformative research. The expectation was that any program that purports to support transformative research defines the term at least indirectly by articulating its goals and criteria for selection, and in general develops a theory of change that guides program design, implementation, and evaluation.

3.0 Review of Programs Funding Transformative Research

Despite the clear articulation of need, as the previous chapter indicated, there is a paucity of research on the topic of transformative research, and a lack of understanding of its theoretical underpinnings. I therefore turned to more empirical information – in particular within programs that fund transformative research. My intent was to discover practical heuristics that guide programs that purport to support transformative research to shed light on their implicit assumptions about the nature and characteristics of transformative research. The intent was neither to evaluate the programs nor to examine their implementation.

This chapter explores twenty programs in the U.S. federal government that are noted for supporting transformative research, and attempts to synthesize common features. In Section 3.1, I begin by describing how the programs were selected and the reviews conducted, and then present relevant details of the programs in Sections 3.3-3.9. In the final two sections, I propose a typology and a framework to characterize transformative research programs.

3.1 Methodology

A list of programs that fund transformative research was developed based on a systematic keyword driven search of US federal R&D programs. Given that there is no pre-existing list of transformative programs in the Federal government, I examined dozens of programs whose goals appeared to use terms such as "high risk high reward" research, pioneering research, game-changing research etc. I also searched the general grey literature using keywords listed in Table 3.2. The focus was on federal programs, not foundations (such as the Howard Hughes Medical Institute (HHMI) or Wellcome Trust, which fund transformative research in the U.S. and Britain respectively), nor on programs in other

nations (like the European Union's New and Emerging S&T (NEST) program), nor other mechanisms for promoting transformative research (e.g., prizes and grand challenges).

I found a total of twenty current programs that claim to fund transformative research. They are listed along with their parent agencies in Table 3.1. Once the programs were selected, in order to understand the mechanisms of funding high-risk, high-reward research in U.S. federal agencies, relevant information on the programs was collected. Aspects of the programs examined included:

- overview of the program, purpose of the award, and research area foci;
- formation of the program and how areas of interest were/are determined;
- funding data, including the size of the program, and size and length of the awards;
- selection process, including information on who is involved at each stage of the process;
- selection criteria and application materials required;
- aspects of the PI/research team considered when making funding decisions;
- aspects of the research problem considered when making funding decisions; and
- postaward support and/or management of the award by the program.

Data about each program were collected via the Internet and publicly available documents, where possible. In addition, a short email-based questionnaire was developed to address more subjective aspects of the programs that may not be available through program documents (see Appendix A for the protocol). Program contacts were identified and asked to both answer the questionnaire, and to check the information collected on their program for accuracy. Program officials were given the option to fill out the questionnaire over email or by phone. Respondents were especially probed to describe internal descriptions/definitions of terms such as *high-risk, high-reward, innovative, transformative*, and *pioneering*.

Table 3.1

U.S. Federal Government Research Programs Examined

Agency/Program	**Acronym**	**Department/Agency**
Advanced Research Projects Agency-E	ARPA-E	Department of Energy
Center Innovation Fund	CIF	National Aeronautics and Space Administration
Creative Research Awards for Transformative Interdisciplinary Ventures	CREATIV	National Science Foundation
Defense Advanced Research Projects Agency	DARPA	Department of Defense
Early-concept Grants for Exploratory Research	EAGER	National Science Foundation
Emerging Frontiers in Research and Innovation	EFRI	National Science Foundation
Exceptional, Unconventional Research Enabling Knowledge Acceleration	EUREKA	National Institutes of Health
Exploratory Advanced Research Program	EARP	Department of Transportation
Exploratory/Developmental Research Grant Award	R21	National Institutes of Health
Game Changing Development	GCD	National Aeronautics and Space Administration
Homeland Security Advanced Research Projects Agency	HSARPA	Department of Homeland Security
Innovative Advanced Concepts	NIAC	National Aeronautic and Space Administration
Intelligence Advanced Research Projects Activity	IARPA	Office of the Director of National Intelligence
National Security Science and Engineering Faculty Fellowship	NSSEFF	Department of Defense
New Innovator Award	NIA	National Institutes of Health
NIH Director's Pioneer Award	NDPA	National Institutes of Health
Outstanding New Environmental Scientists	ONES	National Institutes of Health
Sunshot Initiative	Sunshot	Department of Energy
Technology Innovation Program	TIP	National Institute of Standards and Technology
Transformative R01	TR01	National Institutes of Health

3.2 Overview of Programs

Each of the twenty programs is described in brief below, in alphabetical order by the name of the program. Sections 3.3 – 3.9 drill down into specific aspects of the programs.

The ***Advanced Research Projects Agency-Energy (ARPA-E)*** program within the Department of Energy (DOE) was introduced in legislation in 2007 but was not funded until the American Recovery and Reinvestment Act of 2009. ARPA-E is modeled after DARPA and is intended to support out of the box, transformational R&D that runs the range from basic research to commercialization, in areas that are too multidisciplinary or risky to fit within the traditional Department of Energy system. In its first two years of existence, ARPA-E has awarded amounts ranging from roughly $400,000 to $9 million each to 121 projects, with an average award value of $3 million.

The NASA ***Center Innovation Fund (CIF)*** program began in 2010 with the goal of stimulating and encouraging creativity and innovation "from within the NASA Centers." In its first year, the program funded twenty research proposals at the ten NASA research labs with early stage research ideas in novel technologies and new processes that have "the potential to revolutionize or enable new capabilities in space flight, science, aeronautics, and exploration." The program focuses on what are called early Technology Readiness Levels (TRL) – which implies it emphasizes high-risk research at the applied research and proof-of-concept stages of innovation Since its initiation, the program has grown, and expected total funding for FY2012 is $17.5 million.

The NSF's CREATIV (***Creative Research Awards for Transformative Interdisciplinary Ventures***) program was launched in FY 2012, and unlike traditional NSF proposals, it requires only internal merit review. CREATIV proposals are required to be

interdisciplinary and potentially transformative, with larger than average-sized requests (up to $1 million and up to 5 years duration). The program expects to make 30-40 grants in its first year.

The Department of Defense's ***Defense Advanced Research Projects Agency*** (DARPA) was established in 1958 in response to the Soviet launching of Sputnik. DARPA funds project-based assignments but also aims to create teams and networks to solve those problems. Projects are typically organized around a specific technological challenge. Funded researchers work closely with DARPA program managers, under a range of funding mechanisms, and projects often have specific milestones with required deliverables for continued funding. Projects are typically funded for 3-5 years.

The NSF ***EArly-concept Grants for Exploratory Research*** (EAGER) mechanism was created to support exploratory work in early stages of research on untested, but potentially transformative, research ideas or approaches. This work can be considered especially "high risk-high payoff" in the sense that it involves radically different approaches, applies new expertise, or engages novel disciplinary or interdisciplinary perspectives.

The ***Exploratory Advanced Research Program*** (EARP) in the Department of Transportation was established as part of the Safe, Accountable, Flexible, Efficient Transportation Equity Act-A Legacy for Users (SAFETEA-LU, 2005). It "focuses specifically on longer term and higher risk breakthrough research with the potential for transformational improvements to plan, build, renew, and operate safe, congestion free, and environmentally sound transportation systems." Funding is done through contracts and cooperative agreements, and projects are typically for 2-4 years.

The ***Emerging Frontiers in Research and Innovation*** (EFRI) program is located within the National Science Foundation's Engineering directorate. It was formed in 2007 due to concerns that emerging areas of engineering were not being sufficiently funded through the normal engineering directorate areas. Each year two new emerging areas are identified and solicited. The program emphasizes interdisciplinarity and requires a minimum of two PIs on each grant.

The National Institutes of Health's ***Exceptional, Unconventional Research Enabling Knowledge Acceleration*** (EUREKA) was formed in 2007 with the goal of fostering "exceptionally innovative research that, if successful, will have an unusually high impact on the areas of science that are germane to the mission of one or more of the participating NIH Institutes." The program, defunded toward the end of 2011, stated that it is not intended to fund pilot projects. Teams were allowed to participate in this grant program, and awards were for up to $800,000 over a four year period.

The ***Game Changing Development*** (GCD) Program within the Office of the Chief Technologist at NASA focuses on identifying and supporting innovative, high-impact capabilities and technologies, and complement them with "new start" and competitively selected projects. Unlike CIF, GCD emphasizes mid-level TRLs with the goal of maturing technologies. The program had a slow start in FY 2011 with prior commitments, but it will begin to make significant investments (about $30 million a year) in FY2012.

The ***Homeland Security Advanced Research Projects Agency*** (HSARPA), established as part of the Homeland Security Act of 2002, was modeled after DARPA and focuses on "homeland security research that could lead to significant technology breakthroughs." HSARPA initially focused more on technology prototyping, but since a change in 2006, has encompassed a broader range of R&D activities.

The ***Intelligence Advanced Research Projects Activity*** (IARPA) is located within the Office of the Director of National Intelligence and was formed to coordinate the nation's intelligence R&D portfolio. It was also modeled after DARPA and aims to support high-risk, high-payoff research through the range of funding mechanisms.

The ***NIH Director's Pioneer Award*** (NDPA) was established in 2004 in response to concerns that the NIH was overly conservative in its funding portfolio. The program aims to support individual investigators "of exceptional creativity who propose pioneering and possibly transformative approaches to major contemporary challenges that have the potential to produce a major impact in a broad area of biomedical or behavioral research.". Awardees are given extreme flexibility in their funding ($2.5M over 5 years) and are not tied to a specific project.

The NIH Director's ***New Innovator Award*** (NIA), formed in 2007, has two goals: to stimulate highly innovative research and to support promising new investigators. Like the NDPA, it gives awardees extreme flexibility in how their funds are spent and does not focus on a specific problem to be addressed. Recipients are required to be no more than 10 years out from their latest degree. Grants are for $1.5 million over 5 years.

The ***NASA Innovative Advanced Concepts*** (NIAC) Program is the reconstitution of the 1998-2007 NASA Institute for Advanced Concepts program that funded revolutionary concepts to transform future aerospace endeavors. In its current form, NIAC funds revolutionary concepts with the potential to transform future aerospace missions. Projects are chosen for their innovative and visionary characteristics, technical substance, and early development stage -- 19 years or more from use on a mission.

The Department of Defense's ***National Security Science and Engineering Faculty Fellows*** (NSSEFF) was modeled after the NDPA program, but has an additional goal of fostering a long-term relationship between the fellows and the DoD. Established in 2008, the program awards $3 million over 5 years, and requires awardees to participate in activities designed to give awardees an understanding of the research needs of the DoD.

The ***Outstanding New Environmental Scientists*** (ONES) program at the NIH's National Institute on Environmental Health Sciences (NIEHS) "is intended to identify outstanding scientists who are in the early, formative stages of their careers and who intend to make a long term career commitment to research in the mission areas of NIEHS." Recipients are encouraged to form an advisory committee to help them pursue their research and develop professionally. Awardees may receive up to 5 years of funding at levels higher than traditional NIEHS grants.

The NIH's ***Exploratory/Developmental Research Grant Award***, known as the R21 mechanism, is "intended to encourage exploratory/developmental research by providing support for the early and conceptual stages of project development." Award amounts are smaller than the traditional NIH grant mechanism, and are limited to 1-2 years in duration.

DOE's ***Sunshot Initiative*** at the Department of Energy (DOE) was created to reduce the installed cost of solar energy systems by 75% over the next decade (to $1 per Watt installed) to achieve full competitiveness with fossil fuels for electricity generation. Under the SunShot Initiative, the DOE is funding research and loan guarantees for high risk, high payoff concepts—technologies that promise transformation in the ways solar energy is generated, stored, and utilized.

The National Institute of Standards and Technology (NIST)'s ***Technology Innovation Program*** (TIP) was established as part of the America COMPETES Act of 2007. The program aims "to support, promote, and accelerate innovation in the United States through high-risk, high-reward research in areas of critical national need." Awardees may be small or medium sized enterprises (SME), or joint ventures between SME and academics. A 50% cost-share is required, and TIP provides up to $3 million over 3 years for single-recipients or $9 million over 5 years for joint ventures.

NIH's ***Transformative R01*** (TR01) program is intended to complement the "people-based" programs (NDPA and NIA) by funding "project-based" transformative proposals. There is no limit to the size of the award, up to the program budget.

Each of the programs above is attempting to attain its mission through a combination of approaches. Sections 3.3-3.9 identify notable aspects of these programs, with with Sections 3.1-3.12 providing overarching frameworks that characterize the programs as a whole.

3.3 Definitions of Terms of Interest

Terms such as 'high-risk,' 'innovative,' 'potentially transformative,' and/or 'pioneering' were used liberally through program documents and program announcements. Table 3.2 color codes programs that use and define the terms, and those that use them without definitions. Language tended to be lofty and romantic, to say the least. The NASA Innovative Advanced Concepts (NIAC, 2011) program, for example, described itself as "a program to support early studies of innovative … visionary concepts that could one day "change the possible" in aerospace. Despite use of such terms, programs left their definitions open to interpretation. Program officers were probed to operationalize such

descriptions. They, however, cautioned against narrowly operationalizing the terms due to concerns that doing so would leave out research that is worthy of funding. There was a strong belief that this type of research is best defined as "you know it when you see it" and that, as one program officer put it, "definitions so vague you can drive a truck through them" allow applicants to push the boundaries of what they might normally propose in an application. "We don't need to define innovation – we rely on the scientists in the community to propose what is innovative," said one program manager.

Table 3.2

Terms Used by Programs

High-Risk	Innovative	Game Changing	(Potentially) Transformative / High Impact	Pioneering	Creative
DARPA	EUREKA	ARPA-E	ARPA-E	NDPA	CREATIV
EAGER	NIAC	EFRI	EAGER		EUREKA
HSARPA	ARPA-E	GCD	EFRI		IARPA
IARPA	DARPA	TIP	NIAC		NDPA
NIAC	EARP	CREATIV	TIP		NIA
ARPA-E	EFRI	EARP	EARP		NSSEFF
CREATIV	GCD	EUREKA	EUREKA		ONES
EARP	HSARPA	HSARPA	GCD		TR01
EFRI	IARPA	CIF	HSARPA		
EUREKA	CIF	NDPA	CIF		
GCD	NDPA	NIAC	Sunshot		
NDPA	NIA	TR01	TR01		
NIA	NSSEFF				
R21	ONES				
Sunshot	R21				
TIP	Sunshot				
TR01	TIP				
	TR01				

Dark shading - Term used in program documents, and defined. Lighter shading - Term used in program documents, but not defined

Still, officials noted that the lack of firm definitions did lead to some challenges for the programs. Several program managers complained that some applicants to these programs

did not seem to understand what the programs were aiming to fund – "[the applicants are] still proposing to do what they would propose in a normal grant application" – and also had difficulties in convincing reviewers of applications that the programs were truly intended to fund innovative work: "I've had problems with some of my reviewers saying 'this proposal is too ambitious.' I tell new reviewers that no proposal is too ambitious for this program! But it's a whole new way of examining applications for them." In an attempt to better delineate the types of research they aimed to fund, some of the programs used more measureable requirements, such as requiring the research to be substantially different than what is already going on in the scientist's lab or anywhere else (e.g., NDPA) or requiring the research to be interdisciplinary (e.g., EFRI).

Further complicating issues, several interviewees stated that they believed the definitions of these terms would vary across the federal government. "What DoD believes is innovative may be very different than what NSF believes is innovative, and that could be very different than what NIH thinks. And I believe each agency also has different levels of acceptable riskiness in projects." There was an interest among respondents to better understand how the other federal agencies operationalize these terms.

3.4 Program Origins

Some of these programs were formed as completely new programs, with the intent of communicating to the scientific community that they were aimed at funding wholly new types of research (e.g., NDPA, EFRI), while other programs were designed to be modifications of existing mechanisms (e.g., TR01, EUREKA, NIAC). Program officers associated with programs that were adapted from existing mechanisms stated that they believed they had to work harder to differentiate their programs from the traditional programs, but that doing so was possible.

Both programs starting from scratch and programs that were modifications to existing programs often underwent changes during their existence. For example, prior to 2006, the Homeland Security Advanced Research and Development Agency (HSARPA) managed all extramural R&D at the Department of Homeland Security (DHS), meaning its portfolio included both high-risk, high-reward research as well as more conventional research (DHS Directorate, 2009).

Program staff said a difference between HSARPA and DARPA was that HSARPA was interested in research that could be developed within 6-24 months, as distinct from DARPA research that may take many years to develop. Since 2006, however, HSARPA has evolved to be more like the original DARPA model. HSARPA now has two main programs, one of which is aimed at demonstrating prototypes of high-payoff technologies within a timeframe of two to five years, and the second of which funds high-risk basic research to develop proof-of-concepts for potentially transformative work. (Shea, 2009)

Other programs have made changes to their selection processes or selection criteria over time. The NDPA program, for example, eliminated its pre-proposal phase after reaching a phase where the number of applications was reasonable to review.

3.5 Areas of Funding

Programs varied in their approach to identifying innovative areas. Many programs did not specify which areas were of interest, except for those scientific and engineering areas germane to their mission or capabilities (e.g., NDPA, NIA). Other programs specifically solicited proposals in areas considered to be innovative; and various mechanisms were used to find these areas. NASA's GCD program seeks ideas related to NASA's specific Space Technology Grand Challenges or any of the 14 Technology Areas (TAs) identified in

NASA's draft Space Technology Roadmap. NIST's TIP has as part of its mandate a requirement to fund areas that are not currently addressed by the federal government or private industry. NIST's TIP program staff develop potential areas of interest and then employ outside support to better understand the funding landscape around that area to verify whether there is indeed a funding "gap" that can be filled by TIP funds. The EFRI program at NSF culls ideas from program staff across the Engineering Directorate based on emerging areas the program officers are seeing proposed by academia. The EARP employed a series of workshops involving members of academia, industry, and the government to discuss which areas of science could have the highest impact on the future of transportation.

3.6 Competitiveness of Programs – Funding Rates

One of the tensions that agencies face when launching these new programs is the desire to both pilot and experiment with how funds are disbursed, while not wasting taxpayers' money, but also to send a signal to the scientific community that the funds are worth applying for. For example, one critic claimed that set-aside transformative research programs detract from producing efficient innovative output, due to the very small number of awards given out in a year compared to the time it takes for reviewers and applicants to evaluate and submit applications. This sentiment was echoed by a committee of distinguished scientists (ARISE, 2004) in their recommendation that programs not be launched until "they have enough critical mass to avoid fruitless grant-writing efforts." (p. 3)

The programs examined exhibited a range of funding rates (Table 3.3). The NDPA program at the NIH has a funding rate of approximately 4%, while the TIP program at NIST and the DOT's EARP program have funding rates of approximately 20%. ARPA-E has had a funding rate of 1%, although it is not unusual for a program to be oversubscribed in its

first year, before applicants fully understand the program. For example, in its first year, the NDPA program had a less than 1% funding rate,but that rose to a rate of approximately 4% by its fourth year of implementation.

Table 3.3

Program Funding Rates

Program	Budget	Funding Rate (where publicly available)
ARPA-E	$180 m (FY2011)[4]	1%
CIF	$17.5 m (FY2012)	Most funded (intramural program)
CREATIV	$24 m (FY2012)[5]	n/a
DARPA	$2.8 b (FY2012)[6]	Unknown
EAGER	No pre-set amount	~88%[7]
EARP	$14 m (FY2012)[8]	20%
EFRI	$33 m (FY2012)[9]	25% (4% of pre-proposals)
EUREKA	Program retired	Unknown
GCD	<$155 m[10], $30 m	Unknown
HSARPA	<$ 44 m (FY2011)[11]	Unknown
IARPA	Classified	Unknown
NDPA	$37m (2012)[12]	4%
NIA	$80 m (2012)[13]	5%
NIAC	<$40 m (FY2012)	Unknown
NSSEFF	$21 m (FY2012)[14]	2.2%
ONES	$52 m (FY2012)[15]	Unknown
R21	variable	16%
Sunshot	$111 m (FY2013)[16]	Unknown
TIP	$75 m (FY2012)[17]	20%
TR01	$71 m (2012)[18]	Up to 60

[4] http://arpa-e.energy.gov/About/Budget.aspx
[5] http://news.sciencemag.org/scienceinsider/2011/11/new-nsf-program-sidesteps-external.html
[6] http://www.darpa.mil/NewsEvents/Budget.aspx
[7] http://www.nsf.gov/nsb/publications/2011/nsb1141.pdf
[8] http://www.fhwa.dot.gov/advancedresearch/about.cfm
[9] http://nsf.gov/about/budget/fy2012/pdf/18_fy2012.pdf
[10] http://www.nasa.gov/offices/oct/game_changing_technology/index.html
[11] http://www.dhs.gov/xlibrary/assets/budget_bib_fy2011.pdf
[12] http://commonfund.nih.gov/pdf/Vol_1_tab4_Common%20Fund.pdf
[13] http://commonfund.nih.gov/pdf/Vol_1_tab4_Common%20Fund.pdf
[14] http://www.dtic.mil/descriptivesum/Y2013/OSD/0601120D8Z_1_PB_2013.pdf
[15] http://www.cossa.org/volume30/FY2012budgetupdate.pdf
[16] http://www1.eere.energy.gov/ba/pba/pdfs/fy2013_eere_congressional_budget_request.pdf
[17] http://www.nist.gov/public_affairs/releases/budget_2012.cfm
[18] http://commonfund.nih.gov/pdf/Vol_1_tab4_Common%20Fund.pdf

Programs used a number of strategies to alleviate the challenges accompanying the low funding rates. Several programs utilized a phased approach to the application and review processes to minimize the amount of time required from the proposer, and the program administration and reviewers. Often times, a shorter preproposal or white paper is submitted and very quickly given a review as to whether it fits the program criteria before an applicant is invited to submit a full proposal. Other programs, if not using a phased approach to the application, required a shorter overall application than what would typically be required of applicants to similar programs.

In terms of program processes, administrators of programs that had a preproposal phase seemed to believe it shortened burden on both applicants and program staff, in that quick reviews focused on potential significance could quickly weed out those proposals that would be better suited to traditional programs. Program staff did say they would be open to the idea of being able to refer applicants to other mechanisms within their agencies, or to other high-risk programs if the project more closely fit another program's goals.

3.7 Timeline and Renewals

Seed-like programs such as EaGER, had the shortest project timelines, ranging from 1 to 3 years. Programs aimed at funding projects meeting specific technological or national goals had typical timeframes of 3-5 years. Those programs aimed at funding an individual scientist to undertake research of his or her own choosing had the longest grant durations, all at 5 years. This longer duration was commonly believed to be the minimum amount of time needed to complete or make substantial progress on a high-risk project, and the duration for these awards is longer than the timeline for traditional grants. For example, the average duration of a typical NIH grant is a little under 4 years, and the average duration of

NSF grants is about 3 years. This longer duration of high-risk projects is consistent with the evidence in the literature on how to best support this type of research, as described by Heinze (2008).

Program officials had mixed views on whether a renewal option was beneficial. On the one hand, programs do not want to leave the projects on a lurch and for the investment to be 'wasted,' but on the other hand, program officials believed that the lack of an option for a renewal may lead grantees to be more high- risk in their approach and created more urgency in conducting the research. One program manager said, "If there's the option for a renewal, scientists may be tempted to avoid the real high-risk problems in order to get quick results so they can secure follow-on funding."

Some program officers noted that these high- risk programs are meant to spur the projects and that by 5 years into a project, a PI should have enough preliminary data to either abandon the idea or apply for follow-on funding through more conventional mechanisms. Table 3.4 summarizes program data on timelines and renewals.

Table 3.4

Timeframe for Each Type of Grant, Cooperative Agreement, or Award & Whether Renewal Is Allowed

Program Name	**Timeframe for grant**	**Renewal allowed**
ARPA-E	< 3 years	Unknown
CREATIV	< 5 years	No
DARPA	3-5 years	Unknown
EAGER	<3 years	Yes
EARP	2 to 4 years	case-by-case
EFRI	4 years	No
EUREKA	4 years	No
GCD	2-3 years	
HSARPA	Unknown	Unknown
IARPA	5 years	Unknown
IF	4 months	No
NDPA	5 years	No
NIA	5 years	No
NIAC	<1 year	No
NSSEFF	5 years	No
ONES	5 years	No
R21	<2 years	No
Sunshot	18 months – 5 years	No
TIP	3 years (single); 5 years (joint venture)	No
TR01	<5 years	No

3.8 Selection Criteria

Some program selection processes differed greatly from their agency's normal processes. The NIH study section is the traditional process for funding the majority of biomedical research in the nation. A study section consists of substantial time commitment from members (up to 40-60 hours prior to meeting, and the meetings themselves last 15-20 hours), and study section members must reach a consensus in terms of which applications to fund. Due to the extremely competitive nature of the applicant-to-funds-available ratio and the desire to fund only the most promising of proposals, it is perceived in the biomedical community that an application (typically 25 pages) will only be funded if it has essentially

been shown through preliminary data that the experiment will work. In contrast, the NDPA and NIA programs employ a process in which 5-page applications, not required to contain preliminary data, are reviewed by a set of external reviewers who never meet to discuss applications and who do not have to reach consensus on their recommendations to fund or not fund.

Selection criteria for these programs were broader than they are for typical R&D programs. Several of the programs relied on reviewers to recognize "innovation" in the applications. There was a belief among many of these program managers that innovation can best be defined, like the infamous Supreme Court ruling on obscenity, as "you know it when you see it."

For those programs aimed at funding an individual scientist, applications included information on the track record of the scientist, including evidence of innovativeness. Some of these programs also included an interview round to better assess the applicant. Program officials of these programs stated the interview round, while costly, was an integral part of assessing whether an applicant was competitive or not.

More so than those applying to traditional programs, applicants to these programs were often asked to explain the innovativeness of their proposal and its potential payoff in language that non-specialists can understand. NASA's Center Innovation Fund, for example, asks simply for novelty and potential impact as its primary selection criterion ("The qualities of the idea that make it a new or unique approach.") While preliminary data were not required for most of the programs, they were allowed, and program officials said that a large share of applicants did choose to submit it.

As Table 3.5 shows, some programs required a preapplication proposal, so that a shorter set of application materials were submitted at first to determine their suitability for the program. All programs prescreen applications to ensure that they are complete.

Table 3.5

Length of Application Form and Prescreening of Proposals

Program Name	**Shortened Application (compared to standard for agency)**	**Program uses pre-application to screen proposals**
ARPA-E	No	Yes
CREATIV	No	Yes
DARPA	No	Yes
EAGER	Yes	Yes
EARP	No	No
EFRI	No	Yes
EUREKA	Yes	No
GCD	No	No
HSARPA	Unknown	Unknown
IARPA	No	No
IF	Yes	No
NDPA	Yes	No
NIA	Yes	No
NIAC	No	Yes
NSSEFF	No	Yes
ONES	Yes	No
R21	Yes	No
Sunshot	No	Yes
TIP	No	No
TR01	Yes	No

3.9 Post-Award Support

Some of these programs differ from traditional programs in providing postaward support. Many ofthe programs aimed at funding an individual to hold an annual symposium, designed to honor thescientists, and to create a community of awardees. Program officials felt this created a sense ofprestige for the grant and encouraged the scientists to engage in further high-risk research (Table 3.5).

DARPA is well-known for its hands-on program management. This makes sense given its mission-oriented intent, with a detailed project timeline and specific deliverables. DARPA managers develop metrics for their own projects, and are able to evaluate the effectiveness of the programs given these metrics. The NIH ONES program, designed not only to support innovative research but also to develop new investigators, asks that the awardee develop an external advisory committee to meet once a year and provide support to the investigator in terms of research progress and professional development.

3.10 Types of Programs Found – By Program Goals

Heinze (2008) noted two types of high-risk high-reward programs in his study of creative research. The first type was of programs that "target individual scientists and provide them with the means to engage in the long-term development of sometimes risky ideas... Clearly their focus is on the selection and the support of highly talented individual scientists" (p. 16). The second was programs that "target unconventional ideas that would likely be eliminated under peer review ... Such programs are perhaps best viewed as support vehicles for unconventional ideas that are developed until they are better suited for more traditional follow-up funding elsewhere" (p. 16).

The review of the programs described in the sections above certainly found both types of programs. But the analysis also revealed other modes – one focused around a national need or mission, and another that created new fields or moved current ones forward via giant leaps. In fact, examining the programs above revealed a more nuanced multi-dimensional typology than Heinze's – three types of programs that support two types of performers. These categories are not mutually exclusive, and many programs placed in one category have characteristics of the other typologies, yet using this typology is useful for deriving higher-level insights. Table 3.6 lays out this typology.

Table 3.6

A Proposed Typology of Transformative Funding Programs

		Driving Force		
		End-Game Driven	***Synergy Programs***	***Seed Programs***
		Driven by grand challenges	*Roots-up, topic agnostic, open-ended exploration*	*Exploratory - Could be top-down or roots-up*
Performer	***Individual Researcher***	NSSEFF	CREATIV	All ARPAs
		Sunshot	EAGER	EAGER
			EUREKA	EFRI
	Rock Star		NDPA	
			NIA	
			R21	
			ONES	
			TR01	
	Research Team	All ARPAs	CREATIV	
		Sunshot	EARP	
		CIF	EFRI	
	Rock Band	GCD	EUREKA	
		NIAC		
		TIP		

The analysis revealed "***end-game" programs*** that are driven by grand or societal challenges. These are programs aimed at funding high-risk projects that have the potential to substantially contribute to the *agency's mission or a broader national need* (e.g., CIF, DARPA, ARPA-E, GCD, HSARPA, IARPA, EARP, TIP) - These programs are usually goal-driven

and typically operate using cooperative agreements or contracts; although the projects they fund are high-risk and high-payoff, they are usually not hypothesis-driven exploratory projects, and are often times focused on high-risk technology development. As Table 3.7 shows, these programs tend to be longer-term and large, their selection criteria include attention to project plans, and typically are actively managed to deliverables.

Table 3.7

Major Designs Seen in Transformative Research Programs

Type of HR/HR Program	*Project Funding (per year)*	*Project Timeline*	*Selection Criteria*	*Other Aspects*
End-Game	$500k - $10M	3 years	Work Plan, Relevance to Agency's Mission	Typically uses strong program management
Synergy	$400k - $500k	4 years	Significance of problem to be addressed, interdisciplinary approach	
Seed	$15k - $100k	2 years	Potential impact	Also used to support time-sensitive projects

A second type of programs is ***synergy*** programs that push frontiers in a roots-up way. These are programs aim at funding high-risk projects that will *move a field forward or integrate existing scientific fields* at a rate faster than normal research (e.g., EUREKA, TR01, EFRI) - These programs often encourage or require teaming and interdisciplinarity. As Table 3.7 shows, these programs also differ from conventional programs in that the funding amounts are typically larger and/or allow more flexibility in how funds are spent**.**

The third type of programs can be considered **seed** programs that nucleate and nurture breakthrough but unformed or not fully formed ideas. These programs aim at jump-starting a new research area for a single or small group of scientists (e.g., EAGER, R21) through *"seed funding,"* with the idea that the project may advance enough so that more conventional grants may be sought. The amount of project funding is one third or less than funding obtained through traditional grants, and the funding duration may be shorter.

Table 3.6 also describes the second axis of the typology – whether the program funds an individual or a team. There are programs that fund exceptionally creative individuals or "rock stars." These programs attempt to identify *individual scientists* who have the potential to see a high-risk project to fruition (e.g., NDPA, NIA, ONES, NSSEFF). These programs allow for a single principal investigator (PI) only, and place as much emphasis on the investigator qualifications as the project details when making funding decisions. They often require a minimum effort commitment; these programs may also have an emphasis on bringing new investigators into their agency's fold, or nurturing junior investigators. The funding duration for these grants is typically longer than a normal grant. Then there are programs that fund *teams of researchers* or rock "bands." All the ARPA programs (e.g., DARPA, IARPA, ARPA-E) are typically pursuing large and complex challenges, which are feasible to address only with multidisciplinary teams. The "rock" metaphor in the typology emphasizes the counterculture/maverick nature of researchers who are predisposed to pursuing research intended to displace prevailing wisdom.

3.11 Types of Programs Found – By Strategies

The review of transformative research programs above revealed that there is no one way to support transformative research, and that based on the goals of the program, it can

be supported in many different ways. The strategies can nonetheless be grouped into four categories, which could be the basis of a common framework via which transformative research can be examined. Figure 3.1 begins to build a notional diagram of such a framework. At the most basic level, there are four stages to supporting scientific research. One must design a program, and then use some mechanism to review incoming proposals. Next research must be conducted by researchers. Finally research must be managed. While all stages are connected by feedback loops, the basic flow, illustrated in Figure 3.1, remains. A key insight here is that even in this traditional flow of research activities, there is a possibility of producing transformative research.

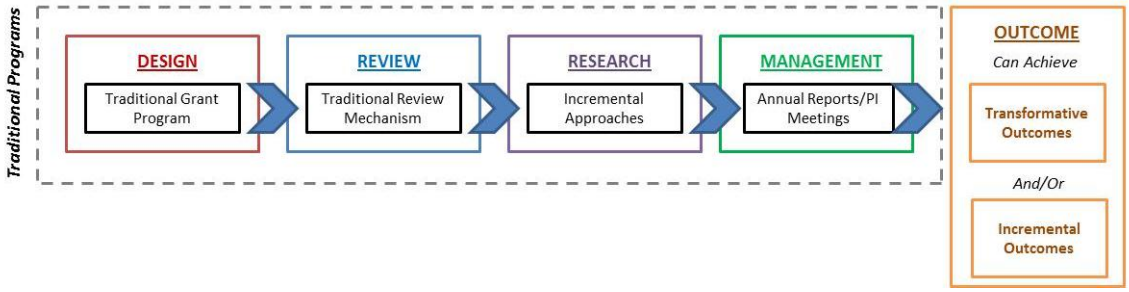


Figure 3.1. A model to illustrate a traditional research funding model

In non-traditional programs, like the ones described in this chapter, the research activities are enhanced, through a combination of "people" and "process" strategies that help them to adapt to the special nature of transformative research (for example, the need to incorporate high-risk research). In the "design" stage, a "people" strategy refers, for example, to creating a process, like the NDPA does, that funds people with a track record of conducting creative research rather than well thought-out experiments or projects. At the "review" stage, the "people" strategy refers to the use of exceptionally creative researchers, mavericks, and other out-of-the-box thinkers not typically invited to evaluate agency proposals. At the research stage, a "people" strategy is about funding innovative thinkers

who may not be competitive in a traditional program, but are especially suited to conducting risky research. At the "management" phase, a "people" strategy may refer to unconventional management of research.

Organizations examined in this chapter also use what could be referred to as "process" approaches to fund transformative research, including use of exceptionally large or small amounts of funds, unconventional review panels, support of teams or interdisciplinary research, and use of aggressive metrics to drive end-game outcomes. The NSF EFRI program, for example, requires interdisciplinary proposals led by two or more PIs as a pre-requisite to apply for the program.

Figure *3.2* illustrates this combination.

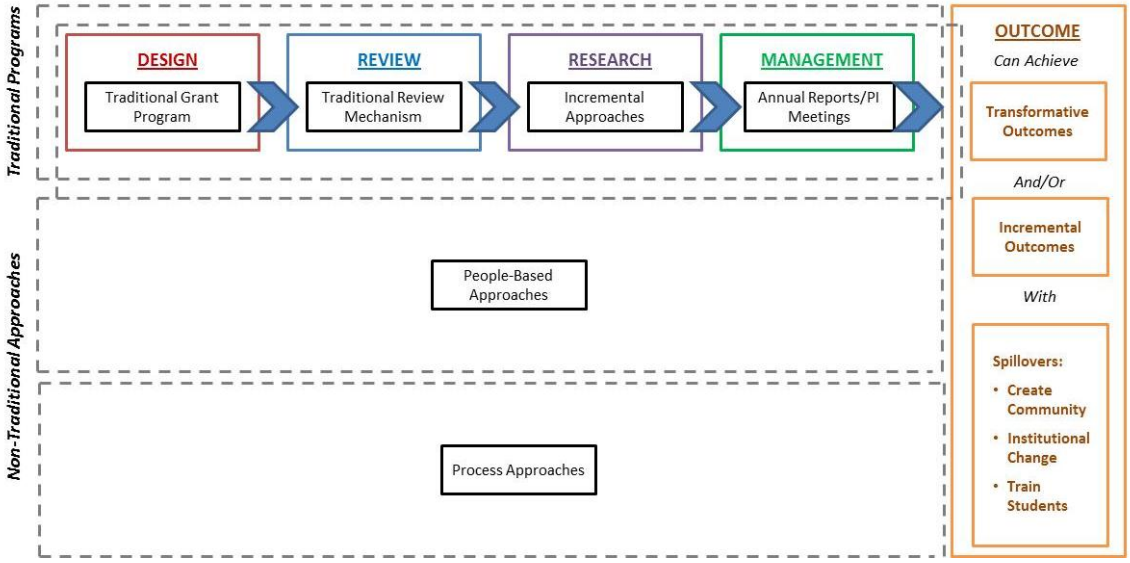


Figure 3.2. A model to illustrate the approach used by transformative research programs

Figure 3.3 inserts all the people and process strategies used by programs included in the program scan, to illustrate the variety in the design, implementation, and management of programs. The topmost section in the Figure – within the dotted lines – is left in to remind the readers that a traditional program can lead to transformative outcomes as well. Most

organizations studied above use some combinations of "people" and "process" approaches to fund their transformative programs – a mix-and-match strategy, so to speak. Figure 3.4 and Figure 3.5show two programs' customizations of these approaches.

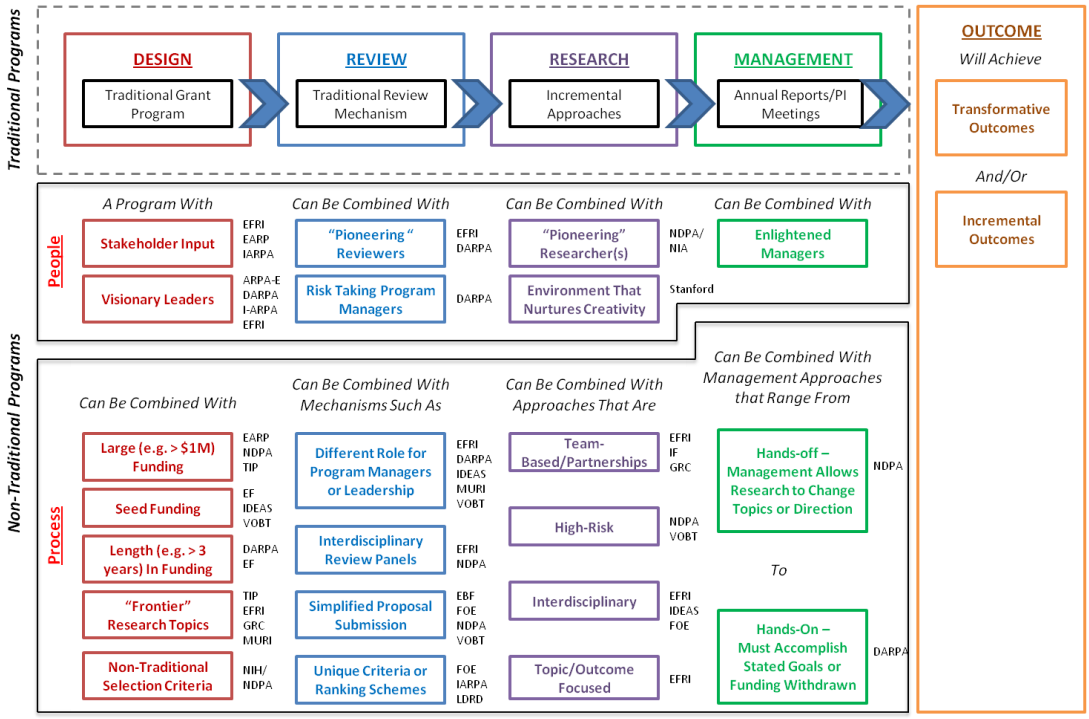


Figure 3.3. A model to illustrate various ways in which transformative research is funded

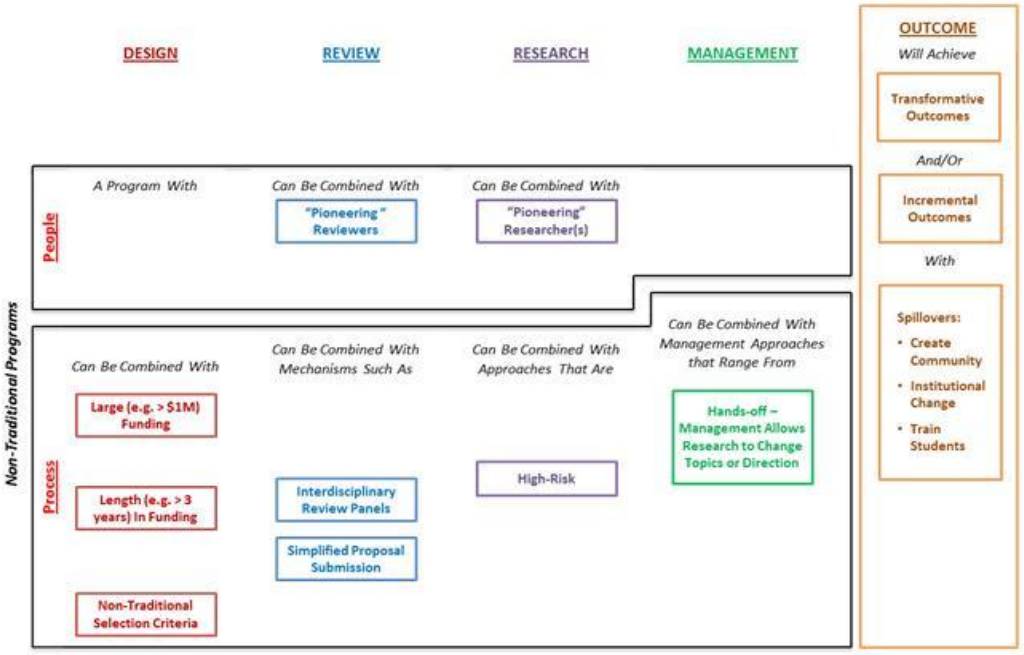


Figure 3.4. NSF's EFRI program – A synergy program

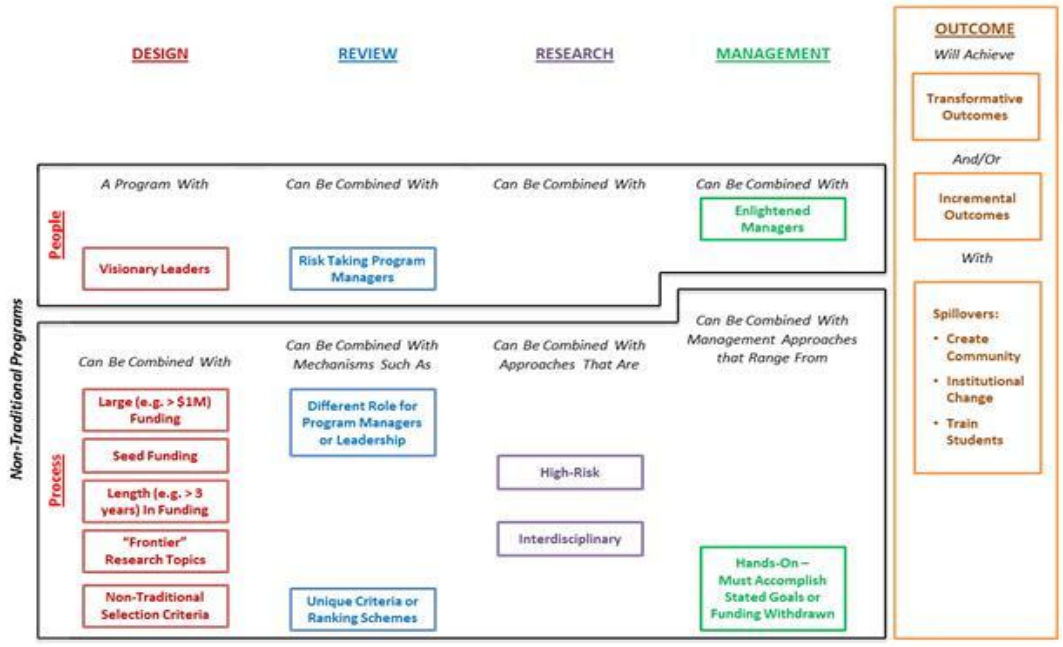


Figure 3.5. ARPA agencies – End game programs

The most interesting aspect of the use of these people-and-process approaches is that no one really knows what will work. Based on their organizational history and culture (not to mention political winds of the time that created programs like NIH's New Innovator Award), some heuristics are developed, and programs created and funded. But there is no theoretical or empirical evidence that these approaches will result in the wished-for outcomes. To visualize this confusion, one could imagine throwing darts in a darkened room, and hoping that one or two of them will hit the bulls-eye.

3.12 Types of Programs Found – by Approach to Research

In Chapter 2, the Kuhnian model of science was introduced. In this model, science proceeds via an incremental process, and accumulates anomalies. These anomalies accumulate to the point where a conceptual revolution (i.e., a transformative event) occurs. In the Kuhnian model, revolutionary science (or paradigm shift or transformative outcome) is fundamentally a consequence of "normal" science. In a Kuhnian interpretation of transformative research, if one wants to encourage more transformative research, then one ought simply to support more normal science. Most mainstream research follows the Kuhnian model, and as expected, every now and then, transformative outcomes occur. This is illustrated in Figure 3.1.

Froderman & Holbrook (2012) introduce three other models to explain transformative research. These were observed in the program review above. They call the first the "quantum' model" – which sees transformative research as different in kind, "a break or phase change independent of normal science." (p. 5) This differs from the Kuhnian model in that there is no build-up beforehand; discoveries arise serendipitously. In a quantum interpretation of transformative research, one would systematically look for radical breaks in approach across a variety of fields, even when there has not been a buildup of

anomalies. Some of the older set-aside transformative research programs, the ARPA agencies, in particular, tend to promote the quantum model.

The second is the "continuum" model, where transformative research is one extreme of a continuum that begins with banal science, with normal, puzzle-solving research activities placed near the mean. The continuum interpretation of transformative research is similar to the quantum approach, with the difference that the policymakers look for instances of transformative research in terms of a sliding scale of transformation. Many of the recent transformative research programs, for example, NSF's EFRI or NIH's NDPA programs are examples of the continuum model.

The third is the "Christensenian" model, based on Clayton Christenson's ideas of "'disruptive innovation" (Christensen 1997), where transformative insights start as inferior to existing interpretations and then over time gain power and momentum. The Christensenian model lends itself to support of ideas that offer alternative approaches for addressing problems, even though they appear worse compared to existing approaches. DARPA and ARPA-E continually seek to create disruptive innovation.

3.13 Conclusion

In this Chapter, I reviewed twenty ongoing federal programs that claim to support transformative research. My goal was two-fold. The first was simply to learn more about transformative research, and how it is defined and operationalized in different contextual settings. The second and more important goal was to discover if the programs use any heuristics – whether clearly-articulated or tacit – that would better help explicate the characteristics of transformative research.

The review showed three types of programs (end-game driven, synergistic, and seed) that support both individuals and teams. It also led to the development of a framework that

organizes the strategies used by programs into two groups – process-based strategies and person-based ones – and shows attributes that program leaders believe they must emphasize to identify and nurture transformative research. Lastly, it helped categorize the funding streams into four models – Kuhnian, quantum, continuous, and Christensenian.

But most of all, the review yielded some basic hypotheses, or propositions, that programs use to fund transformative research. These are discussed in the following Chapter.

4.0 Introducing Propositions

There has been much transformative research in the history of mankind. In the modern scientific enterprise, most of the breakthroughs, from the decoding of the human genome to the discovery of graphene, have come from mainstream and traditional funding programs. As Chapter 1 indicated, in recent years, however, there has been increasing interest in purposefully creating transformative research via set-aside programs, and as Chapter 3 demonstrated, transformative research programs have proliferated across the government. However there is no theoretical basis for what makes for transformative research or researchers. Based on a review of the literature of the science of science policy (summarized in Chapter 2), and a review of federal programs funding transformative research (summarized in Chapter 3), some heuristics come to the fore. These are more perceptions than they are proven principles, and it would be worthwhile to explore them systematically and with data.

In this chapter, I introduce six such heuristics or "propositions." Consistent with study questions two and three, the first three relate to individuals that conduct transformative research, and the second three to transformative research itself.

4.1 Proposition 1: Track Record

Several programs reviewed above select awardees based on their prior record of conducting transformative research. The NDPA solicitation, for example, specifically seeks out evidence for the investigator's claim of innovativeness/creativity. The privately-funded Howard Hughes Medical Institute's Investigator program similarly seeks to researchers with a proven track record.

In all these programs, there appears to be an implicit assumption that researchers likely to conduct transformative research are likely already doing it, and the challenge is to

encourage them to submit their breakthrough ideas, and then identify them in a specially targeted selection process. The proposition has somewhat of an empirical base in a related area: venture capitalists take the academic reputation of researchers are one of the most reliable signals of the potential value of their inventions (Zhang, 2007).

The proposition therefore is: ***Transformative research tends to be conducted by researchers with a track record of conducting transformative research***. If this proposition is supported by data, demonstrating record of prior transformative activity may be an important (though not the sole) marker of those applying for funding to transformative programs. If not, awardees should likely be determined based on the quality of the ideas submitted rather than having to demonstrate a track record of exceptional creativity.

4.2 Proposition 2: Youth

Isaac Newton was 23 when he developed his theory of gravitation, Einstein 26, when he published his Nobel-Prize-winning paper on the photoelectric effect, and. Marie Curie 30 when she discovered the radioactive elements radium and polonium (Luiggi, 2011). Author of the recent book *Imagine: How Creativity Works*, Jonah Lehrer asserted: "Scientific revolutions are often led by the youngest scientists… Youth and creativity have long been interwoven; as Samuel Johnson once said, 'Youth is the time of enterprise and hope.' Unburdened by old habits and prejudices, a mind in fresh bloom is poised to see the world anew and come up with fresh innovations—solutions to problems that have sometimes eluded others for ages" (Lehrer, 2012).

There is support in the funding community. According to the European Commission's Future and Emerging Technologies (FET) scheme[19]:

> Many young scientists have a natural attitude to think out-of-the-box, and to target frontier challenges. Unbiased by received wisdom, they tend to embrace new trends early and to invest energy in high-risk transformative research If this proposition is shown to be true, restricting transformative programs to younger researchers could be an important (though not the sole) criterion of those applying for funding to transformative programs.

Similarly, programs like the NIH New Investigator Award, for example, specifically seek investigators within ten years of a degree. HHMI has similar criteria, and limits its flagship Investigator funding program to researchers who are within ten years of a faculty appointment.

While there are some anecdotes to the contrary – biologist Oswald Avery, who was in his 60s, and already retired from his faculty position at the Rockefeller Institute, when he made his radical claim that genetic information is carried in DNA rather than in proteins (Harman, 2009) – the overwhelming belief relates paradigm-shifting with youth.

The proposition here is: ***Researchers conducting transformative research tend to be young***. If this proposition is supported by data, restricting transformative research programs to younger researchers could be an important (though not the sole) criterion for such programs. However, if not, programs considering age cut-offs would be well-served by not doing so.

---

19 ftp://ftp.cordis.europa.eu/pub/news/research-eu/docs/focus9_en.pdf

4.3 Proposition 3: Productivity

The famous chemist Linus Pauling is thought to have once said that the best way to have a good idea is to have a lot of ideas. Along those lines, Simonton (2004) posited that scientific creativity is a "probabilistic consequence" of research quantity, and the likelihood of a researcher's peers finding his or her work creative is a probabilistic consequence of the number of papers published. Heinze & Bauer (2007) found this to be true in their research on creativity in the field of nanotechnology. There is anecdotal data that some of the best known researchers of our time are also the most published. Therefore, the total number of publications could be used as an indicator for transformative research.

The proposition can be stated as: ***Researchers conducting transformative research tend to be prolific publishers***. If this proposition could be supported by data, knowing if a researcher has a strong publication record may be a good marker for the researcher conducting transformative research in the future.

4.4 Proposition 4: Risk

To achieve scientific breakthroughs, practitioners and policymakers alike believe that risk is inevitable. And risk, they claim, always entails the distinct possibility of failure, which can be measured at many levels, from the individual level, where a scientist might fail to prove a potentially transformative hypothesis at the laboratory bench, to the agency level, where taxpayers might fail to get their money's worth.

But risk-taking, at any level, may not be a critical pre-condition to produce transformation. In fact, some of the most transformative discoveries of our time, whether the discovery of penicillin or the equations guiding search algorithms on the Internet, came from incremental research or in serendipitous ways. However, the view persists that risky research is likely to be paradigm-shifting.

The proposition to be explored can be stated as: ***Transformative research tends to have a high risk [of failure]***. If this proposition is supported by data, risk could be more explicitly incorporated in research programs.

4.5 Proposition 5: Interdisciplinarity

The Keck Foundation, in its justification to support interdisciplinary research, asserted:[20]

> "Advances in science and engineering increasingly require the collaboration of scholars from various fields. This shift is driven by the need to address complex problems that cut across traditional disciplines, and the capacity of new technologies to both transform existing disciplines and generate new ones." (Keck Foundation website)

The rationale behind this sentiment is that people who work at the intersection of scientific communities are more likely to be familiar with selecting and synthesizing alternatives into novel ideas. Stated another way, if a new idea connects previously disparate patches of knowledge, then its transformative potential is higher than the potential of ideas that are limited to well-trodden paths over the existing structure. This suggests that scientists, who create novel connections between two or more previously disparate units of scientific knowledge, have a higher probability of exposure to alternative ways of thinking and behaving. An example is that of string theory, in which its proponent Juan Maldacera proposed a relationship between quantum gravity and quantum field theories, which elucidates various aspects of both theories (Chen 2011).

---

[20] http://www.keckfutures.org/about/study.html,

The proposition can be stated as: ***Transformative research tends to be interdisciplinary***. If this proposition is supported by data, programs can specifically seek interdisciplinary research, either explicitly, or requiring that research be led by multiple investigators from different disciplines.

4.6 Proposition 6: Skepticism

An important characteristic of transformative research is peer skepticism or controversy. There are oft-repeated stories of how discoveries that transcended existing paradigms and challenges current theories languished because peers were not ready to give up old orthodoxies and accept new ideas, or because they were premature. This skepticism can come at the time a proposal is submitted to an agency for review, where reviewers are skeptical about is validity. It can also come at the time research is published. In some cases in fact, researchers may not even be able to find a journal that would publish their research. Chapter 2 introduced the concept of "sleeping beauties" - publications that took a long time to be recognized as breakthroughs. This delay could be because the ideas they propose contradict prevailing wisdom (as with Australian physicians Barry Marshall and Robin Warren, whose discovery that ulcers are caused by bacteria and not stress brought them ridicule in the biomedical community), or simply that the ideas were ahead of their time (as with many of the algorithms funding in the 1970s and 1980s as part of applied math research, and then not having any use until packet switching networks needed them in Internet circuits).

The proposition can be stated as: ***Transformative research garners disagreement at all stages, and may take longer to be accepted by the community.*** If this proposition could be supported by data, one could design programs to ensure that proposals

that get extremely high and extremely low ratings by reviewers are given a second round of consideration.

4.7 Conclusion

Table 4.1 summarizes the propositions, and Figure 4.1 organizes them using the model introduced in Chapter 3. If the propositions were to be supported by data, policy recommendations around the design of transformative research programs could be guided in the way shown in Figure 4.2.

Table 4.1

Propositions About Transformative Research

Propositions about the Researcher	1. Researchers conducting transformative research tend to have a track record of conducting transformative research
	2. Researchers conducting transformative research tend to be young
	3. Researchers conducting transformative research tend to be prolific publishers
Propositions about the Research	4. Transformative research tends to have a high risk [of failure]
	5. Transformative research tends to be interdisciplinary
	6. Transformative research garners disagreement, and may take longer to be accepted by the community

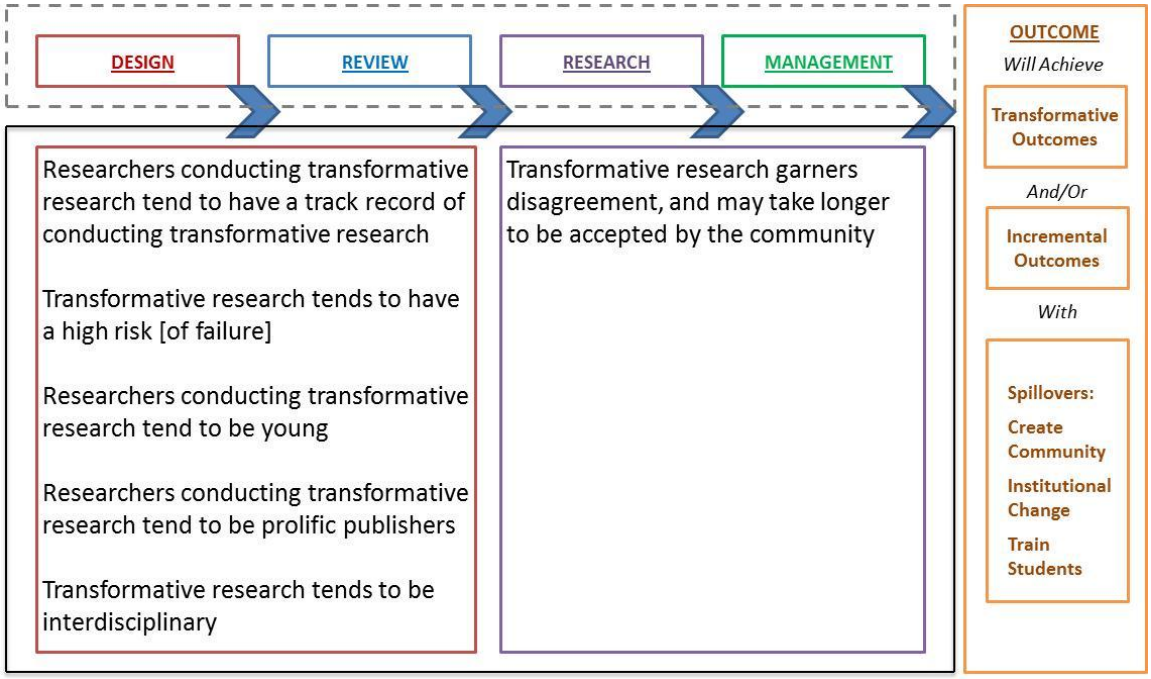


Figure 4.1. Propositions expressed in the proposed framework

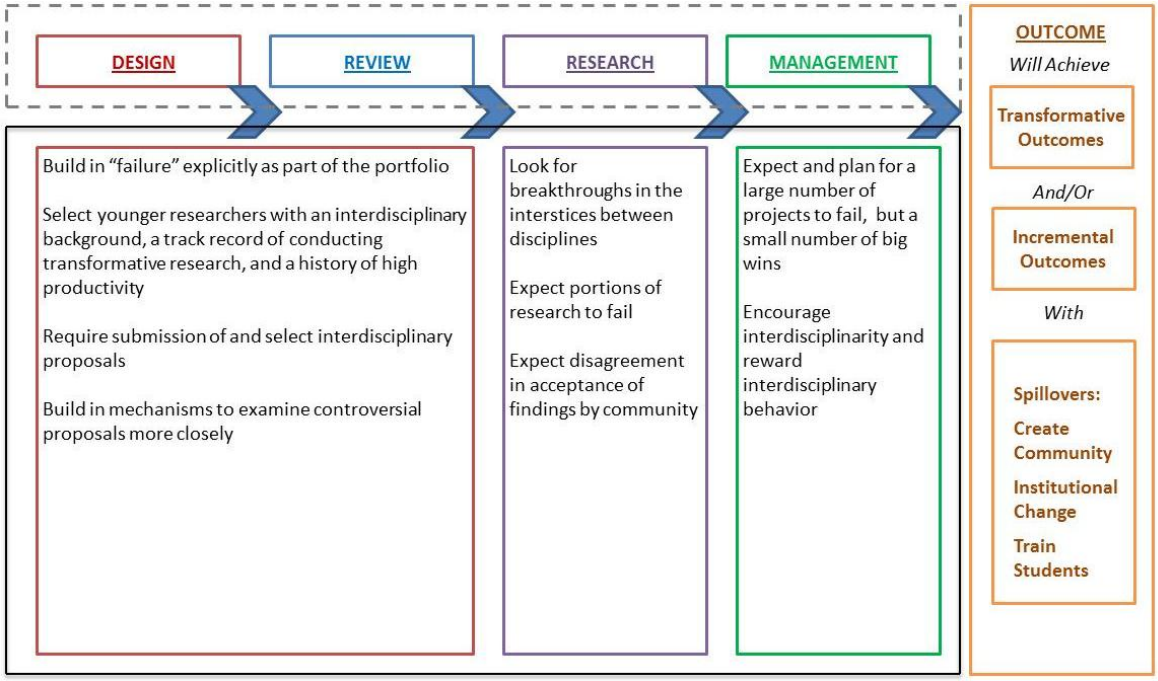


Figure 4.2. Potential policy recommendations if propositions were supported by data

5.0 Methodology

In this Chapter, I describe the methods used to explore the propositions introduced above. Data to explore the propositions come from one of the transformative research programs described in Chapter 3 above; the NIH Director's Pioneer Award (NDPA) program. The program was chosen primarily because data around awardee selection and outcomes was publicly available.

I begin with a brief description of the NDPA program in Section 5.1. Section 5.2 summarizes the specific approach used to examine each proposition. Section 5.3 highlights how each concept of interest in each proposition was operationalized, and Section 5.4 describes how the comparison groups were drawn. Section 5.5 lists the limitations to the analysis. The methods are summarized in Section 5.6.

5.1 Data Source - The NIH Director's Pioneer Award (NDPA)

To explore and test the propositions introduced in the preceding section, data were derived from a research program of particular interest: the NIH Director's Pioneer Award (NDPA) program. This section describes the NDPA program.

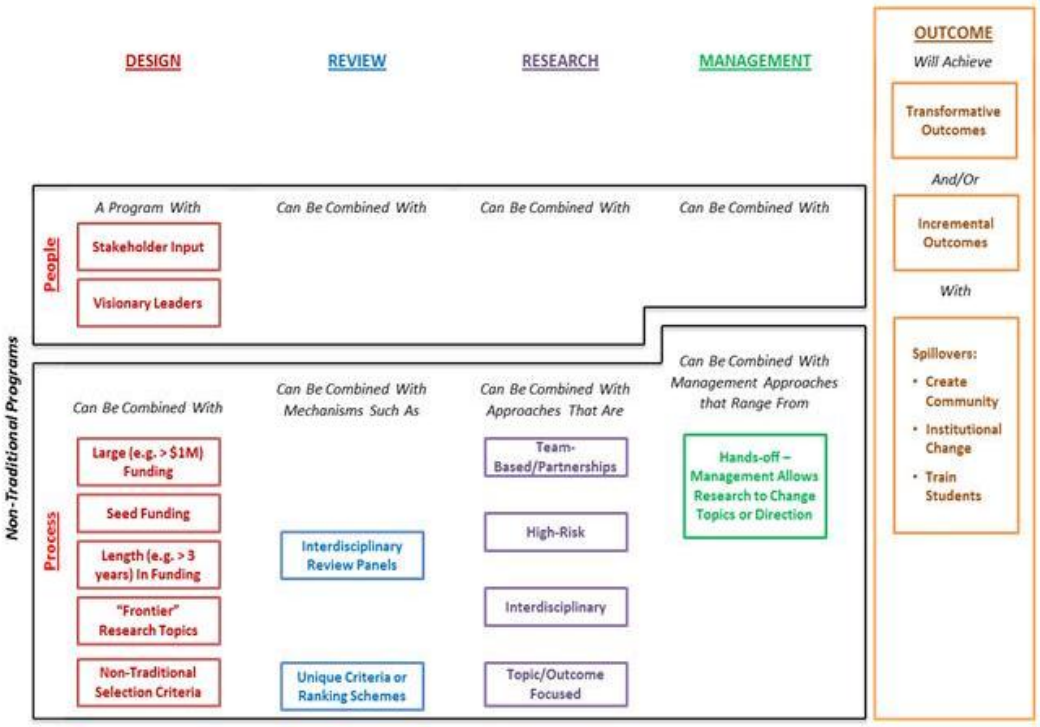
5.1.1 Overview of the NDPA Program

The NDPA is the flagship program of the NIH Roadmap for Medical Research, a framework of NIH priorities designed to optimize the NIH research portfolio and to improve the current state of biomedical and behavioral research through the funding of specific programs. In particular, the Roadmap Initiative aimed to establish programs that promoted high-impact, cutting-edge research, which often did not fall into the interest of a single NIH institution or center. The National Institutes of Health (NIH) Director's Pioneer Award (NDPA) was initiated in fiscal year (FY) 2004 "to support individual scientists of exceptional creativity who propose pioneering approaches to major contemporary challenges

in biomedical research" (NIH, 2004). The NDPA program grew out of concerns that the traditional NIH peer review process had become overly conservative and the belief that NIH required specific means to fund high-risk research (Brenner 1998; Mervis 2004). As such, an implicit goal of the program was to change the culture of NIH to be more innovative and to encourage the development of other high-risk funding programs.

The NDPA program provides individual researchers with $500,000 in direct costs for each of the 5 years, for a total direct cost of $2.5 million (the award size decreased after Year 1). Compared to the average R01 grant, the NDPA provides a greater amount of funding per year; provides funding over a longer duration of time, and does not require a project budget specifying how the funds will be allocated. The NDPA is currently administered through the NIH Office of the Director and is annually awarded to a small number of researchers across a range of scientific disciplines.

A logic model can be used to visually describe the inputs, activities, outputs, outcomes, and external factors that comprise a given program. Figure 5.1 displays a traditional logic model of the NDPA program. Figure 5.2 is an alternative representation of the program based on the model introduced in Chapter 3, and is more nuanced expression of the approach to fund transformative research.



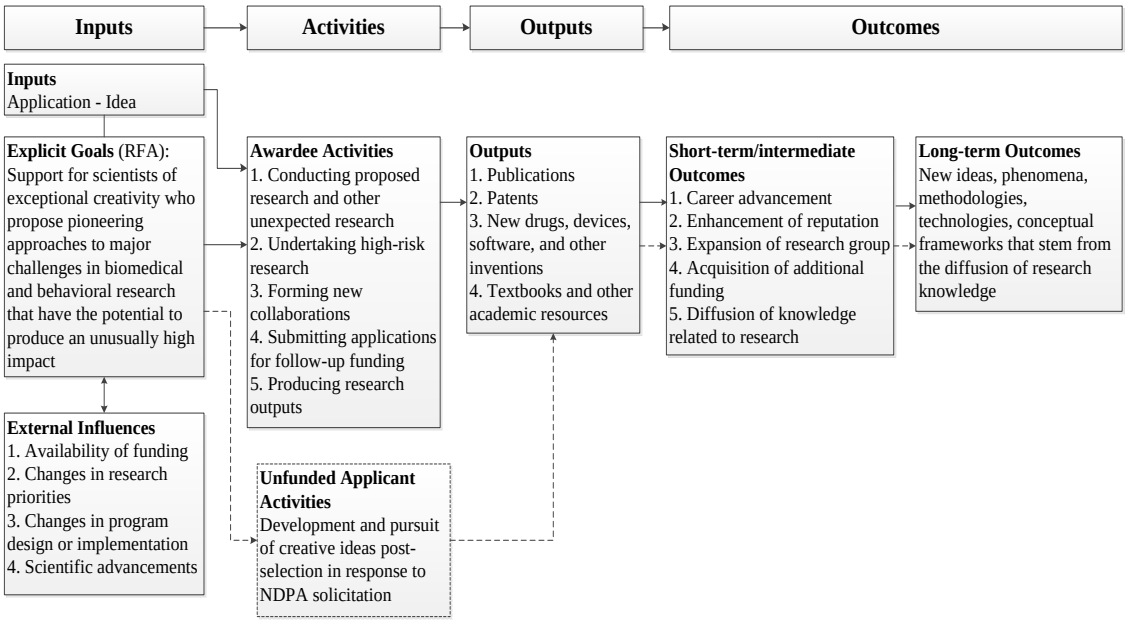


Figure 5.1. NDPA program logic model
Source: Lal et al. (2011)

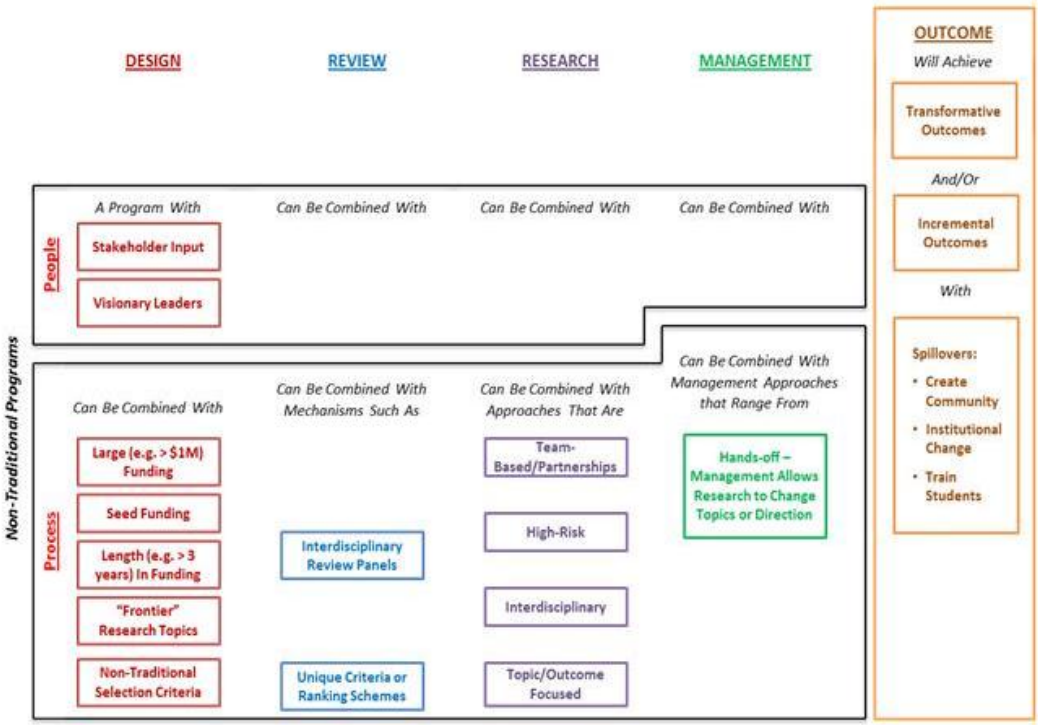


Figure 5.2 NDPA's approach to supporting transformative research

The outputs seen from the NDPA program appear at first glance to be similar to those observed from traditional funding programs. The awardees are expected to publish their findings, to produce research developments such as new drugs, devices, software, and other inventions, and to patent those inventions if appropriate. However, the nature of those outputs is expected to be more pioneering, and expectations regarding the number of outcomes, as compared to traditional R01 grants, differ, as high risk research may not be immediately successful.

5.1.2 Origin of the Program

The roots of the NDPA program can be traced back to FY 2000, when NIH assessed the feasibility of a small initiative run by the Center for Scientific Research. While no action directly resulted from that effort, in 2002, during the development of the NIH Roadmap, a High Risk Research Working Group (HRWG) was created.

In June 2003, the HRWG and its group of consultants came together to design and propose new funding mechanisms at the NIH to promote high risk and innovative research. The HRWG, after a series of meetings with internal and external experts, put forth three potential program designs, two project-based and one people-based. The proposed programs were:

- A *grand challenges program* designed to identify a grand challenge of interest to multiple NIH institutes and centers and fund teams of researchers to meet the challenge;
- An *exceptional projects program* geared toward high-risk projects where individuals submit short applications describing the problem of interest and a proposed approach, which could be funded on a fast-track; and
- A *people-based program* designed to fund individuals with creative approaches to important problems in biomedical sciences.

These recommendations were further discussed at the NIH Director's Budget Retreat. Ultimately the people-based program, which was to become the NDPA program, was approved in FY 2004. in its first year.

To assist in the development of the NDPA criteria, NIH leadership convened a meeting that involved members of the steering committee, as well as academic experts on creativity. Prior to the meeting, the creativity consultants were told neither of the NDPA program nor of why they were being queried on their work. The consultants, asked to advise on what metrics should be used in assessing creativity and innovation, produced the following list:

- The ability to initiate new areas of research or new approaches to research
- A willingness to take scientific risks
- Persistence in the face of adversity
- A willingness to grow scientifically and expand into new areas
- The ability to work in the face of uncertainty
- Scientific vision

- The ability to communicate effectively
- Intrinsic motivation, passion, enthusiasm, and intellectual energy
- Scientific creativity
- Potential for scientific leadership
- A willingness to make mistakes
- The ability to attract first-rate researchers to their labs

In September 2004, the first Pioneer Awards were made, with nine individuals receiving funding under a newly created DP1 activity code. The program announcement emphasized its intent to support individual investigators who display the creativity and talent to pursue high-risk, potentially high-impact ideas in biomedical sciences. It clarified its 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.[21]

NDPA remains a part of the *Research Teams of the Future* theme of the NIH Roadmap. As part of the theme, it "is meant to complement NIH's traditional, investigator-initiated grant programs by supporting individual scientists of exceptional creativity who propose pioneering approaches to major contemporary challenges in biomedical research."[22]

5.1.3 Program Evolution

The NDPA program began in 2004 and is currently awarded every year. Although the selection criteria and number of awardees has changed slightly over the years, the program has largely remained close to its original intention to fund highly innovative people.

---

21 The NIH Director's Pioneer Award Program press release, January 20, 2004. Available online at http://www.nih.gov/news/pr/jan2004/od-20.htm.

22 "NIH Director's Pioneer Award," National Institutes of Health. Available online at http://nihroadmap.nih.gov/pioneer/.

To date, 111 Pioneers have been selected. The Tables in Appendix B summarize the evolution of the program with respect to its recruitment emphasis (Table B-1), its selection process (Table B-2) and its specific selection criteria (Table B-3).

5.1.4 Characteristics of the Pioneers and their Research

Case studies were performed for each of the awardees by Lal et al. (2011) in order to determine whether their research was indeed pioneering, and to examine the impact of their NDPA-funded research on their students, their institutions, the NIH, and the greater research community. The intent was to examine the nature of the outcomes and see how transformative they were. Appendix C summarizes some of the characteristics of the Pioneers and their research. This study focuses on the first three cohorts of the awardees to ensure at least five years have passed since award was received.

5.2 Specific Approach to Explore Propositions

As discussed in the introduction to this Chapter, all propositions were explored using data from the NIH Director's Pioneer Award (NDPA) program. As Section 5.3 illustrates, all concepts were operationalize to make the explorations possible using either bibliometric or programmatic data. To make the exploration meaningful, it was deemed useful to add a comparison group. An ideal comparison group would come from mainstream research programs, and differ only in that that the researchers would not be conducting transformative research. And then only if what *is* true for the test group is *not true* for the comparison group (and vice versa) can help assess the degree to which a proposition could be supported by data. For example, to support the claim that transformative research is interdisciplinary, it needs to be shown not only that a large fraction of transformative research is interdisciplinary, but also that a large fraction of traditional research is not. Section 5.4 describes the attempt to develop such a comparison group.

The first proposition states: *Researchers conducting transformative research tend to have a track record of conducting transformative research*. This is a difficult proposition to address: if we don't know what transformative research looks like now, how can we know what it looked like before. To address this proposition, I therefore looked for differences between a transformative and non-transformative population, and posed the following question: How different do researchers conducting potentially transformative research look with respect to their track records compared to researchers conducting "merely excellent" but traditional research? In particular, how do their prior records differ with respect to the common "currency" of the scientific profession - productivity, quality, and impact? The data used to address these questions were imbedded within the publication records of the researchers. As discussed in Section 5.5, publication-related information was downloaded from Thomson Reuters' Web of Science database.

The second proposition is: *Researchers conducting transformative research tend to be young.* "Young" is a relative term, and the question arises, young compared to whom. Ages of the Pioneers needed to be compared to ages of researchers in mainstream programs for the information to be useful. To address this proposition, I posed the question: Are researchers conducting transformative research (i.e., the "Pioneers") younger than researchers who received mainstream grants in the same time period? Since age-related data are hard to come by, data for a proxy – years since Ph.D. – was obtained from Lal et al. (2010) for the Pioneers, and from public sources for comparison group grantees.

The next proposition is: *Researchers conducting transformative research tend to be prolific publishers*. To address this proposition, the question posed was: How well-published are the researchers in the test group compared with that of researchers in the comparison group? Data to test this proposition was downloaded from the Web of Science database.

The next proposition, *Transformative research tends to have a high risk [of failure],* is difficult to explore, primarily because risk is a difficult concept to operationalize, and also because there are no ready comparison groups. The latter is because typical grant programs do not rate applications on a risk scale, so there is no data with which to compare the riskiness of NDPA proposals. So I explored it in multiple ways, posing a range of questions. The first set of questions was: Was the level of risk (as rated by reviewers of proposals) of the NDPA awardees greater than that of non-awardee applicants? A related question here was: was there a correspondence between what reviewers considered transformative proposals and their levels of risk? To address this set of questions, administrative data from the NDPA program were examined to assess if the Pioneers had "risk scores" that were higher than applicants who did not receive the award. If it appears that Pioneer scores were higher than those of unfunded applicants, it would illustrate that reviewers believed risk to be a virtue when identifying potentially transformative research.

Next, I explored the link between risk and project transformativeness. The questions posed were: Did the riskiest projects show transformative outcomes? Conversely, did projects with transformative results receive the highest risk rankings? In this analysis, risk-related scores were matched with project outcomes to assess if researchers whose proposals were deemed high risk subsequently produced transformative research. I also examined the fraction of projects that was considered as failure and how those projects were ranked in their riskiness.

Lastly, I explored the question: How far did the Pioneers venture with respect to exploring emerging ('non crowded") research areas? To address this question, an NIH-based text mining tool was used to identify how NDPA-funded research was qualitatively different from a comparison group of researchers.

The next proposition was: *Transformative research tends to be interdisciplinary*. To explore this proposition, first the term interdisciplinarity had to be operationalized. This was done by using the concept of the integration score (explained in Section 5.3). Assuming that integration score is a good proxy for interdisciplinarity, the following question was posed: To what extent do Pioneers' publications have higher interdisciplinarity (integration scores) as compared with publications by researchers in the comparison program?

The last proposition was: *Transformative research garners disagreement at all stages, and may take longer to be accepted by the community*. Parts of this analysis had no comparison group, for the same reason as the risk proposition above. Mainstream programs do not formally capture disagreement by peer reviewers and evaluators. Five questions were posed to explore this proposition: To what extent did external reviewers of Pioneer proposals agree with each other in the review of applications? To what extent did external evaluators reviewing the completed research of the Pioneers agree with each other? How do the grant-related citations of researchers in the test group compare with those of grantees in the comparison group? How fast did the publications of the Pioneers increase as compared with the comparison group? How fast did citations of the Pioneers increase as compared with the comparison group? Data sources for these set of questions were programmatic (first two question) and bibliometric (last three questions).

5.3 Operationalizing Concepts of Interest

5.5.1 Track Record

The concept of a track record was operationalized in three ways. The first was the prior publication record of researchers in the test and comparison groups. The second was lifetime quality (as measured by the number of citations to publications and journal impact factors) of prior research. The third was the researcher's lifetime reputation as measured by

his or her h-index (the reader may remember that the h-index is a combined measure of the productivity and impact of a scientist). All observations were between 1980 and year of award. 1980 was selected because it is the earliest year for which data on publications were available in the Web of Science database.

5.5.2 Youth

The best available proxy of youth is years since Ph.D. of the researchers, since that information is more readily available than age-related information. The proxy is not perfect, as many researchers get doctorates late in life. For these researchers, their age does not correspond well with their year of degree. However, I have assumed that the percent of such researchers is small.

5.5.3 Productivity

Productivity was measured by counting the number of papers published by researchers. Again, the proxy is not perfect. There are many ways for the productivity of a researcher to manifest itself (for example, through the education of students, creation of software and other materials not captured in the academic literature). However, publications are the "currency" of science, and therefore not a bad proxy for productivity.

5.5.4 Risk

Risk was operationalized in several different ways. The first was a similarity score, or the degree to which the Pioneers ventured outside traditional areas of research to explore emerging ('non crowded") research areas. The second was an ordinal risk ranking of risk in all NDPA proposals. The third was the researchers' self-perception of the level of risk in the research they undertook. A fourth was the external evaluators' perception of riskiness of Pioneer research.

5.5.5 Interdisciplinarity

As Chapter 2 explains, interdisciplinarity is a difficult concept to operationalize. For the purpose of this analysis, the proxy of interdisciplinarity is embodied in three quantitative metrics. The first is an "integrations score" – this numeric (sections below explain the concept, and how it is computed and visualized) measures the diversity of subject areas that are assigned as keywords in a given publication. The second is a "specialization score" – this score measures the integration scores of the journal in which a publication appears. And the third is a "diffusion score" – this metric measures the integration scores of the publications that are *citing* a given publication.

All require some explaining, The concept of an integration score (I-Score) was developed by researchers Alan Porter, Alex Cohen, David Roessner, and Marty Perreault as a means of measuring the integration of knowledge within a body of research given by the spread of a publication's references (Leydesdorff, 2010; Porter et al., 2007; van Raan, 2002). The measurement of an integration score is based on the diversity of subject categories (SC) assigned to journals (by Thomson Reuter's Institute for Scientific Information (ISI) accessible through the Web of Science database website) and the degree to which a SC relates to a particular journal (de Moya-Anegon et al., 2004, 2007).

While there are other measures of diversity (e.g., Shannon entropy and Simpson Index), Rafols and Meyer note that the I-Score found in Porter et al. (2007) is the only measure that integrates the three aspects of diversity (variety, balance, and similarity) into one index (Rafols and Meyer 2010, Stirling 1998). The I-Score differs from the Shannon and Simpson indices in that the proportions of SCs are weighted by the distance or similarity measure as a normalization factor. Porter calculates the extent of co-citations among SCs and measured the association through Salton's cosine (Porter et al., 2007).

To calculate the I-Score of NDPA and mainstream publications that could be attributed to the grant mechanisms, the publications—the full record plus cited references—were exported into Vantage Point,[23] which was used to assign subject categories to each journal as well as to compute the integration scores.

5.5.6 Skepticism or Peer Disagreement

Like risk, peer disagreement is a subjective concept, and measured in three ways. The first was reviewer agreement of the rankings of NDPA proposals. Each Pioneer application was scored on four indicators of interest – significance, individual, suitability, and an overall score. Since the first three indicators were strongly correlated with the overall score, for this proposition, I examined peer disagreement on overall score.

The second indicator of peer disagreement was expert evaluator rating on which of the NDPA projects were pioneering. This rating came postaward. The third indicator related to peer acceptance of publications that came out of NDPA research. This was measured in three ways. First, how long did it take for researchers to publish – presumably controversial research has a harder time finding a venue for publication. Second, how well did the research get cited - on the whole, if peers agreed with the research, they are likely to cite it. Next, how quickly did NDPA research get cited – presumably research that shifts paradigms penetrates the community at a slower rate than mainstream research.

---

[23] Vantage Point is a text mining software that is used in bibliometric research. Vantage Point includes a program to assign journals the correct WoS subject categories, and a program to calculate the Integration Score. http://www.thevantagepoint.com.

5.4 Developing a Comparison Group

Section 5.1 above describes the specific program – the NIH Director's Pioneer Award (NDPA) program – from which data were drawn. In order to make any reliable statements about the propositions, it was important to have one or more comparison groups. This section describes one primary and three secondary comparison groups.

It would be most meaningful to take NIH's mainstream funding NIH Research Project Grant Program (R01) program as the primary comparison program. The Research Project (R01) grant is an award made to support a discrete, specified, circumscribed project to be performed by the named investigator(s) in an area representing the investigator's specific interest and competencies, based on the mission of the NIH[24].

However, one cannot take the entire mainstream researcher population at NIH as a comparison group. Apart from being intractable (there are over 40,000 R01 awards given every year), it is not a comparable population. PIs come from a range of universities, have had diverse careers, are at different career stages, and work in different areas of research. To make the populations comparable, a series of cuts and matches were made to determine the right population to compare.

5.4.1 Primary Comparison Group – High-Quality Researchers Conducting Research (*n*=35)

First, all R01s that were not "Type 1" and not given in the years of interest (2004-2006) were eliminated from the group. A Type 1 R01 grant is a new R01s (as distinct from a

[24] http://grants.nih.gov/grants/glossary.htm

competitively assigned renewal of an existing grant). Since NDPAs are new grants, it is important that the comparison group grants are also new. This whittled the number of possible grants in a comparison group to 11,170.

Next, the grants were matched to NDPAs by area of research. Matching on research areas was important because transformative research may have subject-area-specific characteristics. For example, a particular research area may be more computationally intensive, and therefore disproportionately represent younger researchers.

R01s that were topically similar to the NDPA awards were selected using the publicly-available text mining NIH Topic Mapping Tool. The tool computes "similarity" scores between grants, and the computation is based on a combination of two techniques (Talley et al., 2011). The first is topic modeling, using Latent Dirichlet Allocation, a Bayesian statistical algorithm that automatically discovers meaningful categories in unstructured text, independent of keywords or preconceived categorical designations. The second is a graph-based layout algorithm, which produces a two-dimensional visualized output in which documents are clustered based on their overall topic- and word-based similarity to one another. These two complementary methods are combined in an interactive web-based format that provides a context in which grants are categorized and clustered based on the language used by researchers. This approach led to the identification of 821 R01 grants that were topically similar to the research being conducted by the NDPA awardees. Some R01s were topically similar to more than one NDPA; overall there were 730 unique R01 researchers identified.

The list of 730 researchers as a comparison group was whittled down further to achieve comparable researcher characteristics. The variables used to cut down the list were:

- year since degree (to make experience levels comparable so as not to compare, say, the productivity of 30-year old and 60-year old researchers)
- institutional prestige (to ensure researchers were similarly qualified, and had access to similar infrastructure)
- prior NIH funding (to ensure at the least the potential of similar productivity)
- terminal degree(s) received (to ensure similar ratios of clinical and research degrees)
- receipt of early career awards and receipt of R01 within 5 years of most recent research or clinical doctorate (to ensure that researchers being compared were similarly qualified)

To select comparable researchers, one R01 from each research stratum was chosen under the constraint that the 35 selected R01s must contain nine that were awarded in 2004, 13 in 2005, and 13 in 2006 (the same distribution as the NDPAs) and be as close as possible to the NDPAs on a set of other characteristics. The final set of 35 matched R01 is compared with the NDPA population below.

Differences between NDPA and Matched Set Grant/Grantees

Figure 5.3 shows the direct costs for the 35 grants in the third comparison group. The average award size is $1.07 million, with 12 grants awarded for 4 years and 21 awarded for 5 years. As the Figure shows, the size of a Pioneer award is about twice as much.

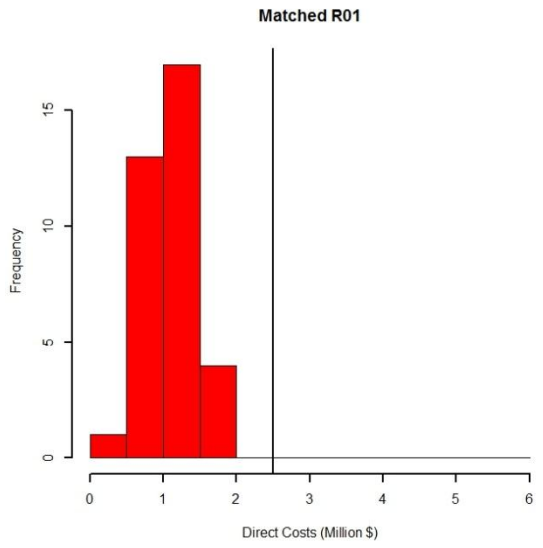


Figure 5.3. Direct costs for 35 R01s in matched comparison group

Figure 5.4 shows the distribution of institutional prestige for the NDPA awardees and the PI for the R01s.

Figure *5.5* shows the distribution of years since degree for the NDPA awardees and the PI for the R01s. Figure 5.6 shows the distribution of prior NIH funding for the NDPA awardees and the PI for the R01s.

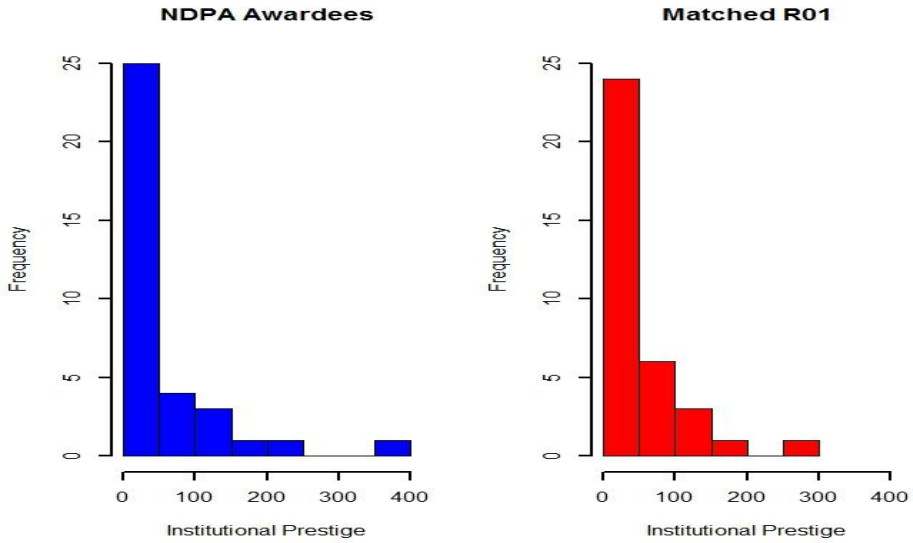


Figure 5.4. Institutional prestige for NDPA and R01 comparison group

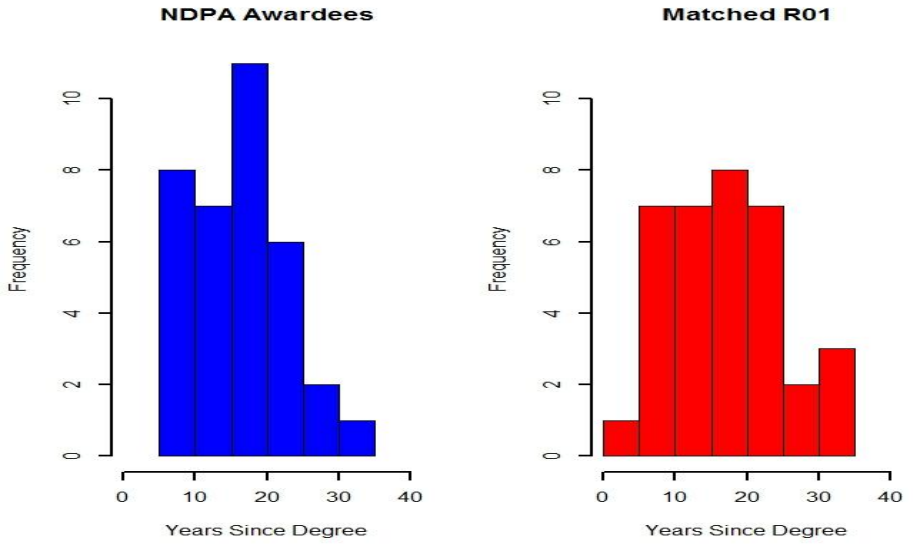


Figure 5.5. Years since degree for NDPA and R01 comparison group

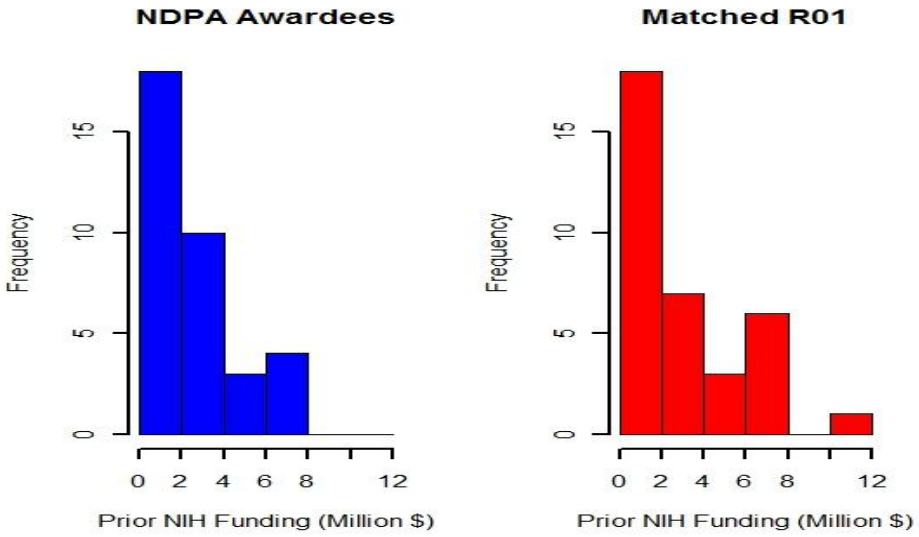


Figure 5.6. Prior NIH funding for NDPA (left) and R01 comparison group

5.4.2 Secondary Comparison Group – All R01 Grantees (*n* = 11,170)

To test the proposition on youth, the primary comparison group was not adequate. This is because selected researchers were matched on age. The second comparison group is therefore the entire set of biomedical researchers who received the mainstream R01 grant in the year in which the Pioneers received funding to conduct pioneering research. For this set, there is no sampling or down selection made, and there are 11,170 members in this group for the years of interest (2004-2006).

5.4.3 Secondary Comparison Group - NDPA Applicants Who Did Not Receive Award (*n* = 891)

Because R01 applications are not specifically expected to propose high-risk ideas, they are not rated on the riskiness or suitability of their proposal. In order to explore the high-risk proposition, I therefore had to add another comparison group, one in which the researchers are also rated on a risk dimension. Applicants to the NDPA who did not receive the NDPA award are therefore an excellent comparison group. Details on the characteristics

of this group, and how it differs from the Pioneers are available in Lal et al. (2010). Table 5.1 summarizes which comparison group is used to explore which proposition.

5.4.3 Secondary Comparison Group – HHMI Investigators (*n* = 30)

HHMI Investigators are included as a comparison group in one particular instance (text mining to explore risk) mainly because of the availability of data on the riskiness of HHMI research of the entire cohort of the class of 2005. HHMI investigators, like the Pioneers, are also high-performing researchers. If their riskiness profile is different than that of the Pioneers, it is a useful finding.

Table 5.1

Relevant Comparison for Each Proposition Explored

PROPOSITION	Comparison Groups		
	Primary Comparison Group Matched-R01 Grantees	**Secondary Comparison Groups** All R01s, Unsuccessful applicants, HHMI Investigators	No comparison group (for some of the questions of interest)
Researchers conducting transformative research tend to have a track record of conducting transformative research	x		
Researchers conducting transformative research tend to be young	N/A*	x	
Researchers conducting transformative research tend to be prolific publishers.	x		
Transformative research tends to have a high risk [of failure]	N/A**	x	x
Transformative research tends to be interdisciplinary	x		
Transformative research garners disagreement, and may take longer to be accepted by the community	X ***		x

*Since matched R01 grantees were selected to represent the age distribution of the Pioneers, this group cannot be used as a comparison group to test the proposition
** Since matched R01 grantees were not selected on a riskiness scale, this group cannot be used as a comparison group to test the proposition
*** Not all subquestions for this proposition could have a comparison group

5.5 Data Sources

Six main data sources were used to test the propositions outlined. The first source was administrative data from the NIH. To gain insights into risk characteristics, external evaluators' scores, and evaluators' qualitative judgment on the applications, data were obtained on candidates' demographic and other characteristics, scores, and prior funding history from publicly available evaluations of the NDPA program (Lal et al., 2010).

The second source of data was interviews with external evaluators conducted as part of Lal et al. (2010). A total of 23 interviews were conducted to gain insights about the reviewers' impressions of the applications they reviewed. These interviews revealed reviewers' insights, among other things, on how they conceptualized transformative research as well as their perceptions of the amount of risk they saw in the applications they reviewed. Data were drawn from these interviews to explore the nature of risk and transformativeness of NDPA research.

The third source of data was coded interviews with awardees conducted as part of Lal et al. (2011). These interviews explored grantees' views on risk and interdisciplinarity. A fourth data source included coded expert review assessments of NDPA outcomes. Insights on risk, impact, and interdisciplinarity were extracted from experts' assessments of NDPA research.

The fifth data source was the Web of Science database which was used to download the relevant publications of the 35 Pioneers and 35 matched set R01 grantees. Once publications were downloaded, related data such as publication dates, citations, number of

co-authors, co-author affiliations, subject areas, and other details were used to compute indices (integration scores, citation rates, etc.) of interest. The sixth source of data was the NIH's public data interface, the Research Portfolio Online Reporting Tool (RePORT)[25]. Other sources included NSF's website and award search database,[26] press releases, and program websites. These are noted in the text when used.

5.6 Potential Limitations

This section outlines the limitations of the analysis along four dimensions: its measurement validity and reliability (which relate to accuracy and repeatability of measurement), internal validity (which relates to establishing causality), and external validity (which relates to the study's generalizability beyond biomedical sciences).

With respect to measurement validity, the biggest concern is how well I am turning abstract concepts (such as interdisciplinarity or skepticism) into empirically observable indicators. A key assumption here was that publications can reflect all the indicators of interest – track record, productivity, interdisciplinarity, riskiness, and peer skepticism. Indeed, there is a nonzero probability that truly transformative research is not able to pass peer review muster, or find journals that are willing to publish research that includes ideas that are at odds with prevailing wisdom, reflect true cognitive integration (a more conceptual measure of interdisciplinarity as compared with the I-score), or reflect quality. As a specific illustration, I have assumed that interdisciplinarity of publications is a close-enough proxy for the interdisciplinarity of research. It is very likely that researchers are bringing

25 NIH RePORT accessed here: http://report.nih.gov/index.aspx
26 NSF's award search database accessed here: http://www.nsf.gov/awardsearch/

interdisciplinarity in their research (for example, through formal or informal interactions with colleagues from other disciplines) without it reflecting in the subject area distribution of their publications. As a result, testing propositions using published work may not be the best way to judge transformativeness of research. Publications being the "currency" of science, however, these are what will be primarily (though not exclusively) used to explore the propositions.

Going beyond the limitation of using publications as a proxy for research, there are other limitations related to the use of proxies that are threats to measurement validity. For example, I have assumed that citations to research publications are a good proxy for community acceptance and buy-in. The propositions explored peer disagreement in many ways, one of which is citations to the publications of researchers conducting transformative research. This is not a bad assumption, but it is an assumption nevertheless that should be made explicit upfront.

I also assumed that an integration score is a good proxy for interdisciplinarity. That may not be the case. The calculation of these scores is based on the diversity of subject categories assigned to journals (by Thomson Reuter's Institute for Scientific Information (ISI)) and the degree to which a subject category (SC) relates to a particular journal. This is somewhat problematic. First, the validity of the I-Score is dependent on the accuracy of the publication dataset used to describe a PI's publication history. The I-Score analysis used the publications from the Web of Science (WOS) database, which generally does not include a comprehensive collection of conference papers and proceedings and the coverage varies by research area. This point is particularly important in calculating the proposal I-Score from cited references, since many of the PI records contained conference proceedings that were not included in the WOS database and, thus, not matched to a SC. Second, the I-Score

calculation is based on the diversity of SCs matched to the journals of a publication dataset; however, the accuracy of the matched SC(s) to classify a journal may vary (Boyack, Klavans, and Borner 2005). Boyack notes that for approximately 50 percent of the SCs, there is a high level of accuracy regarding the matched journal and SC. For the other 50 percent, there is less accuracy in the attribution.

While most operationalizations have one limitation or another, the way I have addressed these limitations is by using multiple indicators as well as multiple data sets (as reflected in Table 5-5). For example, to capture the skepticism, I look not only at the citation uptake (based on publications) but also ratings of reviewers at time proposals were submitted (not based on publications), and the ratings of external evaluators after the end of the award period (again not based on publications). This multi-method strategy has been helpful in addressing the issue of measurement reliability as well. For example, some of the assessments are based on expert evaluators' ratings of projects. This is a subjective approach. Presumably a different set of evaluators might have given the research different ratings. In this case, examining publication-related metadata, such as citations or other indicators such as citation velocity (that measures how long after publications the citations appear), provides an independent assessment of peer skepticism of research. Other limitations related to the reliability of measurement points to my data sources – the Thomson Reuters database that was used as a source of all publications may not be accurate, and include publications that don't belong, or exclude publications that do.

But the two most important limitations that pose threats to measurement validity relate to the dataset used and the timeframe of the analysis. With respect to the dataset, I have assumed that Pioneers are conducting transformative research, and the matched R01 set grantees are not. There are no guarantees that the Pioneers are necessarily conducting

transformative research. Indeed, all we know is that they submitted what they believe to be transformative research proposals, and those reviewing the proposals concurred with them. Similarly, there are no guarantees that researchers in the comparison group are not. For the purpose of this dissertation, however, it was assumed that the test group, funded to conduct transformative research, conducted transformative research and the comparison group did not.

With respect to timing, the challenge is that, as explained in Chapter 2, transformative research is often not evident for years (even decades) after being conducted. This analysis was done for research only five years after award of funds. There is a good chance, therefore, that at least for some of the research, the evidence base – whether the number of citations or expert judgment as to transformativeness - is absent. Conversely, research that is being called transformative, will, in fact, be shown to be incorrect in a few years. And some of what I am measuring is noise rather than signal.

With respect to internal validity, I have made no attempt to establish cause-and-effect. For example, I did not claim that it is the interdisciplinarity of research that causes it to be transformative. I merely examined relationships: is research that is considered transformative interdisciplinary, and conversely, is research that is not, not interdisciplinary. While I did not attempt to establish formal linkages or pathways, I did attempt to seek relationships. As a result, the composition of the comparison group was critical. Every attempt was made to ensure that the researchers in the comparison group were as similar as possible to the NDPA group – they were matched on areas of research, demographics, and other attributes of importance. However there is still a chance that the matched set of R01 grantees was not the best comparison group. This would imply that any differences or lack

of differences cannot be attributed to the "treatment" – which in this case is the receipt of the NDPA award.

Some of the analyses around risk and peer disagreement and skepticism had to be conducted without a comparison group. This further limits my ability to make claims about the nature of transformative research. For example, I found that peers disagreed about the transformativeness of NDPA-funded research. Had these peers looked at the matched R01 grantees' research, they may have disagreed to similar levels. As a result, at least with respect to this measurement, I cannot say if there is greater disagreement for transformative research as for mainstream research.

Lastly, with respect to external validity, the limitation is simple. All analyses are conducted for a single research program – even though there is a well-constructed comparison group in place – within biomedical research. There is no attempt made to generalize it to other areas of research (such as physical science, engineering, social sciences), which surely have very different norms. Strictly speaking, the propositions' veracity is limited only to the biosciences. Similarly, the program examined supports basic research. I cannot be sure if the findings apply to applied research or technology development. Chapter 7 discusses recommendations to the design of a future study in a different domain of research.

5.7 Summary

In this chapter, I discussed the method used to explore the propositions introduced in Chapter 4. The NDPA program – a transformative research funding program at the National Institutes of Health - was selected as a data source primarily because of the ease of obtaining data on transformative research. Table 5.2 summarizes the approach to addressing each proposition.

A key activity was the creation of a comparison group that matched the NDPA PIs in important ways. Researchers in the comparison group wrote proposals in similar areas of research, received their grants in the same years as the Pioneers, were employed by similarly prestigious institutions, and had similar degrees and the same age distribution as the NDPA grantees. This comparability of PIs and their backgrounds ensures that any differences in outcomes could potentially be attributed to the fact that one group is likely conducting transformative research and other is likely not.

Lastly the limitations of the analysis are listed. Principal among those is that the analysis is weak on showing internal validity – this dissertation explores cause and effect, but does not aim to establish causality. Furthermore, given that I did not attempt to generalize beyond the one program of interest, the approach also lacks external validity.

Table 5.2

Proposed Approach to Testing Propositions

Proposition	Specific Probe for the Data Set	Data Source and Analytic Approach				
		Admin. Data	Web of Science Database	Expert Reviewer Interviews	PI Interviews	Expert Evaluator Interviews
Researchers conducting transformative research tend to have a track record of conducting transformative research	How different do the Pioneers look compared to researchers conducting "merely excellent research"?		x			
Researchers conducting transformative research tend to be young	Are Pioneers younger than the researchers who got mainstream grants in the same time period?	x				
Researchers conducting transformative research tend to be prolific publisher	How well-published are the Pioneers compared with researchers in the comparison group?		x			

Proposition	Specific Probe for the Data Set	Data Source and Analytic Approach				
		Admin. Data	Web of Science Database	Expert Reviewer Interviews	PI Interviews	Expert Evaluator Interviews
Transformative research tends to have a high risk [of failure]	Was the level of risk (as rated by reviewers of proposals) of the awardees greater than that of non-awardee applicants? Was there a correspondence between what reviewers considered transformative proposals and their level of risk? Did the riskiest projects show transformative outcomes? Conversely, did projects with transformative results receive the highest risk rankings? *(no comparison group)* How far did the Pioneers venture with respect to exploring emerging ('non crowded") research areas?	x		x	x	x
Transformative research tends to be interdisciplinary (IDR)	To what extent do publications funded by the test program researchers have higher interdisciplinarity as compared with publications funded by the comparison program?		x			

Proposition	Specific Probe for the Data Set	Data Source and Analytic Approach				
		Admin. Data	Web of Science Database	Expert Reviewer Interviews	PI Interviews	Expert Evaluator Interviews
Transformative research garners disagreement at all stages, and may take longer to be accepted by the community	To what extent did external reviewers agree with each other in the review of applications?	x	x	x		x
	To what extent did external evaluators reviewing completed research agree with each other? *(no comparison group)*					
	How do the grant-related citations of researchers in the test group compare with those of grantees in the comparison group?					
	How fast did the publications of the test group increase as compared with the comparison group?					
	How fast did citations of the test group increase as compared with the comparison group?					

6.0 Exploring the Propositions

This chapter explores the six propositions introduced in Chapter 4 using the methods described in Chapter 5. In each of the next six sections, I take each proposition, first describe the specific method used to explore it, and then present the findings regarding whether the data support the proposition. The final section summarizes the findings.

6.1 Proposition 1: Track Record

In chapter 4, I introduced the proposition that researchers conducting transformative research tend to have a track record of conducting transformative research. In this section, I explore the differences between researchers conducting potentially transformative research and the 35 matched R01 researchers. The latter group has been selected such that they can be considered mainstream researchers conducting "merely excellent research."

6.1.1 Review of Method

As mentioned in previous Chapters, it is difficult to explore if NDPA grantees have a track record of prior transformativeness, mainly because there are no extant "indicators" of transformative research (this discovery is precisely the purpose of this dissertation). To address this proposition, I therefore look for any systematic differences between the transformative and non-transformative population, posing the question: How different do the Pioneers look compared to the matched R01 researchers who are conducting "merely excellent research"?

To address this question, I explored how their prior research records differ with respect to three indicators of interest. The first was productivity (number of publications) in the years between 1980 and year of award. The second was quality of work (as measured by the number of citations to publications) of prior research, in the years between 1980 and

year of award. The third was the researcher's impact as measured by the prestige of the journal in which the publication appeared, and as measured by his or her lifetime h-index.

As discussed in Chapter 4, publications and associated data (which includes citation counts and h-indices) for the 69 researchers were downloaded from the Thomson Reuters Web of Science database. Journal impact factors (SCImago journal rank, or SJR) were taken from SCImago (2007), a free resource that provides a ranking of all journals tagged in the Scopus database.

6.1.2 Findings

Figure 6.1 through Figure 6.4 plot the indicators of interest for both groups. The Figures show that both Pioneers and matched R01 grantees as a group had similar records over the time between 1980 and the year they received NIH grants. Since the comparison group researchers were *not* selected for conducting high-risk, high-reward research, it is interesting to note their similarity to the Pioneers.

These findings do not support the proposition that transformativeness is necessarily evident in the prior works of researchers conducting transformative research, and it is unclear if researchers conducting transformative research are any different from researchers conducting "merely excellent research."

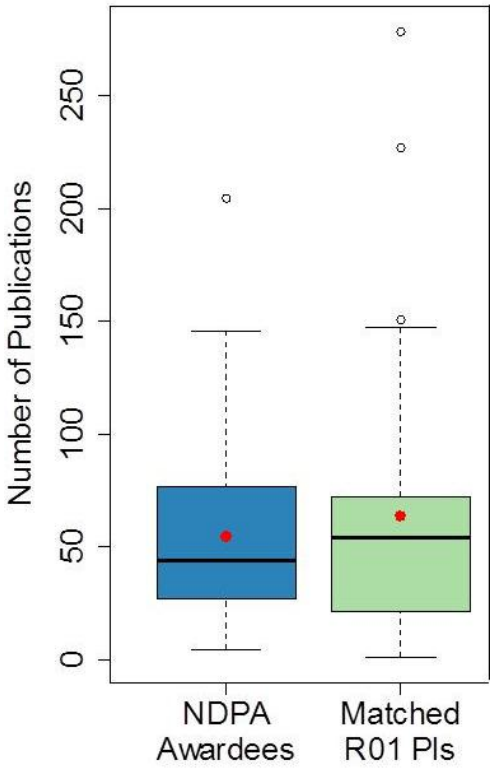


Figure 6.1. Productivity of Pioneers and comparison group researchers, 1980 year of award Kolmogorov-Smirnov p = 0.82

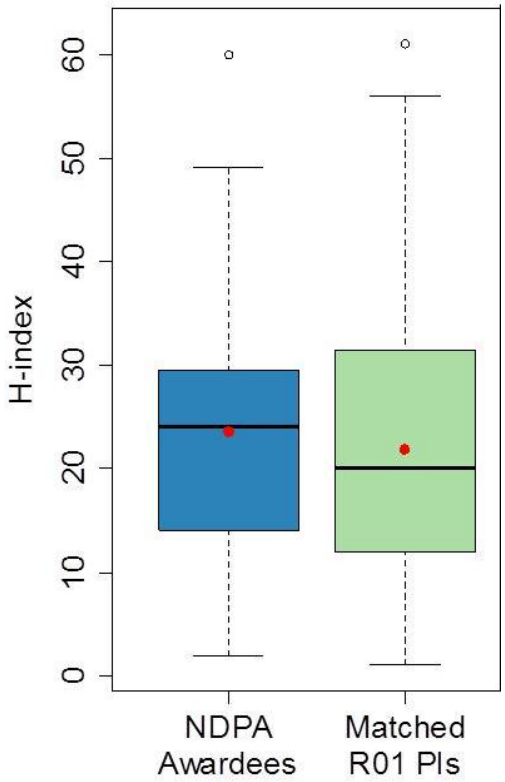


Figure 6.2. H-Index of Pioneers and comparison group researchers, 1980 year of award Kolmogorov-Smirnov p = 0.60

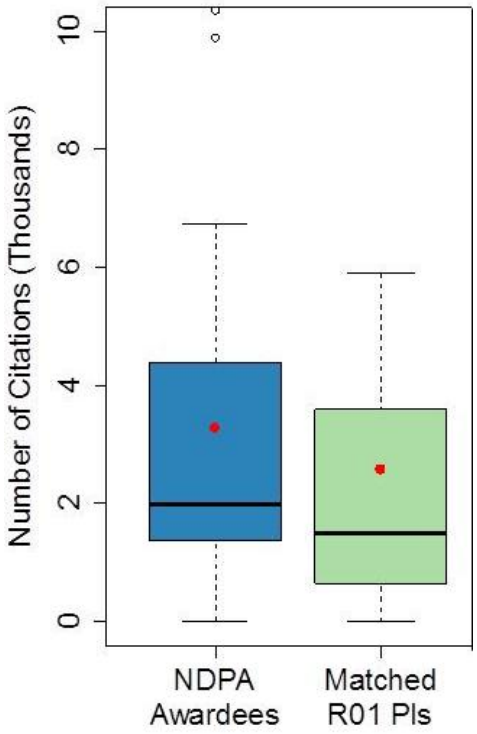


Figure 6.3. Citations of Pioneers and comparison group researchers, 1980 year of award
Kolmogorov-Smirnov p = 0.28

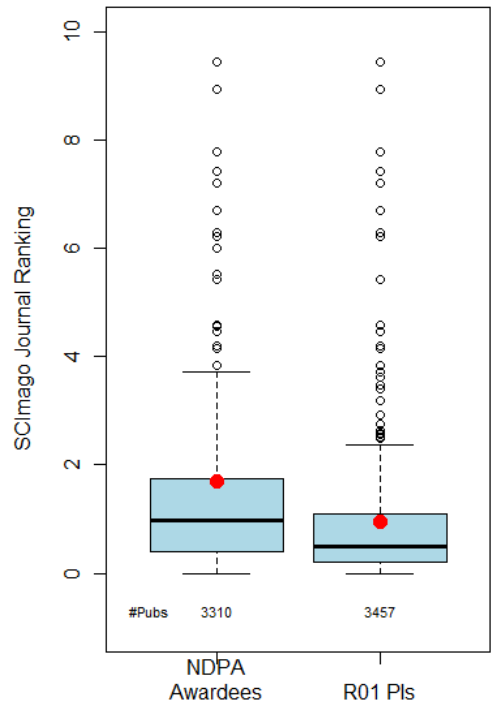


Figure 6.4. Journal impact factors of Pioneers and comparison group researchers, 1980 year of award
Kolmogorov-Smirnov p = 0.28

6.2 Proposition 2: Youth

As discussed in Chapter 4, it is commonly believed that transformative researchers tend to be young. In this section, using years since Ph.D. as a proxy for age, I explored whether researchers conducting potentially transformative research are younger compared with mainstream researchers who received competitive awards in the same years.

6.2.1 Review of Method

The primary comparison group – the matched R01 grantees – is excluded in this analysis because this comparison group was matched on age (or its proxy, years since doctoral degree). The age range for this group will therefore be similar to the NDPA awardees by design. There are three other comparison groups for this analysis.

The comparison group for this analysis therefore comprises the entire universe of mainstream researchers in the R01 program in FY 2004-2006 (n = 11,170). If the proposition is to be supported by data, these researchers should be significantly older than the Pioneers.

6.2.2 Finding

Analysis of the data from each of the groups revealed that the average age of the Pioneers is *not* different from all R01s. The histograms in Figure 6.5 and the boxplots in Figure 6.6 demonstrate these similarities and differences, as does Table 6.1. In other words, there is no particular link between age of a researcher (as measured by years since Ph.D.) and whether he or she conducts transformative research. This proposition is therefore not supported by data.

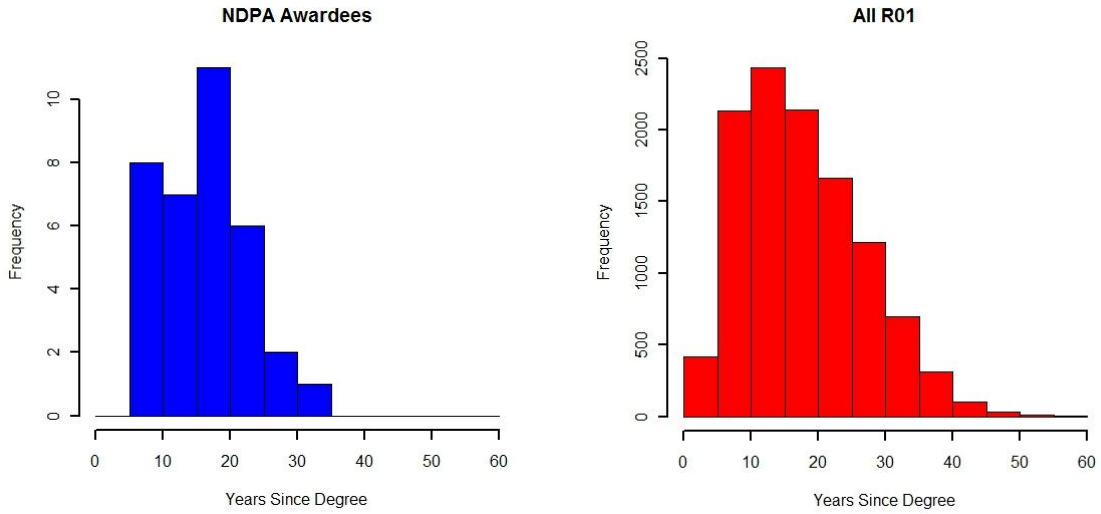


Figure 6.5. Distribution of years since Ph.D. of researchers in groups of interest

Table 6.1

Descriptive Age-Related Statistics for NDPA and All R01 Awardees

	N	Min	25th	Median	75th	Max	Mean	SD
All R01	11,170	0	11	17	24	58	18.1	9.0
Pioneers	35	7	11	18	20.5	31	16.8	6.2

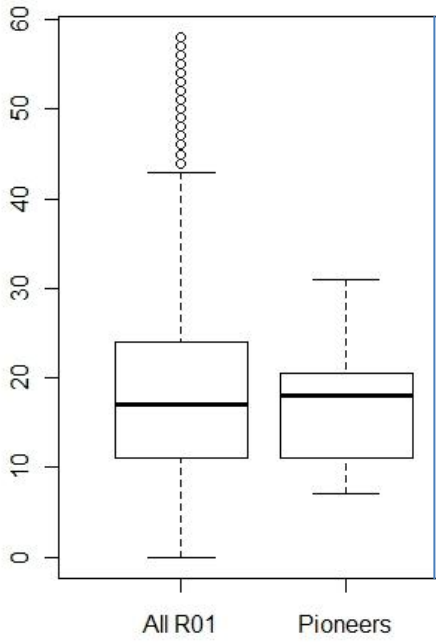


Figure 6.6. Distribution of years since Ph.D. for researchers by group

6.3 Proposition 3: Productivity

As discussed in Chapter 4, creative scientists are considered more productive than their peers, and many experts believe that the likelihood of a researcher's peers finding his or her work creative is a probabilistic consequence of quantity. In this section, using cumulative number of publications as a proxy for productivity, I explore whether researchers conducting potentially transformative research are more productive than a matched set of counterparts conducting excellent but mainstream research.

6.3.1 Review of Method

To explore this proposition, the number of publications of researchers conducting transformative research is compared with that of the matched set of mainstream researchers. Publication data on both NDPA Awardees and matched R01 grantees were collected from the Web of Science database, and synthesized.

6.3.2 Finding

On average, in the years before receipt of the NDPA, a Pioneer had produced 54 publications over his or her lifetime (with a median of 44). However, as might be expected, there is great variation in the number of publications within the group. Within the group, the number of publications over the career (since 1980) range from 4 to 205, as shown in the LHS box-and-whisker plot in Figure 6.7. Comparing the number of NDPA researchers' publications with the publications of researchers in the matched R01 group, however, shows no differences between the two groups. In other words, transformative researchers do not have higher productivity as compared with mainstream researchers who have similar backgrounds.

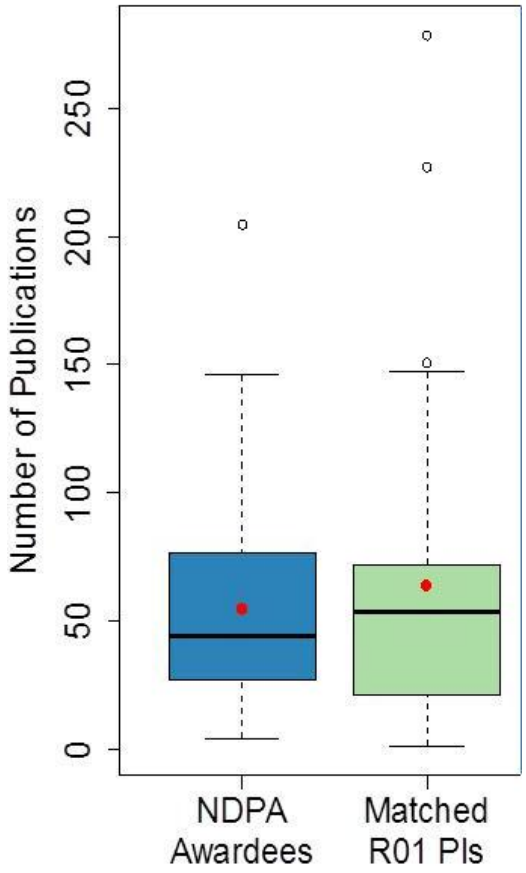


Figure 6.7. Researcher-level publications, 1980 start of grant
Kolmogorov-Smirnov p = 0.82

6.4 Proposition 4: Risk

As the overview of federal programs in Chapter 3 showed, there is an implicit belief that risk, whether technical or otherwise, is a core attribute of potentially transformative research. In this section, I explore this proposition in three different ways.

6.4.1 Review of Method

To examine whether transformative research tends to have a high risk [of failure], data at the proposal and outcome stages were analyzed to address three sets of questions:

The first set of question was: Was the level of risk (as rated by reviewers of proposals) of the awardees greater than that of non-awardee applicants? A related question here is: was

there a correspondence between what reviewers considered transformative proposals and their levels of risk? To address this set of questions, administrative data from the NDPA program were examined to assess if the Pioneers had "risk scores" that were higher than applicants who did not receive the award. The results are presented in Section 6.4.2.3. If it appears that Pioneer scores were higher than those of unfunded applicants, it would illustrate that reviewers believed risk to be a virtue when identifying potentially transformative research.

Because risk scores correlated with overall scores (strong positive correlation between risk and overall score with a Spearman's $r = .848$[27]), and those with the highest overall scores received the NDPA award, it is not logical to demonstrate using this analysis alone that transformative research is high-risk. It (i.e., transformative research) is by design [of the selection process]. I therefore explored the link between risk and project transformativeness (results presented in Sections 6.4.2.4). The questions posed were: Did the riskiest projects show transformative outcomes? Conversely, did projects with transformative results receive the highest risk rankings? In this analysis, risk-related scores were matched with project outcomes to assess if researchers whose proposals were deemed high risk subsequently produced transformative research. I also examine the fraction of projects that was considered as failure and how those projects were ranked in their riskiness. *There is no comparison group to this analysis, but it is insightful nonetheless, primarily because there is no known research that addresses this question.*

---

[27] The Spearman coefficient is used over the Pearson because Spearman is a rank-based coefficient, whereas Pearson is a linearly-based coefficient. With only 5 categories, it is difficult to justify the use of a linear coefficient.

Lastly, I explored the question: How far did the Pioneers venture with respect to exploring emerging ("noncrowded") research areas? To address this question, an NIH-based text mining tool (the methodology, discussed in Chapter 4 above, is based on a Bayesian statistical algorithm that automatically discovers meaningful categories from unstructured text) was used to identify the number of grants at NIH that were similar to the NDPA grants. Findings are presented in Section 6.4.2.5.

To provide richness to the analysis, I begin this section with a qualitative assessment of risk as determined by the Pioneers themselves, and by evaluators who examined their research five years after receipt of award (Sections 6.4.2.1 – 6.4.2.2).

6.4.2 Findings

6.4.2.1 PIs' Self-Assessment of Risk (Qualitative)

As part of the interviews conducted by Lal et al. (2011), the awardees were asked to characterize the risks in their NDPA proposal, following the typology of risk introduced in Chapter 2. Awardees most often characterized their NDPA proposals as incorporating multidisciplinary risks (Table 6.2).

Table 6.2

Awardees' Assessment of the Nature of the Risks of Their Research

Pioneer	Conceptual Risk	Technical Risk	Experience Risk	Multidisciplinary Risk	None of these risks
1		X	X	X	
2	X	X	X	X	
3	X	X	X	X	
4	X		X	X	
5	X	X	X	X	
6	—	—	—	—	
7	—	—	—	—	
8		X	X	X	
9	X	X	X	X	
10	X	X	X	X	
11		X	X	X	
12	X	X	X	X	
13	X	X	X	X	
14	X		X	X	
15	X	X		X	
16	X	X	X	X	
17		X	X	X	
18		X	X	X	
19	—	—	—	—	
20	X		X	X	
21	X	X	X	X	
TOTAL	13	15	17	18	0

Note. Dashes indicate that three awardees (6, 7, 19) could not be reached for comment regarding the nature of their risks. Adapted from: Unpublished analysis of NDPA Program data. The typology is explained in Chapter 2 (Colwell, 2003) Source: Lal et al. (2011).

In the interviews, the Pioneers spoke about the nature of risk taken through their Pioneer awards. One of the PIs, for example, remarked that her hypothesis that epigenetic mechanisms may contribute to "heritable changes in gene expression" is at odds with the prevailing idea that "all gene regulation is monitored by SNPs" and that diseases may be found by "looking at nuclear type changes." The PI also believed that although she continued to study genetics in her NDPA proposal, "human genetics is…different from plant genetics," and the shift required her to "read [new] literature and…collaborate and interact with [new] people," all of which took "an immense amount of…time."

When probed, another PI similarly indicated that her research incorporated technical, experience, and multidisciplinary risks. In the interview, the PI explained that the technique

she proposed for studying DNA damage response, the rapid uncapping of telomere ends, had never before been studied. In fact, the chemical techniques she first used in her attempt to "inhibit telomere function" did not work, so a *ts* mutant had to be developed instead. Her project also required her to build knowledge in DNA damage response, so an experience of risk was involved.

6.4.2.2 Experts' Post-Research Assessment of Risk (Qualitative)

In Lal et al. (2011), three experts were asked to assess the research of each of the awardees from FY 2004-2005, for a total of 63 experts (3 experts per 1 awardee). A feedback form was developed for the expert reviewers to characterize the Pioneer research and to assess the effects and value of the NDPA program. Experts who committed to participate were sent a research summary describing the Pioneer activities, three publications from the NDPA project, the NDPA program notice from the year of the award, and the feedback form to record their assessment of the Pioneer project. Following submission of the completed feedback form, phone interviews were held with selected experts, either because they requested to submit additional feedback or to clarify answers on the feedback form.

Each awardee's expert panel was asked to characterize the risk associated with the research they were evaluating. If at least two of the three experts agreed on a statement, the statement was reported as being chosen by the experts. The responses to this question are shown in Table 6.3.

Table 6.3

Experts' Assessment of the Nature of the Risks of the Pioneer's Research

Pioneer	Conceptual Risk	Technical Risk	Experience Risk	Multidisciplinary Risk	None of these risks
1	X				
2			X		
3			X		
4			X		
5				X	
6	X	X			
7	X			X	
8	X			X	
9	X			X	
10	X			X	
11		X	X		
12		X	X		
13		X		X	
14		X		X	
15		X		X	
16			X	X	
17			X	X	
18	X		X	X	
19		X	X	X	
20	X	X	X	X	
21	X	X	X	X	
Total	8	9	11	14	0

Note. This table displays consensual expert assessments (at least two of three experts in agreement) of Colwell typology risks for each Pioneer. Experts were able to select as many of the risks that applied. Source: Lal et al. (2011).

The most common risks the experts associated with the awardees' proposals were multidisciplinary risk (14/21) and experience risk (11/21). Experts found many of the applicant proposals to be multidisciplinary, in that they incorporated information from a number of different fields and used a combination of different approaches. Experts also agreed that all of the Pioneer proposals incorporated at least one form of risk. These results suggest that all of the funded investigators and projects can be considered risky, although the nature of the risk varied. Below is a selection of comments from experts regarding the awardees they reviewed:

> "The expertise needed to meet the evolving goals of this project required flexible access to a wide variety of potentially changing co-investigators."

> "The work is truly interdisciplinary, incorporating sophisticated theory and mathematics, but also providing serious connection to and adherence to the constraints of experimental data."
>
> "By the nature of the work and its outcomes, new avenues of investigation beyond the investigator's expertise ensued."
>
> "This NDPA application involved a unique combination of antigenic mapping, molecular phylogenetics, and fitness measures. [The awardee] is the only one doing this sort of research, and he has provided fundamental new insights into the basic biology of influenza virus."

Returning to the research of the two PIs identified above, three out of three experts who were asked to review her/his work agreed that the first PI's research contained conceptual, experience, and multidisciplinary risks:

> "Paramutation was thought to be an obscure phenomenon restricted to plants for which there was no mechanistic explanation. The applicant has found that tandem repeats and RNA interference regulate trans-allelic silencing, an unprecedented window into the phenomenon.
>
> The NDPA award permitted Dr. Chandler, renown for her expertise in plant genetics, to work in the area of mammalian genomics and perhaps even disease."
>
> "Chandler is one of only a very few who could have brought the disciplines of plant biology, mammalian biology, genetics, epigenetics, and molecular biology all into focus at once."

For the second Pioneer, two experts recognized the novelty of the PI's approach to studying DNA damage response (i.e., in using telomere inactivation) and stated that her research emphasized experience risk:

> "The time-lapse microscopy analysis of uncapped telomeres… are novel in the field and probably had to be first established in the de Lange lab."
>
> "One fundamental idea…is that the outcome of DNA damage signaling is cell cycle arrest. De Lange revealed an unexpected role for DNA damage signaling and chromatin mobility in DNA repair."
>
> "The detailed insights demonstrated by Dr. de Lange's work…go beyond the expected or standard expertise of a telomere biologist and would have

required an extensive immersion in the general genomic DNA damage response."

"The idea of using controlled telomere inactivation as a mode of inducing site-specific DNA damage as a model for the general DNA damage response is a novel one that stands slightly 'outside the box.'"

6.4.2.3 Risk Scores at Time of Award

To assess if NDPA-funded research was viewed as having higher risk than unfunded research, the risk and overall scores for all NDPA applicants in 2004-2008 were examined.

As discussed in Chapter 5, external experts who reviewed NDPA proposals made subjective judgments in their assessments and gave four scores to each application they reviewed, one of which was "suitability." The term "suitability" was defined in the NDPA RFA as "evidence that proposed project is of sufficient risk/impact to make it more suitable for NDPA than for traditional NIH grant mechanism." (NDPA, 2007). Because traditional R01 applications are not ranked on a "suitability" scale, the comparison group candidates here are not mainstream grants but rather unfunded applicants to the NDPA program.

There were 1,816 scored applications over the 5 year period. Because 2004 applications were judged on a scale from 1-7 and subsequent years were on a scale from 1-5, scores from 2004 were converted to a 1-5 scale for all analyses. Table 6.4 shows the transformation from 2004 scores to 2005 and subsequent years.

Table 6.4

Conversion of 2004 Scores to 2005 Equivalents

2004	2005 Equivalent
1	1
2,3	2
4	3
5,6	4
7	5

The first finding in this analysis was that there is a difference in the distribution of risk scores between awardees and nonawardees: awardees received higher risk scores than nonawardees. Since this is a non-gaussian population, significance was determined using a chi-squared test (X^2=283.3, df =4, p-value ≈ 0). As Figure 6.8 illustrates, the awarded Pioneers tended to have relatively high risk scores. The average risk score for awardees is 4.4, as opposed to 3.1 for non-awardees.

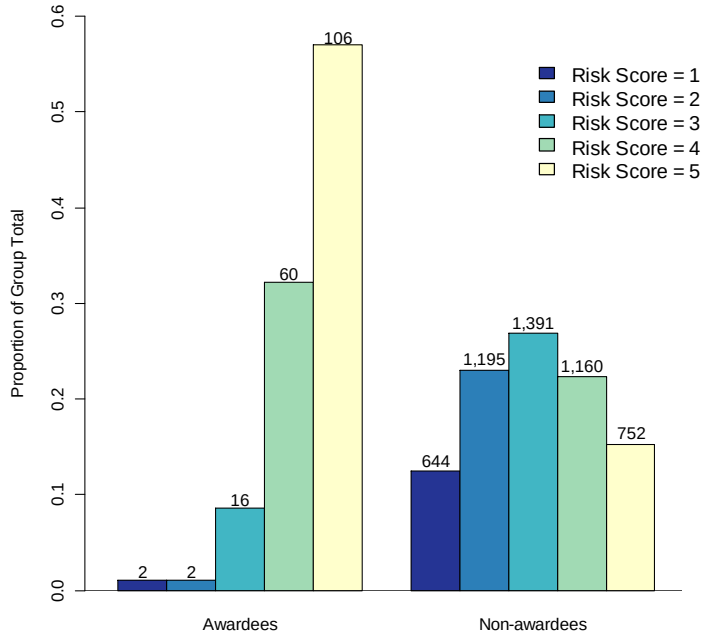


Figure 6.8. Distribution of risk scores for awardees and nonawardees

A second sub-analysis was performed. To examine whether very risky proposals are more likely to be selected for an award than less risky proposals, "very risky" proposals were defined as those with an average risk score *greater* than four. Given this definition, there were 218 very risky proposals, or 12% of the total pool. The second finding is that very risky proposals are significantly more likely to be selected for an award than other proposals. 22.3% of all very risky proposals, or nearly one in four, were selected for an award. On the

other hand, only 0.7% of not very risky proposals were selected for an award. Significance was determined using a chi-squared test (X^2=284.7, df=1, p-value ≈ 0).

A third finding centers on the use of data on "top 4" and "ideal candidate" flags. As discussed in Lal et al. (2010), each reviewer was asked to flag proposals that they felt were ideal for NDPA funding (all funded grants had at least one of these flags). Observing the fraction of high risk proposals that were flagged will illustrate if reviewers consider proposals that are NDPA-worthy to be also risky.

There is a significant relationship between very risky proposals and top-4/ideal candidate selections. Significance was determined using two permutation t tests – one for the top 4 category and one for the ideal candidate category (p-value ≈ 0 in both cases). Very risky proposals were more likely to receive top-4 and ideal candidate votes than other proposals. 48.6% of the riskiest proposals received a top-4 vote, whereas 31.6% of nonriskiest proposals received a top-4 vote. Similarly, 70.4% of riskiest proposals received an ideal candidate vote, whereas 15.5% of nonriskiest proposals received an ideal candidate vote.[28] Figure 6.9 shows the distribution of top-4 and ideal candidates for very risky and all other proposals. It is clear from this analysis that reviewers believed that the best NDPA proposals were the riskiest ones too.

[28] Ideal Candidate votes were only issued in 2008.

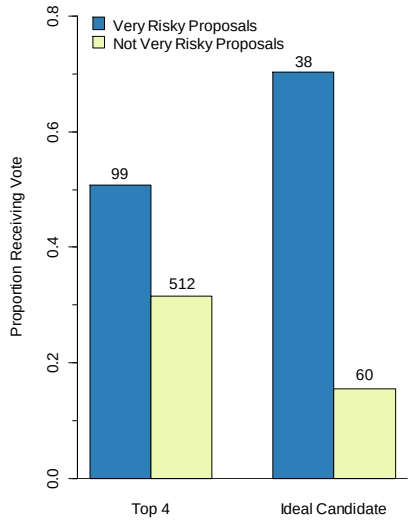


Figure 6.9. Relationship between risk and whether a proposal was an ideal candidate for award

This section shows that selected proposals received higher risk scores that nonselected ones, that very risky proposals are predisposed to being funded, and that risky proposals are associated with being NDPA-worthy. In other words, reviewers considered risky research as being associated with potentially transformative outcomes.

However, because risk scores correlated with overall scores (strong positive correlation between risk and overall score with a Speaman's $r = .848$), and because those proposals with the highest overall scores received the NDPA award, it is not easy to demonstrate using this analysis alone that transformative research is high risk. It (i.e., transformative research) is, by design of the NDPA program. The following section therefore, explores the links between risk at time of award and project success after completion of project.

6.4.2.4 Linking Risk with Transformative Outcomes

Returning now to the question of connecting risk with transformativeness, and risk with failure, I begin by presenting data on how transformative (or using the language of NDPA, "pioneering") the Pioneers' research was, and then connect this data with the risk

rankings at time of award. Figure 6.10 presents the experts' views on the transformativeness or pioneeringness of NDPA-funded research. Each of the three experts recruited to review a Pioneer's work was given the statement: "The accomplished research is pioneering" and then asked to either *agree*, *strongly agree*, *disagree*, or *strongly disagree* with the statement for each Pioneer whose research they were evaluating. The Figure shows that at least two out of the three experts on a panel strongly agreed that the research was pioneering for 14 of the 21 awardees from the first 2 cohorts of the program. In two cases, two out of the three experts on a panel moderately agreed that the research was pioneering; in four cases, there was disagreement across the panel; and in one case, all three experts on the panel strongly disagreed that the research was pioneering. Only for four of the Pioneers, were all experts in agreement that research was pioneering. Figure 6.11 provides the average score for each Pioneer.

To test if the riskiest researchers performed the most transformative research, next I looked at how perceived risk at time of award mapped on to actual performance of transformative or pioneering research at conclusion of award. This analysis is qualitative and conducted without a comparison group. While the numbers here are small (case studies were performed on the first two cohorts of the Pioneers, with an n=21), such the analysis has never before been performed before, and is worth the effort

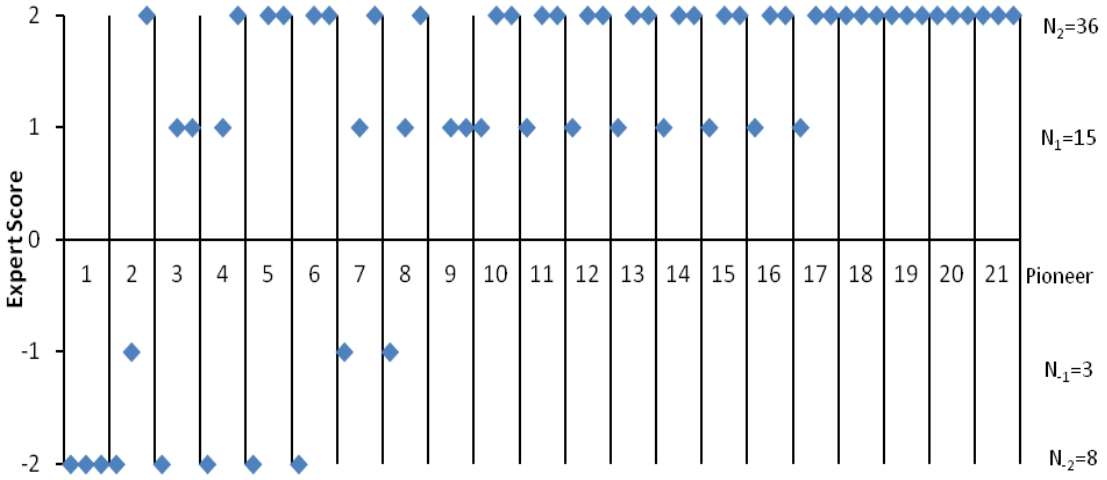


Figure 6.10. Expert scores on the Pioneeringness of NDPA research

Note: Each data point represents the assessment of a single expert. Experts were asked to rate this statement on a scale where 2 is strongly agree, 1 is moderately agree, -1 is moderately disagree, and -2 is strongly disagree.
Source: Lal et al. (2011)

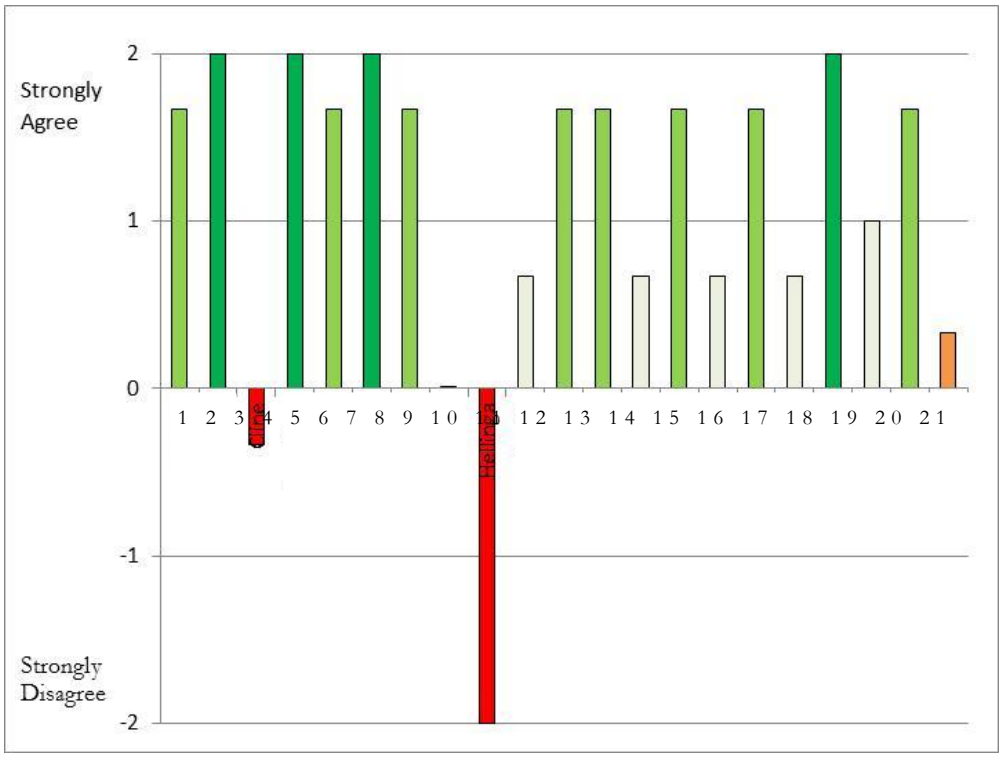


Figure 6.11. Averaged ranking of pioneers by experts

Table 6.6 translates the insight from Table 6.5 into a more compact matrix, which shows instantly that high risk scores did not (by themselves) map onto pioneeringness—and success was evenly divided between high-risk and less risky projects. Indeed as Column 3 shows, of the ten researchers whose research was judged pioneering by at least 2 of the evaluators, eight had not received high risk rankings. Table 6.5 matches the PIs' proposals' risk ratings at the time of award to their pioneeringness ratings five years after receipt of award. The second column in the Table shows that of the 21 Pioneers for whom post-project pioneeringness data are available, only six had, at the time of award, received the highest risk rankings by all three of the evaluators who reviewed their applications.

Table 6.6 translates the insight from Table 6.5 into a more compact matrix, which shows instantly that high risk scores did not (by themselves) map onto pioneeringness—and success was evenly divided between high-risk and less risky projects. Indeed as Column 3 shows, of the ten researchers whose research was judged pioneering by at least 2 of the evaluators, eight had not received high risk rankings.

Columns 3 and 4 show pioneeringness rankings measured narrowly and broadly, respectively (narrowly refers to high agreement among experts, and broadly refers to lower agreement). Table 6-7 shows that of the six Pioneers whose research was considered high-risk, only two were seen as having conducted pioneering research by all evaluators after the fact. Using a broader definition of pioneeringness (agreement among two of three expert evaluators rather that no disagreement), five of these six had also conducted pioneering research. Seen another way, however, of the 14 researchers whose research was seen as less risky (comparatively speaking), nine nonetheless produced transformative outcomes (using the broader definition).

Table 6.5

Pioneer Pre-Award Risk and Post-Award Pioneeringness Rankings

Pioneer Code	**Was risk score high?**	**Did experts "strongly agree" if the research was pioneering?**			
		All	Two	One	None
1			X		
2	X	X			
3				X	
4		X			
5			X		
6	X	X			
7			X		
8					X
9					X
10			X		
11			X		
12			X		
13			X		
14	X		X		
15				X	
16			X		
17				X	
18		X			
20	X				X
21	X		X		
22	X			X	
Total	*6*	*4*	*10*	*4*	*3*

Table 6.6 translates the insight from Table 6.5 into a more compact matrix, which shows instantly that high risk scores did not (by themselves) map onto pioneeringness—and success was evenly divided between high-risk and less risky projects. Indeed as Column 3 shows, of the ten researchers whose research was judged pioneering by at least 2 of the evaluators, eight had not received high risk rankings.

Table 6.6

Mapping Perceived Risk with Pioneering Outcomes

	Pioneer Code			
	Considered Pioneering by All Three Evaluators	Considered Pioneering by Two of Three Evaluators	Not Considered Pioneering by Any of the Reviewers	*Number of pioneers*
Considered High Risk by	2	14	20	*5*

All Three Reviewers	6	21		
Not Considered High Risk by All Three Reviewers	4 18	1 5 7 10 11 12 13 16	8 9	*12*
Total number of Pioneers	*4*	*10*	*3*	

The final analysis conducted was around risk and failure. I examined if the riskiest projects had a high chance of failure. Defining failure in research is complex, because only in exceptional cases does one see true failure – in most cases, research has some outcomes that are fit to be published. In the case of the NDPA program, failure could be defined in two ways. First, researcher did not accomplish the goals of the research as stated in the proposals. A review of the case studies from Lal et al. (2011) showed that of the 21 researchers, six did not accomplish their goals. One could say, therefore, that NDPA had a failure rate of 29%. Given that the program did not have any failure targets, it is difficult to say if this percentage is high or low.

A second definition of failure in the context of a high-risk program like NDPA could be that after starting a high risk project, researcher switched to a "safe" or low risk topic. There is an added complication about making any pronouncements regarding failure. A PI may have failed to achieve his or her stated goals (as six of them did in NDPA), but still achieve pioneering results. In fact, if one were to examine expert evaluator rankings of the six Pioneers who did not accomplish their original objectives, two of them received the highest scores on pioneeringness. So in fact, from a program's perspective, only four awardees could be considered as having failed, a programmatic failure rate of 18%. Table 6.7 shows this result as a 2x2 quad table.

Table 6.7

Success Rates According to Accomplishment of Stated Goals

		Program Level	
		Successful PIs	Unsuccessful PIs
Project Level	Accomplished Stated Goals	All others (15 Pioneers)	N/A
			3
	Did not Accomplish Goals	5	8
	Stated in Proposal	6	9
			22

6.4.2.5 Pioneers' Foray into New Fields

Another measure of risk was the PI's predisposition to propose research in fields of research which were emerging, and therefore risky. Building on work conducted by Podolny (1996) and Okamura &Vonortas (2009), I computed the number of Pioneers that proposed research in fields that did not have many other researchers. In other words, I attempted to identify "brokers" that were forging new ground. Counting all grants with a similarity ranking of >0.4 (arbitrary cut off), I found that very few Pioneers proposed research in new areas.

Figure 6.12 displays the distribution of the number of NIH researchers who were doing research in the same areas as NDPA winners in the years 2008-2011, as compared with HHMI investigators selected in 2005 (this latter group was introduced as it was not feasible to compute similarity rankings of the matched R01 group). If risk were to be measured this way, one could claim that Pioneers' research was not particularly risky.

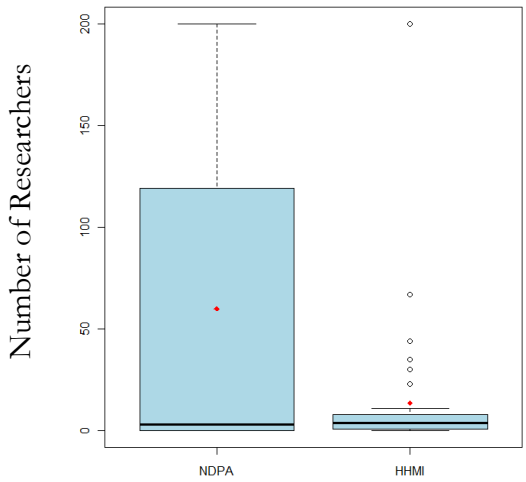


Figure 6.12. Distribution of the number of researchers in similar fields

Figure 6.13 plots the "crowding" of the Pioneers against their "status" in the field, as measured by the number of citations received five years before receipt of award (5-year citations are taken instead of lifetime citations to ensure comparability – since several of the researchers in the Pioneer group are older, they have a large number of citations that disadvantages younger researchers). Figure 6.14 maps "crowdedness" with another measure of status - the researchers' h-index – with the same outcome. The distribution shows that Pioneers are not proposing research in fields with fewer researchers. Using the Podolny (1996) typology, one can see that far fewer Pioneers could be considered breaking new ground, or being "brokers" as compared with being "leaders" in established fields. Using the typology, it could be said that the Pioneers are less likely to build on previously unexploited technology and provide a distinctive foundation for others. They appear to be more "leaders" in well-established technologies. It is worth noting that neither of the groups would be considered "followers" (those engaging in innovative activity in congested regions of the technological space).

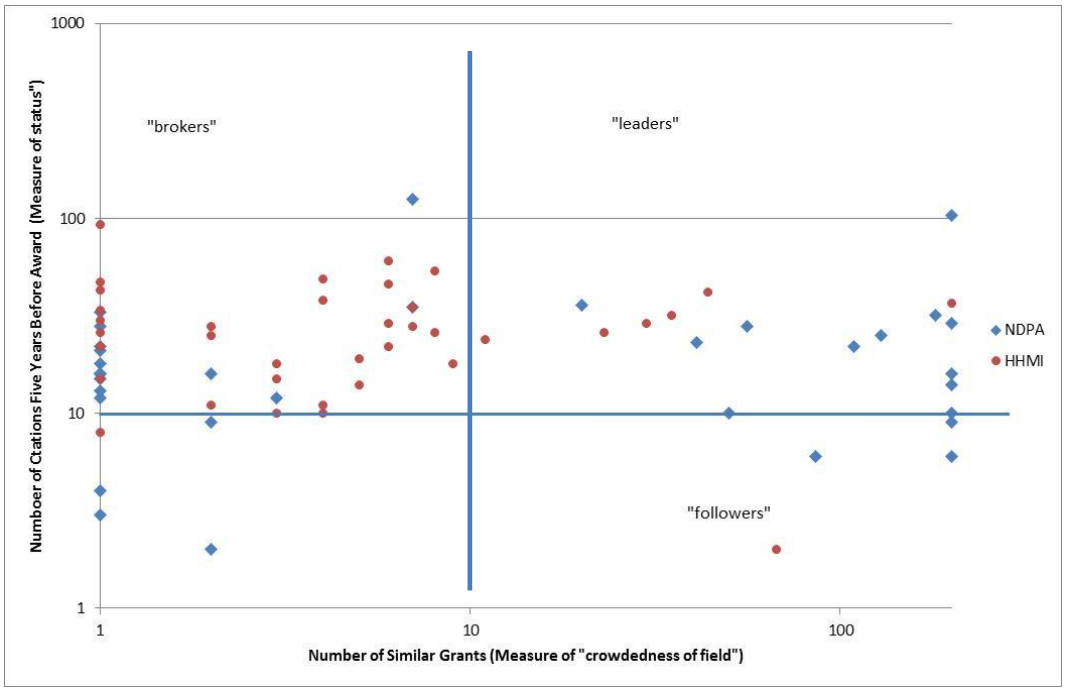


Figure 6.13. Mapping NDPA Citations with Crowdedness
Note: Researchers in the top left quadrant can be considered **brokers** of new technologies. They build on previously unexploited technology and provide a distinctive foundation for others. Those in the top right quadrant are **leaders** in well-established technologies. In the bottom right quadrant are **followers** engaging in innovative activity in congested regions of the technological space. (Typology from Podolny, 1996)

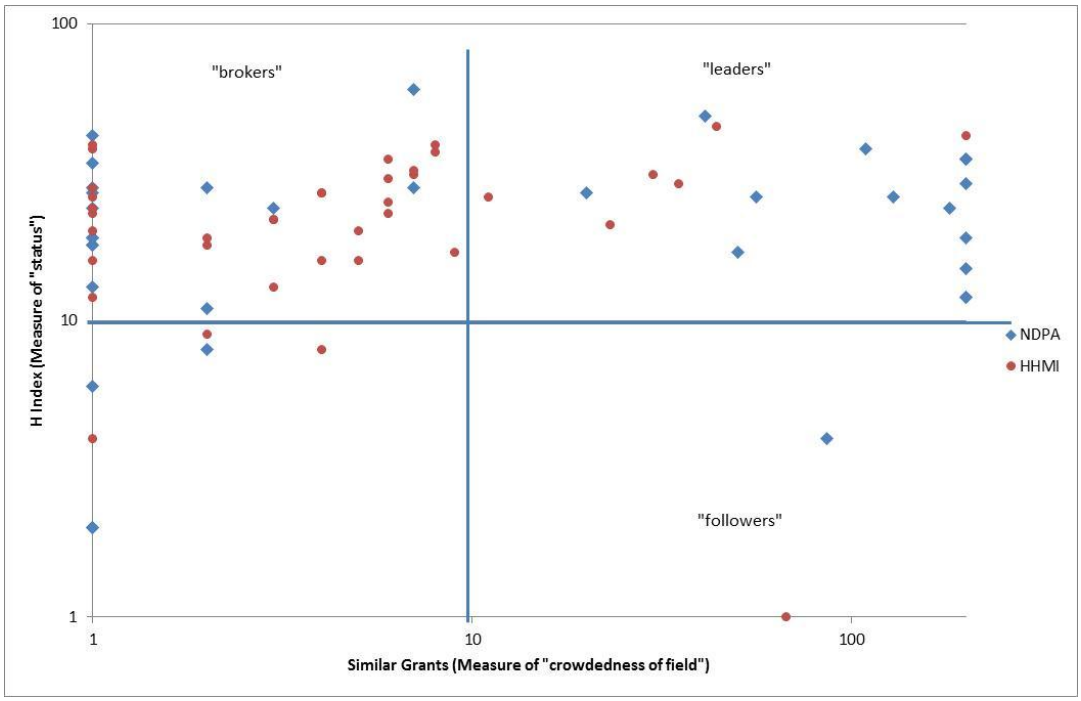


Figure 6.14. Mapping H-Index of Pioneers with Crowdedness
Note: Typology from Podolny (1996)

6.5 Proposition 5: Interdisciplinarity

As discussed in Chapter 4, interdisciplinarity is commonly considered a marker for transformative research, the underlying concept being that people who work at the intersection of scientific communities are more likely to be familiar with selecting and synthesizing alternatives into novel ideas. Likewise, a transformative discovery is more likely to be made when a novel connection is established between two or more previously disparate units of scientific knowledge. In this section, I explore interdisciplinarity as a predictor of transformative research, using three proxies of interdisciplinarity.

6.5.1 Review of Method

This proposition was explored by testing whether research considered transformative research is any more interdisciplinary than mainstream research. Three proxies for interdisciplinarity are used. The "integrations (I) score" measures the diversity of subject areas that are assigned as keywords in a given publication. "Specialization (S) score" measures the integration scores of the journal in which a publication appears, and the "diffusion (D) score" measures the integration scores of the publications that are *citing* a given publication.

As discussed in Chapter 5, to calculate the I-, S-, and D-scores of NDPA and mainstream publications that could be attributed to the grant mechanisms, the publications—the full record plus cited references—were exported into the text mining tool Vantage Point, which was used to assign subject categories to each journal as well as to compute the integration scores.

6.5.2 Findings

Table 6.8 shows the average integration, specialization, and diffusions scores of researchers in both groups of interest. These scores indicate NDPA research is not very

interdisciplinary. More importantly, comparing the scores across the two groups of interest, the test group and the matched R01 grantee set – the most similar researchers, indicates that there are no differences between the two groups on any of these dimensions.

Table 6.8

Average Specialization, Integration, and Diffusion Scores of Pioneers and Matched R01 Grantees

	Average Integration Score – Represents the IDR of the cited references	Average Specialization Score – – Represents the IDR of the publishing Journals	Average Diffusion Score – Represents the IDR of the papers citing research
NDPA Awardees	0.44	0.68	0.48
Matched R01 Grantees	0.48	0.69	0.47

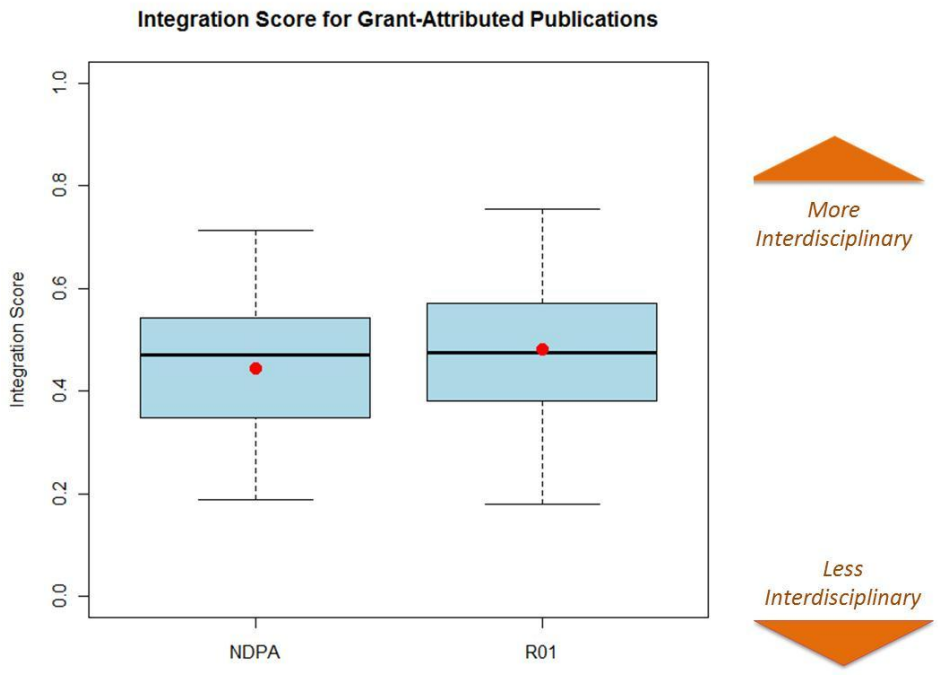


Figure 6.15. Integration Score of NDPA and R01 Programs

Figure 6.15 shows that NDPA publications have about the same or slightly lower integration scores as compared with the comparison group. The distribution of the interdisciplinarity similar to that of the R01 grants.

Figure 6.16 shows the specialization scores of NDPA and R01 publications. NDPA-funded research is similar to that of R01-funded research in that it is published in specialized journals (which may be a measure of lower interdisciplinarity). Figure 6.17 sheds light onto the diffusion of NDPA funded research. The boxplot shows that publications that *cite* NDPA-funded research are likely just as interdisciplinary as publications of the matched R01 set.

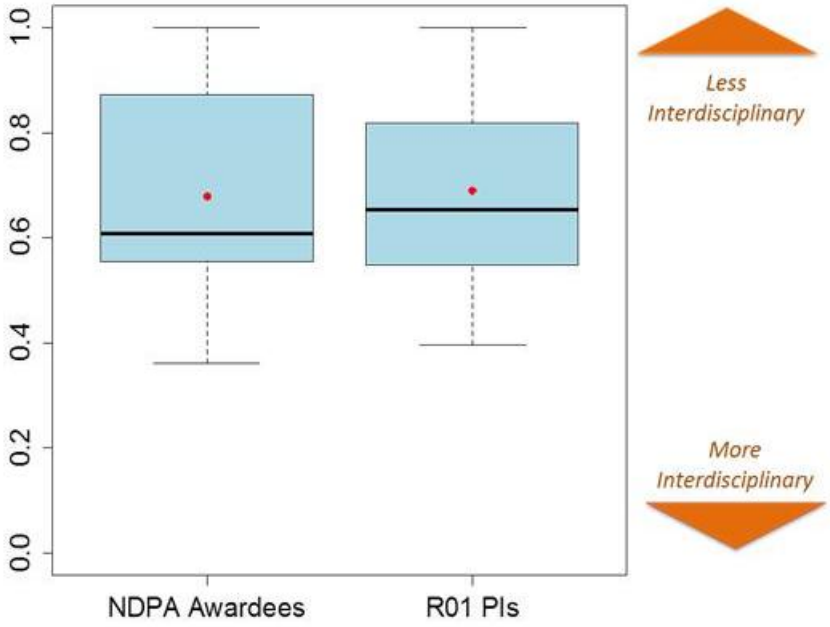


Figure 6.16. Specialization of grant-funded publications

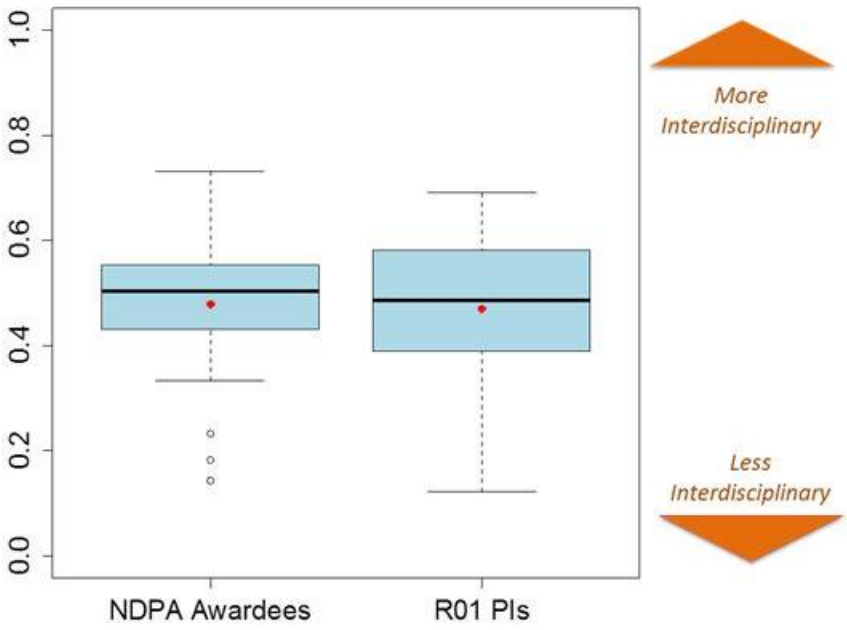


Figure 6.17. Integration scores of publications citing grant-funded research

Examining a paper's I and S-scores together is insightful as the quad-chart in Table 6.9 shows. Ideally, a Pioneer would not just building on interdisciplinary research (high I score) but also developing interdisciplinary outputs (low S score), or being a "renaissance integrators" using the typology introduced by Porter et al. (2007).

Figure 6.18 plots the PIs' integration scores against their specialization scores to explore if transformative researchers are "renaissance integrators." The results are contrary to expectations, and show that for both NDPA and R01 researchers, the higher their integration score, the lower the specialization score – indicating that both groups tend to be either **disciplinarians** or **single interdisciplinary specialists**.

Table 6.9

A Typology of Interdisciplinary Researchers

	Integration (I)		
Specialization (S)		Lo	Hi
	Hi		**Single interdisciplinary specialist** -

		Disciplinarian – a researcher with a sharply focused agenda	researchers who integrate many fields of knowledge into a few specialized areas
	Lo	**Grazer** - researchers who neither integrate knowledge from many different fields nor specialize into one field	**Renaissance integrator -** interdisciplinary researchers who integrate knowledge from many fields and publish in many fields

Source: Porter et al. (2007).

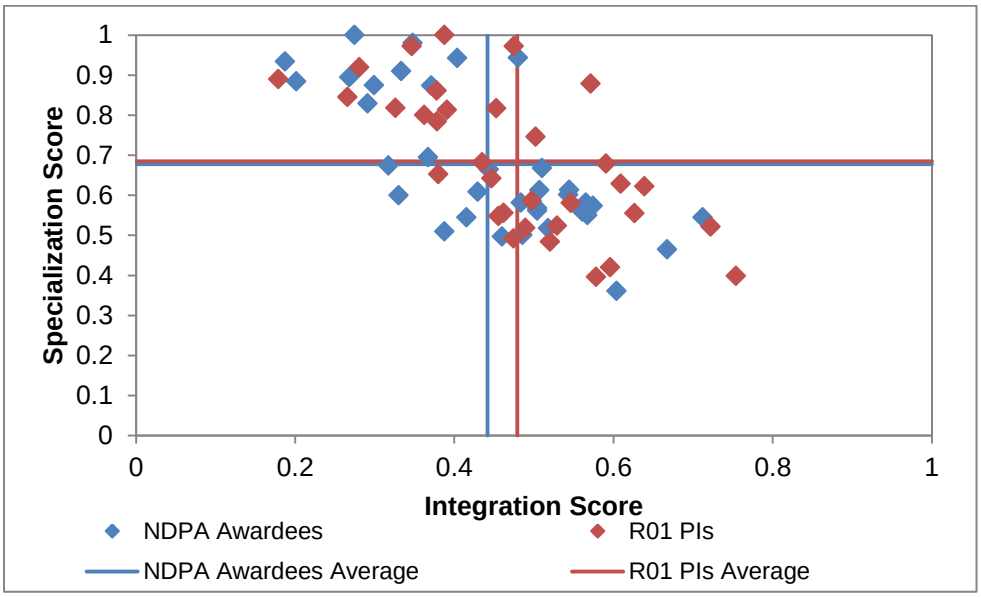


Figure 6.18. PIs' integration scores and specialization scores

Analysis of NDPA-funded and R01-funded publications shows that transformative research is not necessarily more interdisciplinary as compared with excellent mainstream research.

6.6 Proposition 6: Skepticism/Peer Disagreement

Chapter 4 introduced the proposition that research that contradicts prevailing wisdom takes longer to be accepted by the community, whether at the proposal stage where peers are reviewing it to assess its potential for its transformative potential, or at the postpublication stage where peers are assessing its actual impact. In this final section, I explore this proposition.

6.6.1 Review of Method

As outlined in Chapter 4, to explore skepticism or peer disagreement, I addressed four sets of questions. The first question was: To what extent did external reviewers agree with each other in the review of NDPA applications? To address this question, I examined disagreement among raters of NDPA proposals. The data was scoring data as obtained from Lal et al. (2010).

The second question was: How do the grant-related citations of the publications of researchers in the test group (NDPA) compare with those of grantees in the comparison group? To address this question, I examined how the number of citations of transformative researchers differed from those in the comparison group – presumably research that shifts paradigms is harder to penetrate the community, and there would be fewer citations to NDPA research. The data for this were obtained from the publicly available Thomson Reuters' Web of Science database.

The third question was: How fast did the publications of the test group increase as compared with the comparison group? To address this question, I examined how long it takes for researchers to publish – presumably controversial research has a harder time finding a venue. The data for this analysis also came from publicly-available publications from Thomson Reuters' Web of Science database.

The fourth question was: How quickly did citations to Pioneers' NDPA-funded research increase compared with publications of the matched-R01 group? To address this question, I examined how long it takes for publications to get cited – presumably research that shifts paradigms is penetrates the community at a slower rate than mainstream research. The data for this analysis also came from publicly-available publications from Thomson Reuters' Web

of Science database. Lastly, without a comparison group, I also looked at the disagreement among expert evaluators who identified which of the NDPA projects were pioneering.

6.6.2. Findings

6.6.2.1 Disagreement Among Reviewers–Proposal Stage

To address community acceptance, I begin at the proposal stage and examine the judges' agreement on the rating of suitability (for an NDPA award) across proposals. In the context of this analysis, the most suitable proposals were defined as those that received a suitability score of 5 at least once. Similarly, the least suitable proposals were those that received a suitability score of 1 at least once. Because very suitable proposals have a greater likelihood of having a high overall score, which have a greater likelihood of being selected for an award, one might expect high agreement among suitable proposals. Data show the opposite to be true. Using the Weighted Brennan Prediger Coefficient, an indicator of inter-rater reliability (Nunez, 2011), it was seen that "suitable" proposals have an inter-rater reliability score of .034[29] over the 5-year period (i.e., agreement was very poor). The least suitable proposals have an even lower score of -.083, indicating disagreement in suitability scores among judges. These two groups are well below the inter-rater score for all proposals of 0.322.

Table 6.10 summarize these findings that suggest that judges have a difficult time accurately judging the notion of "suitability," but in cases where all judges *can* agree, those proposals are likely to be selected for an award.

---

[29] Inter-rater reliability scores closer to 1 indicated better agreement among judges, a score of zero indicates no agreement, and scores less than zero indicate systematic disagreement.

Table 6.10

Inter-rater Reliability Scores for Different Groups of Proposals

Year	Proposals with High Suitability Scores	Proposals with Low Suitability Scores	All Proposals
2004	0.108 (n=106)	-0.468 (n=37)	0.272 (n=239)
2005-2008	0.022 (n=626)	-0.054 (n=493)	0.330 (n=1,577)
Weighted Average	0.034 (n=732)	-0.083 (n=530)	0.322 (n=1,816)

Qualitative data underscore this disagreement. In their scoring of NDPA proposals, reviewers were allowed to enter optional comments. A review of these comments reveals that there was a high degree of variability in scoring of applications. For example, one applicant was given overall scores of 5 and 1 and received the following two comments, respectively:

- "This is the most original proposal I saw, and it is by a PI who has a history of constant innovation. Though the idea seems very novel to me, [he/she] is in an excellent position to make great progress."

- "While [his/her] earlier work appears very innovative, this project cannot accomplish what is proposed."

Another applicant with overall scores of 5, 5, and 2 received the following comments:

- "Exciting proposal and very novel. This goes against current dogma for cancer treatment and if it works, it could be a major advance."

- "This is very much high risk science, but with a potential for very high gain. It is unlikely to be funded through other mechanisms. [His/her] letters of recommendation are exceptionally strong, and clearly indicate that this is an ideal proposal for a Pioneer award."

- "An old story that in this case may represent a unique animal model with little generalizability. Not really where [this research field] is likely to go."

Disagreement among raters was explored in another way. As Lal et al. (2010) discusses, the final reviewers binned finalist proposals in three categories: Must fund, fund if funds available, and do not fund. NDPA funded many of the applicants in the second two bins. And as Table 6.11 shows, about half the funded Pioneers were not the most highly recommended ones. One of the applicants placed in the "do not fund" bin, is one such researcher. As Rutgers University neuroscientist György Buzsáki says of his research: "It is a fantastic revolution. If [Pioneer name] doesn't do anything else, if he just sits in his office, he will get a Nobel Prize--there is no question in my mind." [30]

[30] http://www.openoptogenetics.org/images/b/bf/The_Light_Fantastic.pdf

Table 6.11

Proposals Recommended for Funding by Reviewers at Time of Award and their Postaward Rating

	Pioneer Code				
		All evaluators "strongly agreed" that PI was Pioneering	Two evaluators "strongly agreed" that PI was Pioneering	Only one evaluator "strongly agreed" that PI was Pioneering	No evaluator "strongly agreed" that PI was Pioneering
Applicants Highly Recommended for Funding (Top Bin)	1		1		
	2	2			
	3			3	
	5		5		
	6	6			
	8				8
	9				9
	10		10		
	11		11		
	12		12		
	21		21		
	22			22	
Subtotal	*12*	*2*	*6*	*2*	*2*
Applicants Not Highly Recommended for Funding (Middle and Bottom Bins)	4	4			20
	7		7		
	13		13		
	14		14		
	16		16		
	18	18			
	15			15	
	17			17	
	20				
Subtotal	*9*	*2*	*4*	*2*	*1*
Total Number	*21*	*4*	*10*	*4*	*3*

What is noteworthy in Table 6-11 above is that of the 14 Pioneers who were seen by postaward evaluators as conducting pioneering research, about half had not received the highest recommendations from the panelists.

6.6.2.2 Findings: Peer Recognition - Citations

If NDPA research is controversial, as one might expect paradigm-shifting research to be, the citation of NDPA-funded publication may be lower, or slow to grow. This section

reviews how the citations to NDPA publications compare with citations to R01-funded publications. As Figure 6.19 shows, the number of citations to NDPA-funded publications is comparable to the R01-funded ones, both in absolute terms, as well as when normalized per dollar.

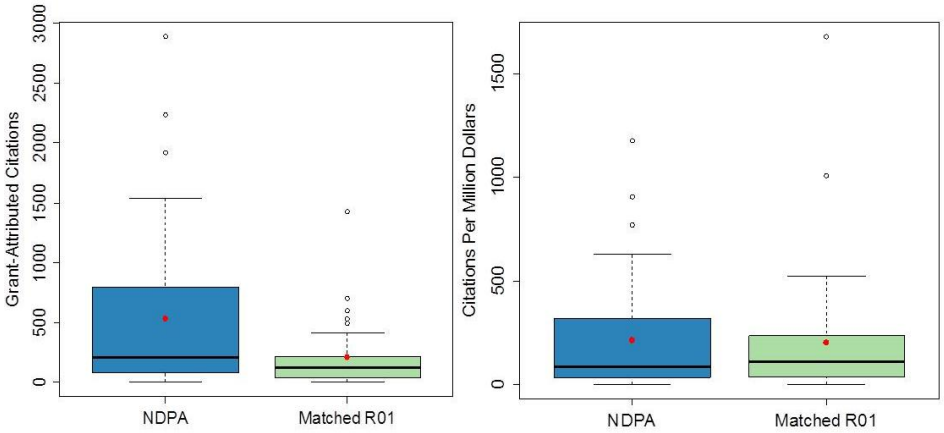


Figure 6.19. Grant attributed citations through 2011
Kolmogorov-Smirnov p = .06, p = 0.89

Breaking down these data by individual investigator could shed light on whether aggregating by program removes individual-level differences, and help explore if there are some Pioneers who are doing more controversial work than others. Figure 6.20 and Figure 6.21 present citation data at the PI level, and show the variability across the PIs for both groups of researchers but demonstrate no pattern that NDPA research was systematically more controversial than similarly prestigious researchers in similar areas of research.

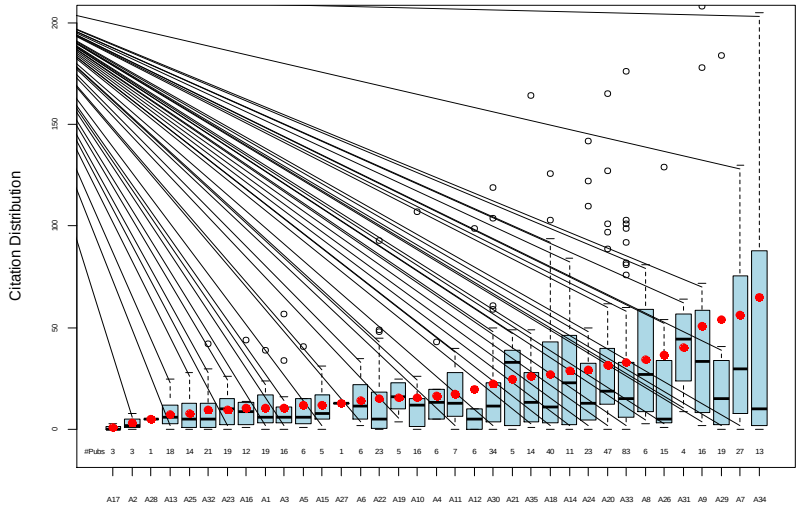


Figure 6.20. Citations to NDPA-funded publications, broken down by PI

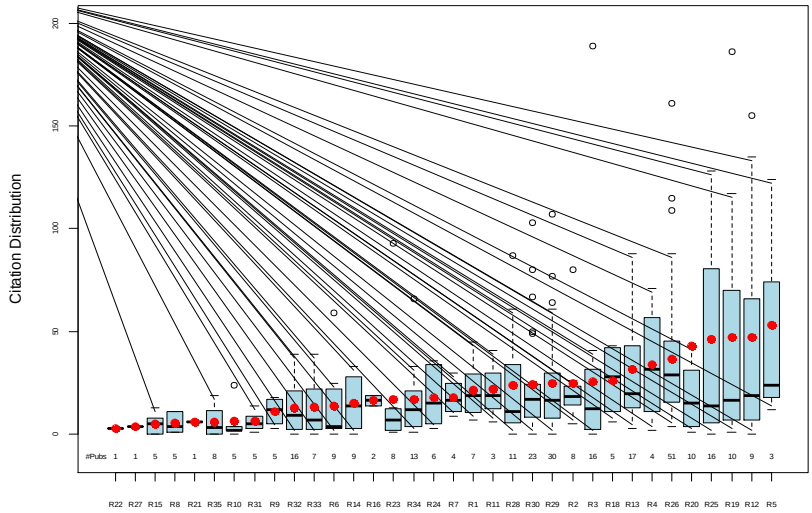


Figure 6.21. Citations to Matched R01-funded publications, broken down by PI

6.6.2.3 Findings: Growth of Publications and Citations

To see if NDPA publications take longer to appear (slower appearance might mean that the research is controversial, and therefore it is more difficult to find journals or publishers that would publish the findings), the time to publication was plotted for both

NDPA and R01 grants. Figure 6.22 plots these data and clearly shows that it takes NDPA grantees longer to publish than it does R01s.

While it may be tempting to believe that the delays originated in the controversy within the research conducted, it may be entirely possible that there are other reasons for this difference. Among other reasons, transformative research may simply take longer to conduct, and the slowness of publications is a consequence of the speed of research; R01 research may also be underway by the time a grant is received, and publications may thus emerge faster.

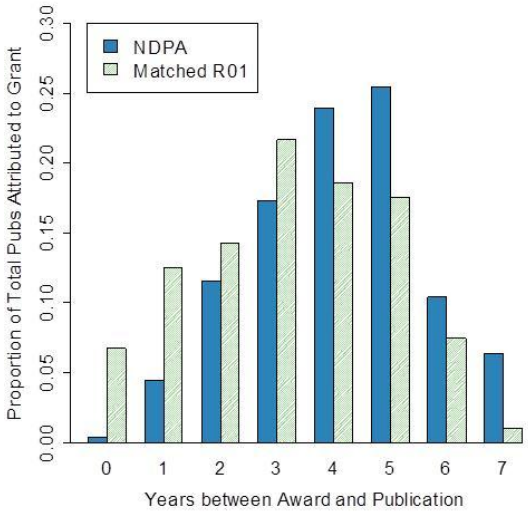


Figure 6.22. Time to publication from time of grant
Pearson's Chi-Squared $p < .0001$

An important measure of the recognition of peers is the speed with which citations grow. 'Citation latency' is the time between a paper being published and its being cited.

Plotting the normalized citation latency of NPDA and R01 publications, one sees that it does not take the community longer to accept research results emerging from the NDPA program (as compared to the mainstream R01 program; see Figure 6.23). Indeed the plot shows that NDPA citations were faster to grow. Indeed, contrary to what one might

expect from controversial research, it does not take the community long to recognize the contributions of NDPA-funded research.

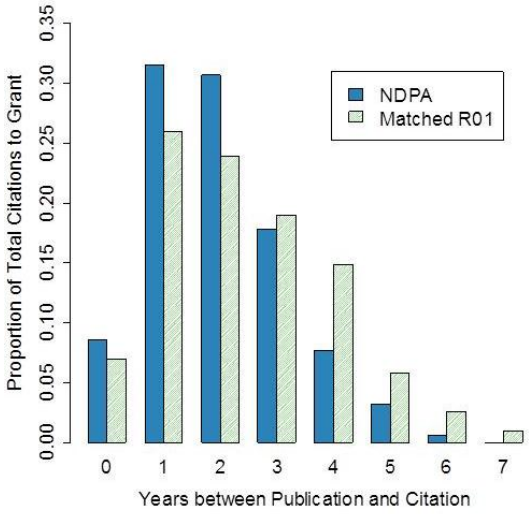


Figure 6.23. Citation latency of the test and comparison group publications
Pearson's Chi-Squared $p < .0001$

6.6.2.4 Findings: Disagreement Among Reviewers Post-Award

As Figure 6.10 displayed, only for four of the 21 Pioneers was there complete agreement among external evaluators with respect to the pioneeringness of NDPA-funded research (or lack thereof). For all of the other awardees, the experts could not agree.

Figure 6.24 illustrates this disagreement differently – in terms of a notional "distance" among the evaluators for each Pioneer. There are four possible scores: *strongly agree* that research is pioneering (a score of +2), *agree* that research is pioneering (+1), *strongly disagree* that research is pioneering (a score of -2), *disagree* that research is pioneering (11). Assuming a distance of 1 between each score leads to the chart below. A distance of 0 implies no disagreement, and a distance of 3 implies the most. A distance of 3 for Pioneer 12, for example, comes from him getting scores of -2, 2 and 2.

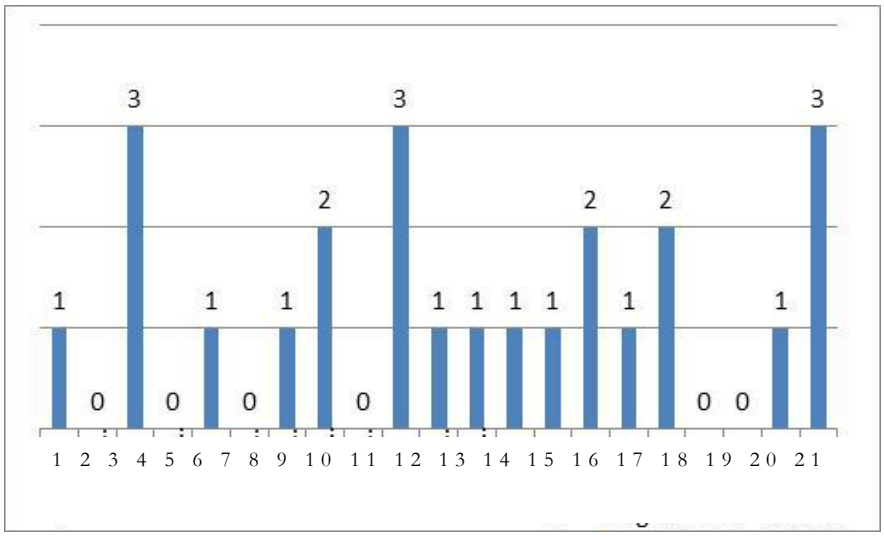


Figure 6.24. Notional agreement "distance" among reviewers

Qualitative review of the coded open-ended responses of the reviewers underscored the ratings. For one of the Pioneers, for example, only one of three experts strongly agreed that her research was pioneering. The other two experts moderately and strongly disagreed with that statement. This negative assessment is likely due to the failure of the PI's original idea. Below is a selection of comments from experts about why Cline's research was or was not pioneering:

> "The resulting methodology for examining functionally connected neuronal cells is either not as radically different from previous methods…or is still at an early stage...The novel method for combining time-lapse imaging and electron microscopy is potentially very important and in some ways more of a novel tool. However, it does not address the question of the organization of interconnected cells, one of the main questions in the original proposal. In several respects, therefore, the technical improvements achieved represent an important but somewhat incremental progress in methodology."
>
> "The research accomplished under the NDPA was solid, and in the case of the insulin receptor signaling work on dendritic plasticity, mechanistically novel. There was little by way of transformative biology or technology that resulted."

Similarly, for another Pioneer, two experts strongly agreed and one strongly disagreed that his research was pioneering. Below is a selection of comments from experts about why the research was or was not pioneering:

> "[Pioneer] has that rare combination of empirical rigor and technical savvy, theoretical understanding and breadth, and creativity and thinking "out of the box", that makes for a truly great researcher. He has consistently pushed the envelope of neurobiology."
>
> "I don't know what was accomplished. As far as I know, he has not made any major discovery or pointed out any major differences."

For another Pioneer, one expert strongly agreed, one moderately agreed, and one strongly disagreed that the accomplished research was pioneering. Below are examples of explanations given by the experts on why they thought the research accomplished was or was not pioneering:

> "The stuff she cited in the Cell paper was fairly novel, but not really an outgrowth of this program. Maybe she was funded by something else and this was the most novel thing she was working on during the time period so she decided to go with it. What she proposed to do was way out there, maybe kinda crazy. She set up a bold hypothesis but she didn't stick to it."
>
> "The results are unexpected and highly novel."
>
> "The proposal that self propagating (prion-like) aggregates play a broader role in neurodegenerative disease has gained some traction (although the jury is still out on this one). However, the PIs research was NOT responsible for these advances. Instead the PI focused on interesting but less conceptual novel areas and made solid progress."

Disagreement among reviewers was also evident in their assessment of the uniqueness of the Pioneer mechanism. Figure 6.25 shows how external evaluators rated each Pioneer on being probed on the uniqueness or value of the NPDA mechanism. As the counts on the RHS scale show, the evaluators were pretty evenly split on the value of the NDPA program.

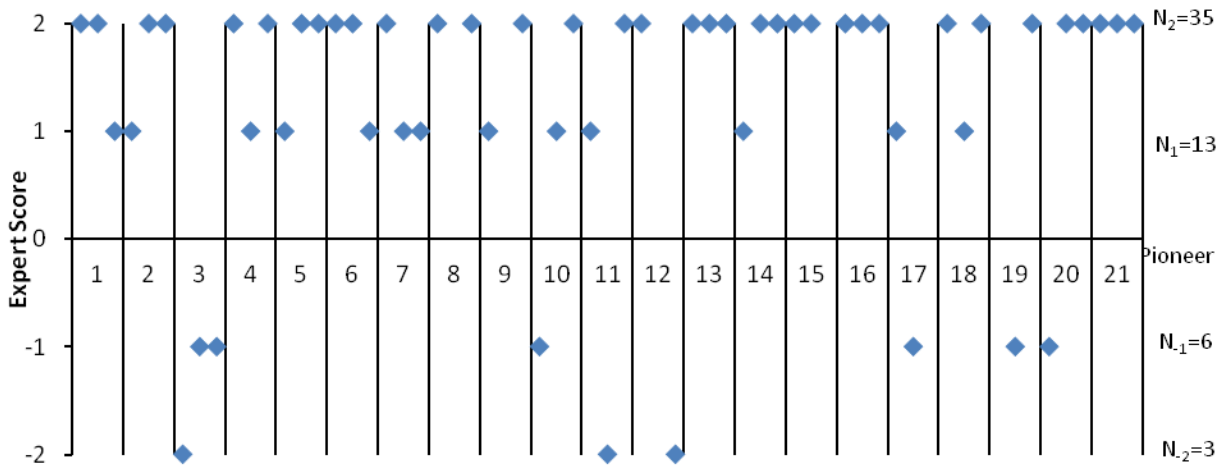


Figure 6.25. Evaluator assessment of the uniqueness/value of the Pioneer mechanism

Note: Each data point represents the assessment of a single expert. Experts were asked to rate this statement on a scale where 2 is strongly agree, 1 is moderately agree, -1 is moderately disagree, and -2 is strongly disagree.

Source: Lal et al. (2011)

6.7 Summary

In this chapter, I used both qualitative and quantitative methods to explore six propositions in the context of the NIH Director's Pioneer Award Program (NDPA) program – a transformative research program at the National Institutes of Health. To make the exploration more meaningful, each proposition was tested with a relevant comparison group.

The first proposition explored the track record of transformative researchers, and compared it to that of researchers that were as similar as they could be to the test group. If transformative researchers had a distinct pattern to their prior success, this could be made a criterion for research selection. However, data analysis found that researchers in both groups had a similar track record from the point-of-view of productivity, quality of research and impact on the scientific community.

The second proposition pointed to the predisposition of young researchers in conducting research that is at odds with prevailing wisdom. Several federal programs

subscribe to this maxim, and target young researchers. If this proposition were supported by data, programs could more formally slant transformative research programs toward younger researchers. Analysis of data from the NDPA and mainstream research programs showed that there seemed to be no particular link between age of researchers (as measured by years since Ph.D.) and whether they were funded to conduct transformative research. Both researchers in the test and comparison groups were of comparable ages.

Indeed, what is likely confused with youth is simply the bringing of insight from outside of a research field's "orthodoxy." This could be accomplished by switching fields. For example, a Pioneer switching to immunology from chemical engineering even at a later stage in life may be having the same effect on innovation in his new field as a new postdoctorate fellow entering the field.

The third proposition explored if prior productivity of researchers was a predictor of transformativeness. Were this proposition supported by data, research programs could make high productivity an explicit criterion for transformative research programs. The productivity of researchers in the test program, NDPA, was compared with that of similarly excellent researchers in the mainstream R01 program. The comparison found no evidence that transformative researchers are more productive than mainstream researchers.

The fourth proposition explored the riskiness of potentially transformative research. Based both on the literature and the number of programs that include "high risk" as a selection criterion, it was important to explore how central risk is to the conduct of transformative research If this proposition were supported by data, risk could be more formally measured and emphasized in the review of a research application. An analysis of scoring-related data showed that the proposals of the awardees of the test transformative research program had higher perceived risk as compared with the proposals of the applicants

to the program. Analysis based on text mining techniques however showed that Pioneers were not any more likely to enter research areas that were already not too crowded with other researchers. Analysis of expert perception also showed that there was a weaker mapping between research that was considered risky at time of award, and "pioneering" at time of evaluation (5 years after award date). It is likely that while there may not be a link between pioneeringness and risk, there may be one between non-pioneeringness and lack of risk. This analysis also suggested that while there is a strong *perceived* link between risk and transformativeness, it may not be as strong in practice, and that there are likely other factors at play that determine a research project's transformativeness.

The next proposition explored the concept of interdisciplinarity. Given the presumption that cutting-edge research occurs at the interface between disciplines, and many programs that require incoming research proposals to be interdisciplinary, this appeared to be an important predictor of transformative research. Indeed if interdisciplinarity were found to be central to the conduct of transformative research, integration scores could be computed at time of application and become an important metric for project selection. Surprisingly though, data analysis found that transformative research is no more interdisciplinary than excellent mainstream research.

Disagreement, both at the time of award as well as time of publication, seemed another important feature of transformative research. Analysis of citations and their growth for NDPA-funded research showed minimal disagreement and rapid assimilation of transformative ideas, comparable to those of researchers in the comparison group. However, a noncomparison-based analysis of disagreement amongst reviewers was notable. Analysis of agreement regarding suitability scores of transformative proposals, for example, revealed great inter-rater disagreement amongst reviewers. Similarly, expert evaluators disagreed

regarding pioneeringness of research and the value of the NDPA mechanism at the end of the project period.

To summarize the findings above, it appears that only one of the propositions – disagreement by peers – was somewhat supportable by data. All others seemed not to hold up to closer data-driven scrutiny. While the findings have their limitations, as Section 5.7 outlined, and there is need for future research, the findings may have policy relevance. These are discussed in the final chapter of this dissertation.

7.0 Overall Summary and Conclusion

7.1 Summary

In this dissertation, I attempted to answer three primary questions: First, what is transformative research, and how is it defined and operationalized in the literature and in the S&T policy community? Second, to what extent are certain attributes – expressed as propositions in this document – about transformative research supported by data? Third, to what extent are attributes – again posed as propositions – about those who conduct transformative research supported by data? In the exploration, my expectation was that examining these propositions analytically could help decide which have an empirical basis, and if (once verified in other domains and parts of the innovation continuum) they could subsequently be used by federal agencies to improve the odds of prospectively identifying transformative research, and developing metrics to evaluate transformative research programs

To address the first question, I conducted a brief review of the relevant literatures in five areas: history and philosophy of science, the sociology of science, psychology, organizational theory, and finance. I also examined twenty programs in the Federal government that purport to support transformative research. The review revealed that the term is used interchangeably with many others, and while there is loftiness to these definitions, the terms have not been operationalized well. In other words, no indicators of transformative research exist. Building on the literature, I posit that the concept of transformative research is a social construct: an agreed upon, or implicit, idea or set of measures coming from within a community, intended to inspire and motivate both performers specifically and society in general, rather than to characterize a specific type of research. Ultimately, I suggest that given the policy goal at hand - nurturing high-impact research - it almost does not matter

what the definition of transformative research is - what matters more is how it can be identified prospectively, managed during its conduct, and evaluated retrospectively.

Six propositions served to address study questions two and three. Three of them targeted transformative research itself (specifically its riskiness, interdisciplinarity, and peer disagreement), and addressed the second study question. Three of them targeted its performers (through the lens of the researchers' track-record, youth, and prior productivity), and addressed the third study question.

I used a range of qualitative and quantitative methods to explore these propositions in the context of the NIH Director's Pioneer Award Program (NDPA) program – a transformative research program at the National Institutes of Health. To make the exploration more meaningful, each proposition was examined with a relevant comparison group. The primary comparison group was a set of researchers comparable in their areas of research from NIH's mainstream research program. This matched set of researchers was not only conducting research in similar areas of research as the Pioneers, but also "looked" demographically (i.e., with respect to age, institutional affiliation) similar to them. For some of the analyses for which this matched set was not appropriate, unsuccessful applicants to the NDPA program, the full set of R01 grantees, and HHMI Investigators were used as comparison groups.

To sum, in an exploratory analysis of the six propositions in the specific context of basic biomedical research showed that only one – disagreement by peers regarding whether research is high-risk, pioneering, or valuable – held up to closer scrutiny. The remaining five propositions were not supported by data. Table 7.1 summarizes the findings. It is important to emphasize that the propositions were explored within a single program within a single

agency, and any interpretation of the findings must take that into account. In fact, Section 7.2 outlines how the propositions can be tested more broadly.

Table 7.1

Summary of Findings from the Analysis

Propositions about the Researcher	Researchers conducting transformative research tend to have a track record that is no different than merely excellent researchers
	Researchers conducting transformative research tend to be no younger than merely excellent researchers
	Researchers conducting transformative research tend to have similar publication patterns as compared with merely excellent researchers
Propositions about the Research	Transformative research tends to be perceived as risky by peers, but may not be strongly associated with transformative outcomes
	Transformative research tends to be no more interdisciplinary than merely excellent research
	Transformative research tends to garner disagreement in the peer review process, but may not take longer to be accepted by the community

7.2 Recommendations for Future Research

The research led to several recommendations for future studies. These could be organized in four general areas, and a final conclusion.

First the propositions need to be tested more broadly. They were explored in one area of research (biomedical research), but all research fields have their own norms and characteristics. What applies to one area of research needn't apply to others. As a result, in order to be sure the propositions are valid for other areas of research, they need to be tested in other fields of research, especially the physical sciences and engineering. Furthermore, even within the field, the propositions were explored with data from *basic* biomedical research (as distinct from applied or translational research, clinical research, drug

development, or other areas). Again to ensure that the findings have broad relevance, the propositions should be explored in other parts of the research continuum, in particular applied research and technology development. Challenge- or end-game oriented research funded by DARPA, for example, is quite different from research funded by NIH, and attributes such as interdisciplinarity may be quite differently expressed than in basic open-ended exploratory research programs like NDPA.

I also recommend that some of the attributes be operationalized differently, two of them in particular. The I-, S- and D-score concepts are sound, and statistically tested in the S&T policy community. However, given their emphasis on the use of subject categories to determine interdisciplinarity, the computation is mechanical, and may under- or overestimate the cognitive integration that might occur if insights come from the same subject area. For future studies, I recommend a more qualitative treatment of the interdisciplinarity, perhaps through interviews, with specific probes to unpack the idea of cognitive integration.

In a similar vein, other concepts might be treated in more quantitative ways. In the literature review in this research, I explored the idea of quantifying risk, but was not able to implement it given limited availability of data. In future studies, I propose operationalizing the concept of risk in similar ways to the finance world – paying particular attention to how the banking industry might do so. Researchers would need to work with federal programs to operationalize risk in more quantitative terms.

One of the most time-consuming aspects of the research was data cleaning – ensuring the correct publications are included and incorrect ones excluded. It is unclear if, beyond a certain point, clean data is helpful with respect to the big picture findings, and there are diminishing returns to the effort. To the extent possible, this aspect of the research ought to be automated or minimized. For a future study, I recommend that the data cleaning

aspect of analysis be reduced, and more data points added. This would also ensure better use of statistics.

Last but not least, I propose that the elapsed time between assessment and the funding of research be around ten years. This will ensure that enough time has passed that there is balance between transformative research having had the opportunity to become evident, and not so much time passed that the lessons learned would no long be applicable to future activities.

7.3 Recommendations for R&D Managers

As the previous section discusses, the propositions above cannot be translated into policy action without being explored in other research areas and parts of the research continuum. Assuming this is done, and they hold true, certain lessons emerge for policymakers and managers of R&D. Indeed, "merely excellent research" may be a better moniker than any other, in a discussion of what type of research to fund. Such research has been conducted since the earliest days of science – when no set-aside transformative research programs existed. When Galileo published *Dialogo sopra i due massimi sistemi del mondo* in 1632, his treatise declaring the triumph of the Copernican system over the traditional Ptolemaic system, there were no transformative research funding programs egging him on. The discovery of penicillin by Alexander Fleming was a serendipitous one, as was the discovery of X-Rays by Wilhelm Röntgen. More recently, Graphene, the one-atom-thick crystal with unusual quantum conductive properties, was discovered by Konstantin Novoselov and Andre Geim after working hours in a very serendipitous manner. Indeed much of breakthrough research in the 19th and 20th centuries has come from mainstream funding programs, and it is important not to discard the role of serendipity in creating

breakthroughs. Does the scientific enterprise really need to *specifically* fund transformative research?

One cannot help but wonder: Is the concept of creating set-aside programs to fund transformative research a ruse of sorts, and just a new justification to continue the almost monotonic fifty-year growth of the S&T enterprise? Are there indeed any differences between programs that fund potentially transformative research and merely excellent "normal" science?

My findings, at least in the realm of basic biomedical research, begin to chip at the "mystique" of transformative research. By design or luck, transformative research has an aura around it; the vision of society standing behind the young rebel garage tinkerer on whom society took a chance is heart-warming indeed, and society has a habit of generalizing from anecdotes. Policymakers are not immune to the romance of funding the next game-changing discovery or technology – no matter how preposterous - that would cure cancer, reach the stars, or enhance human cognition. And it is highly likely that transformative research programs may therefore be here to stay. Given this, it is probably worth building on the insights from this dissertation to derive some policy lessons to make the best of such programs. Some of these insights are presented below using the design-review-perform-manage framework introduced in Chapter 4.

7.3.1 Design of Transformative Research Programs

With respect to program design, the dissertation provides two lessons – program design must follow program goals, and riskiness of proposals need not be an explicit program goal. If a program is about solving a mission-oriented challenge, its design has to be different than if it is about creating open-ended synergies to produce new paradigms or jump-starting breakthrough new research. In designing programs, therefore, program

officials should clearly articulate what they are trying to accomplish. Is the goal to meet a specific technological goal or national need (for example, develop an autonomous airplane that can evade detection). If so, design as an "end-game" program. Is the goal to fund exceptionally creative individual scientists who can engage in what's called "stochastic tinkering"? (stochastic tinkering is a commonly used term that refers to experimenting in small ways, noticing the new or unexpected, and using that to continue to experiment). If so, design to support exceptionally creative researchers. Are there opportunities for teams of scientists to create a new subfield or bring new insights to an existing problem (such as with DOE's Sunshot Initiative to dramatically reduce the cost of solar power). If so, design a team-driven "synergy" program. Or is there need to stimulate new research or new researchers to jump-start a project so that more traditional funding can subsequently be used (such as with IARPA's Automated Location Identification program). If so, design as a "seed" program.

It may also be more productive to design programs that fund excellent and deserving research, whether mission-driven or open-ended. As this dissertation showed, transformative outcomes can emerge both from high and less-high-risk ideas, and while risk can surely be a factor to be considered, it should be considered as one of many criteria in the design of a program.

7.3.2 Review of Proposals for Transformative Research

Chapter 3 describes the review process of many transformative research programs, and identifying transformative research is an important policy activity. For set-aside programs funding such research, this study has a few lessons.

Transformative research program managers should not overly focus on seeking high-risk research, or younger researchers, or highly-productive ones. And while interdisciplinarity

of the proposals should be considered, it should likely not be an overriding factor in the review process. This is not to say interdisciplinarity is not a major driving force of cutting-edge research – it is likely not unique to transformative research any more than it is to just high-quality research.

Disagreement on suitability of funding should, however, be taken seriously. Proposals over which reviewers are in disagreement should be re-examined. I propose there be an "escape hatch" to re-review proposals with disagreements. An escape hatch refers to a mechanism which could give these proposals a second chance with reviewers more well-versed in examining controversial research.

Program managers who run transformative research programs should be carefully selected for meeting high standards of excellence, and then be allowed (as the quote from Charles Herzfeld below emphasizes) to override the recommendations of external peer reviewers if they believe research to be deserving of support.

7.3.3 Management of Transformative Research

Implementation. Just as with program design, the implementation of the program should follow its goals. An end-game program needs a strong and motivated program manager, and a lot of flexibility. In a 2008 *Nature* article, Herzfeld (2008), a former Director of DARPA, said of wanting to build a real agent for change:

> Give it very hard, broad problems to work on. Hire the brightest leaders, problem-solvers and risk-takers, and give them significant resources, guidance, encouragement and freedom. Establish an open, demanding and easy style, with adequate turnover. Hire staff who will build, find more good people, and move on. Then let them work, let them succeed or fail, and give lots of praise. Finally, defend them against friendly fire, for there will be lots.[31]

31

This is easier said than done. NASA's Game Changing Development program, for example, claims to be modeled on DARPA (and will likely be evaluated with the DARPA model in mind), but with none of the flexibility that DARPA programs and program managers have. To some extent, such a mismatch is a recipe for disaster. A program's design and implementation has to follow from its goals.

Individual researcher oriented programs may benefit from the opposite approach, in leaving the researcher completely free from management oversight, with plenty of flexibility in how funds are spent or papers published. "Synergy" programs might benefit from a hybrid approach, or they might consider holding workshops for the awardees to discuss options for overcoming the barriers inherent in interdisciplinary or team research. For "seed" programs, program managers should work to help researchers develop their ideas into full proposals for more traditional mechanisms, and to help identify future funding.

Evaluation. Transformative research programs are still a new construct, and suffer from the challenge – as chapter 3 illustrated – that most of them have not defined terms such as "high-risk, high-reward," pioneering, etc. As Table 3.2 showed, the program may use similar words but may mean different things, and as the typology in Table 3.6 illustrated, despite having defined their goals similarly, the programs have been designed quite differently. Evaluation of transformative research programs have three challenges that go beyond challenges in evaluation traditional programs: operationalization of metrics, timing, and expectations. Each is discussed in turn below.

With respect to operationalization of metrics, the first challenge, as discussed in detail in the limitations section, is a threat to measurement validity. Program managers (and evaluators) model the program by operationalizing the program as designed. Evaluators

operationalize the measures of outcomes, and then examine how the model works by collecting and analyzing data. Based on the model, they make inferences about the success of the program. However, all they then show in an evaluation is if the *model* worked (or did not). The evaluator cannot say if the actual construct – the program being modeled – works. For transformative research programs, which do not have a strong "theory of change" or a strong understanding of the program "construct," this challenge is more daunting than in a traditional evaluation.

Some of the threats to validity can be overcome, or at least ameliorated, by designing evaluations tailored to the type of high-risk, high-reward research program (the typology introduced in this report, shown in Table 3.6 should help with that.) Table 7.2 suggests some strategies.

Table 7.2

Overcoming Threats to Measurement Validity for Transformative Research Evaluations

Threat to Measurement Validity	***Meaning***	***Steps to Address Threat***
Inadequate Preoperational Explication of Constructs	The construct is not well defined operationally	Use methods (e.g., concept mapping) to articulate concepts such as program goals Get experts to critique operationalizations
Mono-Operation Bias	Not capturing the full breadth of the concept of the program	Employ case study approach to understand breadth of program
Mono-Method Bias	Not capturing the full breadth of the measures of the program	Implement multiple methods to test program effectiveness
Interaction of Different Treatments	The "real" program may actually be the *combination* of the separate programs	Employ case study approach to disambiguate multiple streams of funding for PIs
Interaction of Testing and Treatment	Testing is in effect a part of the treatment, and therefore inseparable from the effect of the treatment.	Blind expert panels
Confounding Constructs and Levels of Constructs	Slight increases or decreases of the "dosage" may radically change the results – lack of result may not indicate lack of potential effect	Inferences about constructs that best represent study operations fail to describe the limited levels of the construct studied.

The second lesson under the rubric of metrics relates, again, to the unique nature of transformative research programs. For programs focused on individual researchers, for example, where "stochastic tinkering" may be the norm, metrics should not be based on the project goals that researchers set out in their applications. Evaluation measures for such programs should be those of paradigm shifts (nucleation of new fields like optogenetics, which came out of the work of Pioneers like Karl Deisseroth of Stanford University) rather than traditional outcomes of research. Similarly, a large number of citations, for example, is not an ideal measure of transformation – it could simply be a measure that other researchers are likely building on the work which could well be incremental in nature. It could be a measure of a paradigm shift as well, but not just by itself.

The next policy lesson for evaluation of transformative research is that it likely should not begin until well after the funding has been disbursed. Some evaluation of "seed" programs can be done shortly after the projects end. For example, were the projects able to secure follow-on funding? In most cases, however, research intended to conquer far out frontiers may show little progress if measured too soon. The report of the American Academy of Arts and Sciences (ARISE, 2008) recommended that no outcome evaluations of transformative research programs be performed until at least 10 years after a program begins, with the thought being that any truly transformational work will take that long to be evident as being transformative. Sometimes for policy reasons, however, evaluations must begin sooner than the program is ready. In that case, it might be useful for the evaluation to be a multiphase process.

Last but not least, it is important to emphasize that paradigm-shift research happens but once in a generation, so expecting every project in a transformative research program to be transformative is a fatal flaw. It would indeed make the entire NDPA investment

worthwhile if *one* project found a cure for cancer. DARPA has frequently been derided for wasteful spending, but just one of its thrusts – the creation of the Internet – is priceless and likely worth all the investment since its inception. The lesson here is to take a portfolio approach, and focus on a small number of big successes. Indeed for transformative research, its markers are in the tails and outliers rather than in the means and medians. Figure 7.1 summarizes the lessons learned in the now familiar framework introduced in Chapter 3.

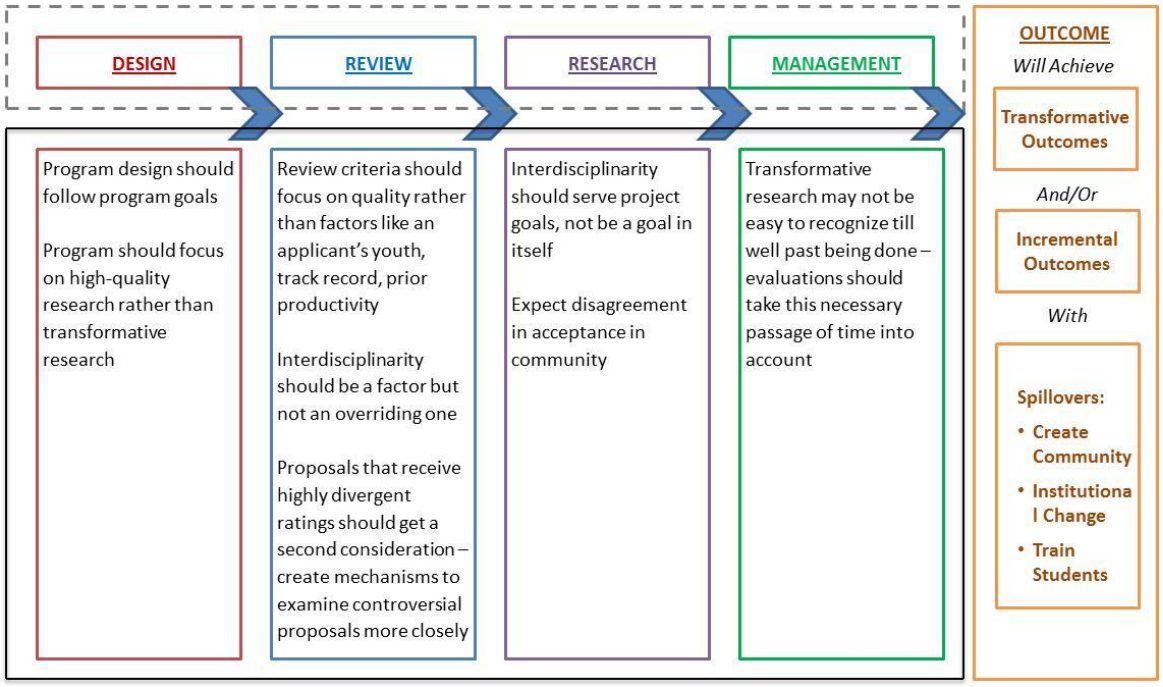


Figure 7.1. Policy recommendations emerging from research and analysis

7.4 Conclusion

Based in part on findings from this dissertation, and pending confirmation by repeating the study in other areas of research, one could reasonably believe that transformative outcomes may be black swan events, unpredictable prospectively (although obvious in retrospect), and could result serendipitously from nurturing merely excellent "normal" science. Seen another way, given that there are no reliable heuristics on what makes research or its performers transformative, it is likely that the odds of identifying

transformative research prospectively cannot be improved. Indeed, at the proposal stage, potentially transformative research may be indistinguishable from "merely excellent research," and in the realm of basic research, R&D managers may be better off funding excellent people to conduct excellent research, regardless of its risk, predicted impact, or transformative potential.

This is my primary takeaway from the analysis (albeit with its many limitations) conducted, and it is not an unreasonable one. But there is also evidence, especially in more technology-oriented domains, that set-aside transformative research programs and agencies can create and have created transformative results. The development of the first computer at the Department of Defense, NASA's support of mankind's first voyage to the Moon, and NIH's funding of the decoding of the human genome, among others, were outcomes of directed efforts to produce transformative outcomes. Ongoing efforts in high-risk high-reward research areas like whole brain emulation continue the tradition. The small probability of their success is balanced by the enormous pay-off if the efforts succeed.

Thomas Jefferson, in a letter to James Madison once said: *I hold it that a little rebellion now and then is a good thing, and as necessary in the political world as storms in the physical* (Jefferson, 1904). Perhaps Jefferson's comment is as applicable to science as in politics, and nations must, on principle, fund research that does indeed shake up the establishment.

It may therefore stand to reason and not just political expedience to say that we need a hybrid approach, balance between traditional and transformative research programs. And then the real question becomes *how much* transformative research ought to be funded? If transformative research is "paradigm-shifting" in a Kuhnian sense, and as such comes about but once in a generation, would throwing more money on it create more of it? Is there some critical mass of transformative research funding that we must reach before leveling off? How

much is that critical mass? Is it seven percent of an agency's total research budget, as Congress demanded of NSF, or eight percent as the National Academies did? Or is it even more? These are unanswered questions, and ought to continue to be debated in the corridors of universities and government alike.

8.0 References

Akerlof, G. A. (1994). Commentary. In R. Swedberg (Ed.), *Economics and sociology. Redefining their boundaries: Conversations with economists and sociologists* (pp. 61–77). Princeton, NJ: Princeton University Press.

Amabile, T. M. (1982). Social psychology of creativity: A consensual assessment technique. *Journal of Personality and Social Psychology, 43*, 997-1013. doi:10.1037//0022-3514.43.5.997

Amabile, T. M. (1996). Creativity in context: Update to the social psychology of creativity. Boulder, CO: Westview Press.

America COMPETES Act of 2007, Pub. L. No. 110-69, §1008a, 121 Stat. 572. Retrieved from http://www.gpo.gov/fdsys/pkg/PLAW-110publ69/pdf/PLAW-110publ69.pdf

America COMPETES Reauthorization Act of 2010, Pub. L. No. 110-69, §1008a, 121 Stat. 572. Retrieved from http://www.gpo.gov/fdsys/pkg/PLAW-110publ69/pdf/PLAW-110publ69.pdf

American Academy of Arts and Sciences. (2008). Investing in early-career scientists and high-risk, high-reward research. ARISE: Advancing research in science and engineering. Retrieved from http://www.amacad.org/ariseFolder/ariseReport.pdf

Arrow, K. J. (1965). *Aspects of the theory of risk bearing*. Helsinki: Yrjö Jahnsson Säätiö.

Artzner, P., Delbaen, F., Eber, J-M., & Heath, D. (1999). Coherent measures of risk. *Mathematical Finance*, *9*, 203-228. doi:10.1111/1467-9965.00068

Austin, F. C. (2008). High-risk high-reward research demonstration project. NIH Council of Councils. Retrieved from http://dpcpsi.nih.gov/council/pdf/CoC-112008-Austin-HRHR.pdf

Avery, D. (2006). Saving People and Nature with the Green Revolution. *The Steward (Hawaii)* January 21, 2006. Retrieved from http://www.agbioworld.org/biotech-info/topics/borlaug/greatest.html

Azoulay, P., Ding, W., & Stuart, T. (2009). The impact of academic patenting on the rate, quality, and direction of (public) research output. *Journal of Industrial Economics, 57*, 637-676. Retrieved from http://www.nber.org/papers/w11917

Azoulay, P., Graff Zivin, J., & Manso, G. (2009). Incentives and creativity: Evidence from the academic life sciences. Retrieved from http://faculty.haas.berkeley.edu/manso/agm.pdf

Ball, P. (2011) Science funding tends to favour mediocrity over grand ideas, The Guardian, Friday 9 December 2011, http://www.guardian.co.uk/commentisfree/2011/dec/09/science-funding-creativ-philip-ball

Barendse, W. (2007). The strike rate index: A new index for journal quality based on journal size and the h-index of citations. *Biomedical Digital Libraries, 4*(3). doi:10.1186/1742-5581-4-3. Retrieved from http://www.bio-diglib.com/content/4/1/3

Batagelj, V., & Mrvar, A. (2002). Pajek—analysis and visualization of large networks. Retrieved from http://www.imfm.si/preprinti/PDF/00871.pdf

Battelle. (2011). 2012 Global R&D funding forecast. Retrieved from http://www.battelle.org/ABOUTUS/rd/2012.pdf

Becker T. J. (2005, Spring/Summer). Wake-up call for innovation. *Research Horizons*, 6-13. http://gtresearchnews.gatech.edu/reshor/rh-ss05/at-risk.pdf

Bement, A. L., Jr. (2007). Important Notice No. 130: Transformative Research http://www.nsf.gov/pubs/2007/in130/in130.jsp

Bement, A. L., Jr. (2007i). Transformative research: The artistry and alchemy of the 21st century. Retrieved from http://www.nsf.gov/news/speeches/bement/07/alb070104_texas.jsp

Bents, R. (2004). Creatrix: History and analysis: Current status. Retrieved from http://rapidbi.com/wp-content/uploads/resourcezone/Creatrix-validity.pdf

Bornmann, L., & Daniel, H. D. (2008). What do citation counts measure? A review of studies on citing behavior. *Journal of Documentation, 64*(1), 45-80. doi:10.1108/00220410810844150 Retrieved from http://www.lutz-bornmann.de/icons/BornmannLutzCitingBehavior.pdf

Boyack, K. W., Klavans, R., & Börner, K. (2005). Mapping the backbone of science. *Scientometrics, 64*, 351-374. Retrieved from http://scimaps.org/exhibit/docs/05-boyack.pdf

Braben D. (2004). Pioneering Research: A Risk Worth Taking. Hoboken: Wiley.

Brainard, J. (2007). NIH Calls for Big Changes to Peer Review and Grants. *The Chronicle of Higher Education,* http://chronicle.com/article/NIH-Calls-for-Big-Changes-to/24325

Braun T., Glänzel W., & Schubert, A. (2006) A Hirsch-type index for journals. *Scientometrics, 69*(1), 169-173. Retrieved from http://sci2s.ugr.es/hindex/pdf/Braunetal2006.pdf

Brenner, S. (1998, November 20). The impact of society on science. *Science*, *282*, 1411-1412. Doi:10.1126/science.282.5393.1411. Retrieved from http://www.sciencemag.org/content/282/5393/1411.full

Burawoy, M., et al. (1991), Ethnography Unbound, *University of California Press.*

Burrell, Q. L. (2007). Hirsch index or Hirsch rate? Some thoughts arising from Liang's data. *Scientometrics, 73*(1), 19-28. doi:10.1007/s11192-006-1774-5

Burt, R. S. (2004). Structural holes and good ideas. *American Journal of Sociology, 110*, 349-399. Retrieved from http://www.econ.upf.edu/docs/seminars/burt.pdf

Bush, V. (1945). Science: The endless frontier. Washington, DC: Government Printing Office. Retrieved from http://www.nsf.gov/od/lpa/nsf50/vbush1945.htm

Cook, T. & Campbell, D. (1979). Quasi-Experimentation: Design and Analysis for Field Settings. Rand McNally, Chicago, Illinois.

Carnes, M., Geller, S., Fine, E., Sheridan, J., & Handelsman, J. (2005). NIH director's pioneer awards: Could the selection process be biased against women? *Journal of Women's Health, 14*(8), 684-691. Retrieved from http://handelsmanlab.sites.yale.edu/sites/default/files/NIH%20Director%E2%80%99s%20Pioneer%20Awards.pdf

Case, R. H., & Shane, S. (2007). Fostering risk taking in research and development: The importance of a project's terminal value. *Decision Sciences, 29*(4), 765-783. doi:10.1111/j.1540-5915.1998.tb00876.x

Chima J. (2005), What's the Utility of the Case-Study Method for Social Science Research?: A Response to Critiques from the Quantitative/Statistical Perspective, Paper for delivery at the 2005 Annual Meeting of the American Political Science Association, September 1-4, 2005.

Chen, C. (2005). The centrality of pivotal points in the evolution of scientific networks. In R. St. Amant, J. Riedl, & A. Jameson (Eds.), *Proceedings of the 10th International Conference on Intelligent User Interfaces* (pp. 98-105). doi:10.1145/1040830.1040859. Retrieved from http://dl.acm.org/citation.cfm?id=1040830&picked=prox

Chen, C. (2011). Predictive effects of structural variation on citation counts. *Journal of the American Society for Information Science and Technology*, *63*, 431-449. doi:10.1002/asi.21694

Chen, C. (2011). *Turning points: The nature of creativity.* New York, NY: Springer.

Chen, C., Chen, Y., Horowitz, M., Hou, H., Liu, Z., & Pellegrino, D. (2009). Towards an explanatory and computational theory of scientific discovery. *Journal of Informetrics, 3*(3), 191-209. Retrieved from http://www.sciencedirect.com/science/journal/17511577

Christensen, Clayton M. (1997), The innovator's dilemma: when new technologies cause great firms to fail, Harvard Business School Press, Boston, MA.

Chubin, D., Hackett, E. (1990). Peerless Science: Peer Review and U.S. Science Policy. State University of New York Press, 1990.

Chubin, D., Hackett, E. (forthcoming). Peer Review. Entry for the Encyclopedia of Science, Technology, and Ethics, Carl Mitcham, ed., forthcoming.

Colwell, R. R. (2003). Keynote address to the International Life Sciences Summit at Georgetown University, Washington, DC. Retrieved from http://www.nsf.gov/news/speeches/colwell/rc031020lifesci_summit.htm

Cozzens, S. (1996). *Assessment of fundamental science programs in the context of the Government Performance and Results Act (GPRA).* Washington, DC: RAND. Retrieved from http://www.rand.org/pubs/monograph_reports/MR707z0.html

Crane, D. (1980). An exploratory study of Kuhnian paradigms in theoretical high energy physics. *Social Studies of Science, 10*(1), 23–54. doi:10.1177/030631278001000102

Dasgupta, P., & Maskin, E. (1987). The simple economics of research portfolios. *The Economic Journal, 97*, 581–595. Retrieved from http://www.blackwellpublishing.com/journal.asp?ref=0013-0133

De Moya-Anegón, F., Vargas-Quesada, B., Chinchilla-Rodríguez, Z., Corera-Álvarez, E, Munoz-Fernández, F. J., & Herrero-Solana, V. (2007). Visualizing the marrow of

science. *Journal of the American Society for Information Science and Technology, 58*, 2167-2179. Retrieved from http://eprints.rclis.org/bitstream/10760/11095/1/Visualizing_the_marrow_of_science.pdf

Dehejia R. H., & Wahba, S. (2002). Propensity score-matching methods for nonexperimental causal studies. *The Review of Economics and Statistics, 84*(1), 151–161. Retrieved from http://www.nber.org/~rdehejia/papers/matching.pdf

Denison, R. F. (2007, August 15). Defining 'transformative research' for NSF [Web log post]. Retrieved from http://blog.lib.umn.edu/denis036/thisweekinevolution/2007/08/something_doesnt_add_up.html

Dietz, J. S., & Rogers, J. D. (2012). Meanings and policy implications of "Transformative Research": Frontiers, hot science, evolution, and investment risk. *Minerva, 50*(1): 21-44. doi:10.1007/s11024-012-9190-x

Dowd, K. (2000). Adjusting for risk: An improved Sharpe ratio. *International Review of Economics and Finance, 9*, 209–222. Retrieved from http://elmu.umm.ac.id/file.php/1/jurnal/I/International%20Review%20of%20Economics%20And%20Finance/Vol9.Issue3.July2000/65.pdf

Felli, J. C., & Andersen, J. S. (2011). R&D risk management. In *Wiley encyclopedia of operations research and management science*. doi:10.1002/9780470400531.eorms0703

Fischhoff, B., Watson, S. R., & Hope, C. (1984). Defining risk. *Policy Sciences 17*, 123-139. Retrieved from http://www.hss.cmu.edu/departments/sds/media/pdfs/fischhoff/DefiningRisk.pdf

Floricel, S., & Ibanescu, M. (2008). Using R&D portfolio management to deal with dynamic risk. *R&D Management, 38*, 452-467. Retrieved from http://www.blackwellpublishing.com/journal.asp?ref=0033-6807

Foster, E. M. (2003). Propensity score matching: An illustrative analysis of dose response. *Medical Care, 41*, 1183-1192. doi:10.2307/3767893

Froderman, R., Holbrook J. (2012) The Promise and Perils of Transformative Research, *Report on the workshop 'Transformative Research: Ethical and Societal Implications'* held in Arlington, VA March 8 & 9, 2012

Garfield, E. (2005). The agony and the ecstasy-the history and meaning of the journal impact factor. Presented at the meeting of the International Congress on Peer Review and Biomedical Publication, Chicago, IL. Retrieved from http://garfield.library.upenn.edu/papers/jifchicago2005.pdf

Garfield, E. (1972). Citation analysis as a tool in journal evaluation. *Science, 178*, 471-479. Retrieved from http://www.elshami.com/Terms/I/impact%20factor-Garfield.pdf

Garfield, E. (1979). *Citation indexing. Its theory and application in science, technology and humanities*. New York, NY: Wiley.

Garfield, E. (1983). Citation indexes for science: A new dimension in documentation through association of ideas. *Science, 122*, 108-111. (Reprinted from *Science, 123,* pp. 108-111, 1955). Retrieved from http://garfield.library.upenn.edu/essays/v6p468y1983.pdf

Garfield, E., Malin, M. V., & Small, H. (1978). Citation data as science indicators. In Y. Elkana, J. Lederberg, R. Merton, A. Thackray, & H. Zuckerman (Eds.), *Toward a metric of science: The advent of science indicators* (pp. 179-209). New York, NY: John Wiley

& Sons. Retrieved from http://garfield.library.upenn.edu/metricofscience/metricofsciencecontents.pdf

Gentile, J. (2010) Let's encourage research daring enough to fail. *Chronicle of Higher Education*. Retrieved from http://chronicle.com/article/Lets-Encourage-Research/65781/

Gerlin, A. (2011, October 5) Tecnion's Shechtman Wins Nobel in Chemistry for Quasicrystals Discovery, http://mobile.bloomberg.com/news/2011-10-05/technion-s-shechtman-wins-chemistry-nobel-for-discovery-of-quasicrystals

Giere, R. N. (2006). *Scientific perspectivism*. Chicago, IL: University of Chicago Press.

Glänzel W., & Schubert, A. (2003). A new classification scheme of science fields and subfields designed for scientometric evaluation purposes. *Scientometrics. 56*, 357-367. Retrieved from http://www.mii.lt/support/matematika/straipsniai/str/GlanzelClassification.pdf050529966.pdf

Glänzel W., Debackere, K., Thijs B., & Schubert, A. (2006). A concise review on the role of author self-citations in information science, bibliometrics and science policy. *Scientometrics, 67*, 263-277. doi:10.1007/s11192-006-0098-9

Glänzel W., Leta, J., & Thijs, B. (2006). Science in Brazil Part 1: A macro-level comparative study. *Scientometrics, 67*(1), 67-86. doi:10.1007/s11192-006-0055-7

Glänzel W., Schlemmer, B., Schubert, A., & Thijs, B. (2006). Proceedings literature as additional data source for bibliometric analysis. *Scientometrics, 67,* 457-473.

Glänzel W., Thijs, B., & Schlemmer, B. (2004). A bibliometric approach to the role of author self-citations in scientific communication. *Scientometrics, 59*(1), 63-77. doi:10.1023/B:SCIE.0000013299.38210.74

Glänzel, W. (2006). On the opportunities and limitations of the H-index. *Science Focus. 1*(1), 10-11.

Gläser, J. (2007). The social orders of research evaluation systems. In R. Whitley & J. Gläser (Eds.), *The changing governance of the sciences* (pp. 245-266). Springer. Retrieved from http://www.springerlink.com/content/r8283q021855t165/

Glenisson P., Glänzel W., Janssens, F., & De Moor, B. (2005). Combining full text and bibliometric information in mapping scientific disciplines. *Information Processing & Management, 41,* 1548-1572. Retrieved from ftp://ftp.esat.kuleuven.ac.be/sista/ida/reports/04-205.pdf

Godin, B., & Doré, C. (2004). Measuring the impacts of science: Beyond the economic dimension. Retrieved from http://www.csiic.ca/PDF/Godin_Dore_Impacts.pdf

Grant, J., & Allen, L. (1999). Evaluating high risk research: An assessment of the Wellcome Trust's Sir Henry Wellcome Commemorative Awards for Innovative Research. *Research Evaluation, 8*, 201-204. doi:10.3152/147154499781777513

Guo, S., Barth, R., & Gibbons, C. (2004, January). Introduction to propensity score matching: A new device for program evaluation. Presentation at the meeting of the Society for Social Work Research, New Orleans, LA. Retrieved from http://ssw.unc.edu/VRC/Lectures/PSM_SSWR_2004.pdf

Hackett, E. (1997). The Place of Peer Review in Science and Science Policy, in Mark Frankel and Jane Cave, eds., Evaluating Science and Scientists. Budapest: Central European University Press.

Hage, J., Meeus, M., & Edquist, C. (2006). *Innovation, science and institutional change: A research handbook*. Oxford: Oxford University Press.

Hansson, S. (2011). Risk. The Stanford Encyclopedia of Philosophy (Fall 2011 Edition), Edward N. Zalta (ed.), http://plato.stanford.edu/archives/fall2011/entries/risk/

Harman, O., & Dietrich, M. (2008). *Rebels, mavericks, and heretics in biology*. New Haven, CT: Yale University Press.

Häyrynen, M. (2007). Breakthrough research funding for high-risk research at the Academy of Finland Publications of the Academy of Finland 8/07

Heinze, T., & Bauer, G. (2007). Characterizing creative scientists in nano-S&T: productivity, multidisciplinarity, and network brokerage in a longitudinal perspective. *Scientometrics, 70*(3). 811-830. doi:10.1007/s11192-007-0313-3

Heinze, T., Shapira, P., Senker, J., & Kuhlmann, S. (2007). Identifying creative research accomplishments: methodology and results for nanotechnology and human genetics. *Scientometrics, 70* (1):125-152. doi:10.1007/s11192-007-0108-6

Heinze, T. (2008). How to sponsor ground-breaking research: a comparison of funding schemes. *Science and Public Policy*, *35*(5), 302–318 doi:10.3152/030234208X317151

Hemphill, T. A. (2006). US innovation policy: creating (and expanding) a national agenda for global competitiveness. *Innovation: Management, Policy & Practice*, Innovation: Management, Policy & Practice: Vol. 8, No. 3, pp. 288-295.

Hicks D., Tomizawa, H., Saitoh, Y., & Kobayashi, S. (2004). Bibliometric techniques in the evaluation of federally funded research in the United States. *Research Evaluation. 13*, 1, 76-86. Retrieved from http://works.bepress.com/diana_hicks/10

Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences*,*102*, 16569-16572. doi:10.1073/pnas.0507655102

Hollingsowrth, J. R., & Hage, J. (2000). A strategy for the analysis of idea innovation. *Organization Studies, 21*, 971-1004.

Hollingsworth, J. R. (2000). Doing institutional analysis: implications for the study of innovations. *Review of International Political Economy, 7*(4), 595-644. Retrieved from http://faculty.history.wisc.edu/hollingsworth

Hollingsworth, J. R. (2002). Research organizations and major discoveries in twentieth-century science: a case study of excellence in biomedical research. Discussion Papers, Presidential Department P 02-003, Social Science Research Center Berlin (WZB).

Holton, G. A. (2004). Defining risk. *Financial Analysts Journal, 60*(6), 19-25. Retrieved from http://riskexpertise.com/papers/risk.pdf

Hughes, M. E., & Lal, B. (2010). Understanding High-Risk, High-Reward Research Programs. Poster presented at the Gordon Research Conference on Science and Technology Policy, Waterville Valley, NH.

Iglesias, J. E., & Pecharroman, C. (2007). Scaling the h-index for different scientific ISI fields. *Scientometrics 73,* 303-320. doi:10.1007/s11192-007-1805-x.

Investing in high-risk, high-reward research, 111th Cong. (2009). Retrieved from http://www.gpo.gov/fdsys/

Jefferson, T (1904) The Works of Thomas Jefferson, Federal Edition (New York and London, G.P. Putnam's Sons, 1904-5). Vol. 5. Chapter: TO JAMES MADISON 1 Accessed from http://oll.libertyfund.org/title/802/86654 on 2012-05-28

Jensen, M. C. (1968). The performance of mutual funds in the period 1945-1964*, Journal of Finance, 23*, 389-416. Retrieved from http://www.afajof.org/journal/browse.asp

Johnston, S. C., & Hauser, S. L.(2008). Transformative research. *Annals of Neurology,* 63: A11–A13. doi: 10.1002/ana.21414

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47*, 263-291.

Katz, M. (1986). An analysis of cooperative research and development. *RAND Journal of Economics, 17*, 527–543. Retrieved from http://www.jstor.org/stable/2555479

Klavans, R., & Boyack, K. W. (2009). Toward a consensus map of science. *Journal of the American Society for Information Science and Technology, 60* (3), 455-476.

Klein, J. T. (1990). *Interdisciplinarity: History, theory, and practice*. Detroit, MI: Wayne State University Press.

Kuhn, S. (1962). *The structure of scientific revolutions*. Chicago: University of Chicago Press. p. 206.

Lal, B., Chaturvedi, R., Zhu, A., Hughes, M., Shipp, S., Kang, C., Marshall, M., Yglesias, E., Singpurwalla, D. (2010). Process Evaluation of the National Institutes of Health (NIH) Director's Pioneer Award (NDPA), FY 2004-2005, Available at http://commonfund.nih.gov/pdf/PioneerAwardProcessEvaluation_2004-2008.pdf

Lal, B., Hughes, M., Shipp, S., Lee, E., Richards, A., Zhu, A. (2011). Outcome Evaluation of the National Institutes of Health (NIH) Director's Pioneer Award (NDPA), FY 2004-2005, Available at http://commonfund.nih.gov/pdf/Pioneer_Award_Outcome%20Evaluation_FY2004-2005.pdf

Langfeldt, L. (2001). The decision-making constraints and processes of grant peer review, and their effects on the review outcome. *Social Studies of Science, 31*(6), 820–841.

Langfeldt, L. (2006). Risk avoidance: The policy challenges of peer review: managing bias, conflict of interests and interdisciplinary assessments. *Research Evaluation, 15*(1), 31-41.

Langfeldt, L., & Brofoss, K. E. (2005). *Evaluation of the European young investigator awards scheme*. Working Paper.

Laudel, G. (2006). The art of getting funded: how scientists adapt to their funding conditions. *Science and Public Policy, 33*(7), 489–504.

Lehrer, J. (2010), Fleeting Youth, Fading Creativity, http://online.wsj.com/article/SB10001424052748703444804575071573334216604.html?KEYWORDS=%22Scientific+revolutions+are+often+led+by+the+youngest+scientists%22

LeRoy, S. (1987). Knight on Risk and UncertaintyAuthor(s): Stephen F. LeRoy and Larry D. Singell, Jr. *Journal of Political Economy, 95*(2), 394-406.

Leta, J., Glänzel W., & Thijs, B. (2006). Science in Brazil. Part 2: Sectoral and institutional research profiles. *Scientometrics, 67*(1), 87-105. doi:10.1007/s11192-006-0051-

Leydesdorff, L., & Rafols, I. (2009). A global map of science based on the ISI subject categories. *Journal of the American Society for Information Science and Technology 60*(2), 348-362.

Leydesdorff, L., & Rafols, I. (2010). Indicators of the interdisciplinarity of journals: Diversity, centrality and citations. *Journal of Informetrics*, Volume 5, Issue 1, January 2011, Pages 87-100, ISSN 1751-1577, 10.1016/j.joi.2010.09.002. available at http://www.sciencedirect.com/science/article/pii/S1751157710000854

Leysdesdorff, L. (2007). Betweeness centrality as an indicator of interdisciplinarity of scientific journals. *Journal of the American Society for Information Science and Technology 58(9)*, 1303-1319.

Lundberg, J. (2007). Lifting the crown-citation z-score. *Journal of Informetrics, 1*, 145-154.

Luiggi, C. (2011). Q&A: Aging Geniuses - A new study shows that over the past century, the age at which scientists produce their most valuable work is increasing. *News & Opinion*, November 8, 2011http://the-scientist.com/2011/11/08/qa-aging-geniuses/

Luukkonen, T. (2012). Conservatism and risk-taking in peer review: Emerging ERC practices, *Research Evaluation* (2012) 21 (1): 48-60. doi: 10.1093/reseval/rvs001

March, J. G., & Shapira, Z. (1987). Managerial perspectives on risk and risk taking: *Management Science, 33*(11), 1404-1418.

Markowitz, H. M. (1959). Portfolio selection: efficient diversification of investments. New York: John Wiley & Sons. (Reprinted by Yale University Press, 1970; 2nd ed. Basil Blackwell, 1991)

Masterman, M. (1970). The nature of a paradigm. In I. Lakatos & A. Musgrave (Eds.), Criticism and the growth of knowledge. *Proceedings of the international colloquium in the philosophy of Science, London, 1965. 4* (pp. 59–89). Cambridge: Cambridge University Press.

Melin, G., & Danell, R. (2006). The top eight percent: development of approved and rejected applicants for a prestigious grant in Sweden. *Science and Public Policy, 33*(10), 702–712

Mervis, J. (2004). 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, pp. 220–221.

Modigliani, F., & Modigliani, L. (1997). Risk adjusted performance: How to measure it and why. *The Journal of Portfolio Management, 23*(2), 45-54. doi:10.3905/jpm.23.2.45

Moed, H. F. (2010). Measuring contextual citation impact of scientific journals. *Journal of Informatics, 4*, 265-277.

Moed, H. F., Glänzel W., & Schmoch, U. (2004). *Handbook of quantitative science and technology research*. New York, NY: Kluwer Academic.

Moore, M. (1982). Evaluating creativity. *The English Journal, 71*(8), 33-35.

Moya-Anegon, F., Vargas-Quesada, B., Herrero-Solana, V., Chinchilla-Rodriguez, Z., E. Corera-Alvarez, E., &. Munoz-Fernandez, F. J. (2004). A new technique for building maps of large scientific domains based on the cocitation of classes and categories. *Scientometrics, 61*(1), 129-145.

Murdock, M.C., & Pucclo, G. J. (1993). *A contextual organizer for conducting creativity research." understanding and recognizing creativity*. In S.G. Isaksen, M.C. Murdock, R.L. Firestien, & D.J. Treffinger (Eds.), Understanding and recognizing creativity: The emergence of a discipline. Norwood, NJ: Ablex.

National Academy of Sciences (1999). Evaluating federal research programs. Research and the Government Performance and Results Act. Washington, D.C.: Author.

National Academy of Sciences. (2005). Facilitating interdisciplinary research. Committee on Facilitating Interdisciplinary Research, National Academy of Sciences, National Academy of Engineering, Institute of Medicine. Washington, D.C.

National Academy of Sciences. (2007). Rising Above the Gathering Storm: Energizing and Employing America for a Brighter Economic Future. Committee on Science, Engineering, and Public Policy (COSEPUP) Retrieved from http://www.nap.edu/openbook.php?record_id=11463&page=1

National Academy of Sciences. (2010). By the President At The National Academy Of Sciences Annual Meeting. Retrieved from http://www.nationalacademies.org/morenews/20090428.html, and http://www.whitehouse.gov/the_press_office/Remarks-by-the-President-at-the-National-Academy-of-Sciences-Annual-Meeting.

National Academy of Sciences (2011). Measuring the impacts of federal investments in research: A workshop summary." Washington, D.C.

National Institute of Standards and Technology. (2011). Retrieved from http://www.nist.gov/public_affairs/releases/shechtman-100511.cfm, and http://news.businessweek.com/article.asp?documentKey=1376-LRXLB40YHQ0X01-1E7U2VUHSGIVADSR3FA29L1R15.

National Institutes of Health Office of Portfolio Analysis and Strategic Initiatives. (2006, April 7). NIH Roadmap for Medical Research: High-risk Research.

National Research Council (2005). Assessment of Department of Defense basic research. Washington, D.C.: National Academies Press.

National Science Board. (2007). Enhancing support of transformative research at the National Science Foundation. Retrieved from http://www.nsf.gov/pubs/2007/nsb0732/nsb0732.pdf

National Science Foundation. (2004). Summary of National Science Board Workshop: Identifying, reviewing and funding transformative research. Unpublished Manuscript.

NDPA (2007) RFA-RM-07-005: 2007 NIH Director's Pioneer Award Program (DP1) http://grants.nih.gov/grants/guide/rfa-files/RFA-RM-07-005.html#SectionV

NSF Advisory Council (1980) "Report of Task Group 10: Funding of Innovative High Risk Proposals." NSF Historian's Files. As quoted in Frodeman & Holbrook (2012). The Promise and Perils of Transformative Research. Report on the workshop 'Transformative Research: Ethical and Societal Implications' held in Arlington, VA March 8 & 9, 2012, Available at http://csid.unt.edu/files/TRW_FinalReport.pdf

NSF. (2005). Strategic Planning Overview: Strategic Directions for Engineering Research, Innovation and Education. National Science Foundation, Directorate for Engineering.

NSF. (2006). High-risk research in anthropology. Program Announcement. 7 Sep. 2006. National Science Foundation. 10 Sep. 2008

NSF. (2007i). Impact of Proposal and Award Management Mechanisms Final Report August 1, 2007. National Science Foundation. http://www.nsf.gov/pubs/2007/nsf0745/nsf0745.pdf

NSF (2007). Notice No. 130 September 24, 2007 Important Notice To Presidents Of Universities And Colleges And Heads Of Other National Science Foundation Awardee Organizations on Transformative Research, Available at http://www.nsf.gov/pubs/2007/in130/in130.pdf, Accessed May 25, 2011

NSF (2009) NSF-wide Facilitating Transformative and Interdisciplinary Research (FacTIR) working group report (Internal working document)

National Institutes of Health [NIH] (2008). Enhancing peer review: A self-study by the NIH in partnership with the scientific community to strengthen peer review in changing times.

NIH Roadmap, Available at http://nihroadmap.nih.gov

NIH. (2004). "Proceedings of the 88th Meeting of the Advisory Committee to the Director," National Institutes of Health, May 6, 2004.

Nunez, M. (2011). The Weighted Brennan-Prediger Coefficient for Multiple Raters in Low Variance Data with Missing Values, Section on Government Statistics – Joint Statistical Meeting.

Okamura K. and Vonortas N. (2009). Partnership networks and knowledge networks in five sectors, in Malerba F. and Vonortas, N.S. (eds.), Innovation Networks in Industries, Edgar Elgar Publishing, Chap. 6, pp.125-156.

Olsen K L and Carlson A B. (2007) "Federal Science Policy and University Research Agendas." Defining Values for Research and Technology. Rowman & Littlefield. 2007. (117-128).

Oppenheim, C. (1995) "The correlation between citation counts and the 1992 research assessment exercise ratings for British library and information science university departments." Journal of Documentation, 51(1995), 18–27.

Oppenheim, C. (1997) "The correlation between citation counts and the 1992 research assessment exercise ratings for British research in genetics, anatomy and archaeology." Journal of Documentation, 53(1997), 477–487.

OSTP (2010), Science and Technology Priorities for the FY 2012 Budget. 2010, Office of Science and Technology Policy: Washington, D.C. http://www.whitehouse.gov/sites/default/files/microsites/ostp/fy12-budget-guidance-memo.pdf

Persson O., Glänzel W., & Danell, R. (2004). Inflationary bibliometric values: The role of scientific collaboration and the need for relative indicators in evaluative studies. *Scientometrics, 60*, 421-432.

Petsko, G. A. (2008). The right to be wrong, *Genome Biology, 9*. doi:10.1186/gb-2008-9-2-102

Pion, Georgine M., and Charles B. Hammond. (2005). The American Association of Obstetricians and Gynecologists Foundation Scholars Program: Additional data on research-related outcomes. American Journal of Obstetrics and Gynecology 193 (5):1733-1739.

Pion, Georgine M., and David S. Cordray. (2008). The Burroughs Wellcome Career Award in the Biomedical Sciences. Evaluation & the Health Professions 31 (4):335-369.

Polanyi, M., (1969). Knowing and Being. Chicago University Press, Chicago.

Pollack A. (1994) "Technology; The creative force behind Japan's Computer U." The New York Times. 26 Jun 1994.

Podolny J. (1996) "Networks, Knowledge, and Niches: Competition in the Worldwide Semiconductor Industry, 1984-1991" American Journal of Sociology, Vol. 102, No. 3 (Nov., 1996), pp. 659-689.

Ponomarev, I., D. Williams, J. Schnell, and L.Haak. (2011) "Quickly identifying breakthrough papers in scientific research." Presentation at the 1st Global TechMining Conference, Atlanta, GA, September 14, 2011.

Porter, A. L., A. S. Cohen, D. J. Roessner, and M. Perreault. 2007. Measuring researcher interdisciplinarity. Scientometrics, Vol. 72, No. 1 (2007) 117–147

Porter, A., Roessner, D., Heberger, A. 2008. How interdisciplinary is a given body of research? Research Evaluation 17 (4):273-282.

Pratt, J. (1964). Risk aversion in the small and in the large. Econometrica, 32

Prendergast, P.J., Brown, S.H., Britton, J.R. (2008), Research programmes that promote novel, ambitious, unconventional and high-risk research: an analysis INDUSTRY & HIGHER EDUCATION Vol 22, No 4, pp 215–221

Providing for consideration of the bill (H.R. 1868) to authorize appropriations for the National Institute of Standards and Technology for fiscal years 2008, 2009, and 2010, and for other purposes, H.R. 350, 110th Cong. (2007). Retrieved from http://www.gpo.gov/fdsys/pkg/BILLS-110hres350rh/pdf/BILLS-110hres350rh.pdf

Public Law 111 - 358 - America Competes Reauthorization Act Of 2010" Or The ``America Creating Opportunities To Meaningfully Promote Excellence In Technology, Education, And Science Reauthorization Act Of 2010

Rafols, I, & Meyer, M. (2010). Diversity and network coherence as indicators of interdisciplinarity: Case studies in bionanoscience. Scientometrics 82 (2):263-287.

Rafols, I., A. L. Porter, and L. Leydesdorff.(2010). Science Overlay Maps: A New Tool for Research Policy and Library Management. Journal of the American Society for Information Science and Technology 61 (9):1871-1887.

Renn, O. (1998). Three decades of risk research: Accomplishments and new challenges, *Journal of Risk Research, 1*(1), 49-71.

Roe A. (1983). Psychological approaches to creativity in science: The creativity question. Eds. Albert Rothenberg and Carl Hausman. Durham, NC: Duke University Press Duke University Press. (165-173).

Roessner, D. (2000). Quantitative and qualitative methods and measures in the evaluation of research. *Research Evaluation, 9*(2), 125-132.

Roessner, D. et al. (2007). Measuring the Interdisciplinarity of a Body of Research, Presented at the annual meeting of the American Evaluation Association, Baltimore, MD, November 8, 2007, http://higherlogicdownload.s3.amazonaws.com/EVAL/271cd2f8-8b7f-49ea-b925-e6197743f402/UploadedImages/2007-02-Roessner_D-Measuring.pdf

Roessner, R. (2007): News - Editorial: Show me the data, *Journal of Cell Biology* 2007 179:1091-1092., doi:10.1083/jcb.200711140, Available at http://jcb.rupress.org/content/179/6/1091.full

Rose, S., Shipp, S., Lal, B., & Stone, A. (2008). Measuring innovation and intangibles: A business perspective. Retrieved from http://www.athenaalliance.org/pdf/MeasuringInnovationandIntangibles-STPI-BEA.pdf

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika, 70*, 41-55.

Rosenbaum, P. R. (2005), An exact, distribution free test comparing two multivariate distributions based on adjacency, *Journal of the Royal Statistical Society*, B, 67, 515-530.

Rothschild, M., & Stiglitz, J. E. (1970). Increasing risk I: A definition. *Journal of Economic Theory*, Elsevier, vol. 2(3), pages 225-243.

Rosalie R. (2002) edited, "Benchmarking Evaluation of Public Science and Technology Programs in the United States, Canada, Israel, and Finland," Proceedings of a Workshop, TEKES--National Technology Agency of Finland, Embassy of Finland, Washington, DC.

Schubert, A., Glänzel W., & Thijs, B. (2006). The weight of author self-citations: A fractional approach to self-citation counting. *Scientometrics, 67*, 503-514. doi:10.1007/s11192-006-0056-6

Schumpeter, J. (1942). *Capitalism, socialism and democracy*. New York, NY: Harper & Row.

Sharpe, W. F. (1966). Mutual fund performance [Supplemental material]. *Journal of Business*, *39*(1), 119-138. Retrieved from http://www.jstor.org/page/journal/jbusiness/about.html

Shea, D., et al. (2009) The DHS Directorate of Science and Technology: Key Issues for Congress, CRS Report RL34356

SCImago. (2007). SJR — SCImago Journal & Country Rank. Retrieved January 24, 2012, from http://www.scimagojr.com

Simonton, D. K. (2004). *Creativity in science: Chance, logic, genius, and zeitgeist*. Cambridge: Cambridge University Press.

Small, H. (1973). Co-citation in the scientific literature: A new measure of the relationship between two documents. *Journal of the American Society for Information Science, 24,* 265-269.

Sosa, R., & Gero, J. S. (2005). A computational study of creativity in design: the role of society. *Artificial Intelligence for Engineering Design Analysis and Manufacturing. 19*(4). 229-244. doi:10.1017/S089006040505016X

Sternberg, R. (1990). *Wisdom: Its nature, origins, and development*. New York, NY: Cambridge University Press.

Sternberg, R. J. (2003). *Wisdom, intelligence, and creativity synthesized*. Cambridge: Cambridge University Press. 2003.

Stirling, A. (1998). On the economics and analysis of diversity. Retrieved from http://www.uis.unesco.org/culture/Documents/Stirling.pdf

Stoskopf, M. (2005). Observation and cogitation: How serendipity provides the building blocks of scientific discovery. *Ilar Journal, 46*(4), 332-7. Retrieved from http://dels-old.nas.edu/ilar_n/ilarjournal/46_4/pdfs/v4604Stoskopf.pdf

Stumpf, H. (1995). Scientific creativity: A short overview. *Educational Psychology Review. 7*(3), 225-241.

Swanson, D. R. (1986). Fish oil, Raynaud's syndrome, and undiscovered public knowledge. *Perspectives in Biology and Medicine* 30(1): 7-18.

Talley, E. (2011). Database of NIH grants using machine-learned categories and graphical clustering. *Nature Methods*, 8, 443-444.)

The National Academies. (2007). Rising Above the Gathering Storm: Energizing and Employing America For a Brighter Economic Future, Committee on Prospering in the Global Economy of the 21st Century. An Agenda for American Science and

Technology, National Academy of Sciences, National Academy of Engineering, Institute of Medicine. Retrieved from http://www nap edu/catalog/11463 html

Thijs, B., & Glänzel W. (2006). The influence of author self-citations on bibliometric meso-indicators. The case of European universities. *Scientometrics, 66*(1), 71-80. doi:10.1007/s11192-006-0006-3

Tobin, J. (1958). Liquidity preference as behavior towards risk, *The Review of Economic Studies*, 25, 65-86.

Treynor, J. (1965). How to rate management investment funds. *Harvard Business Review, 43*(1), 63-75.

U.S. Supreme Court JACOBELLIS v. OHIO, 378 U.S. 184 (1964) 378 U.S. 184, Jacobellis V. Ohio. Appeal From The Supreme Court Of Ohio. No. 11. Argued March 26, 1963. Restored to the calendar for reargument April 29, 1963. Reargued April 1, 1964. Decided June 22, 1964.

van Raan, A. F. J. (2002). Assessment of the scientific basis of interdisciplinary, applied research Application of bibliometric methods in Nutrition and Food Research. *Research Policy, 31*, 611-632.

van Raan, A. F. J. (2003). The use of bibliometric analysis in research performance assessment and monitoring of interdisciplinary scientific developments. *Technology Assessment-Theory and Practice*, *1*, 20-29.

van Raan, A. F. J. (2004). Measuring science: Capita selecta of current main issues. In H. F. Moed, W. Glänzel, & U. Schmoch (Eds.), *Handbook of Quantitative Science and Technology Research*, (pp. 19-50). Dordrecht: Kluwer Academic Publishers.

van Raan, A. F. J. (2004). Sleeping beauties in science. *Scientometrics, 59*, 461-466. Retrieved from http://www.cwts.nl/tvr/documents/avr-slbeau-scientom.pdf

van Raan, A. F. J. (2005). Fatal attraction: Conceptual and methodological problems in the ranking of universities by bibliometric methods. *Scientometrics, 62*(1), 133-143.

van Raan, A. F. J. (2006). Comparison of the Hirsch-index with standard bibliometric indicators and with peer judgment for 147 chemistry research groups. Scientometrics, Vol. 67, No. 3, 491–502. Retrieved from http://arxiv.org/ftp/physics/papers/0511/0511206.pdf

Wang, N. (2012). Transformative Research, Chinese Perspectives, *Workshop on Transformative Research: Ethical and Societal Implications'* held in Arlington, VA March 8 & 9, 2012 http://digital.library.unt.edu/ark:/67531/metadc84363/m1/46/

Wageman, S. (2004), Risk Management on Research and Development Projects AACE International Transactions, Morgantown, Pages R171.

Wagner, C. S., Roessner, J. D., Bobb, K., Thompson Klein, J., Boyack, K. W., Keyton, J., Rafols, I., & Borner, K. (2009). *Approaches to understanding and measuring interdisciplinary scientific research (IDR): A review of the literature*. Arlington, VA: SRI.

Waltman, L., van Eck, N. J., van Leeuwen, T. N., Visser, M. S., & van Raan, A. F. J.(2010). Towards a new crown indicator: Some theoretical consideration. Retrieved from http://arxiv.org/PS_cache/arxiv/pdf/1003/1003.2167v2.pdf.

Werner, M., & Lutz, B. (2010). How accurately does Thomas Kuhn's model of paradigm change describe the transition from the static view of the universe to the big bang theory in cosmology? A historical reconstruction and citation analysis. *Scientometrics, 84*, 441–464 doi:10.1007/s11192-009-0107-x

Yin, R. K. (1984). Case study research: Design and methods. Newbury Park, CA: Sage.

Zhang, J. (2007). A study of academic entrepreneurs using venture capital data. IZA Discussion Papers No. 2992. Available from ftp://repec.iza.org/RePEc/Discussionpaper/dp2992.pdf.

Zucker, L. G., & Darby, M. R. (1997). Individual action and the demand for institutions. *American Behavioral Scientist, 40*(4), 502-513.

Zucker, L. G., Darby, M. R., & Armstrong, J. (1998) Geographically localized knowledge: Spillovers or markets? *Economic Inquiry, 36*(1), 65-86.

Zucker, L. G., & Darby M. R (2001) Capturing technological opportunity via Japan's star scientists: Evidence from Japanese firms' biotech patents and products. *The Journal of Technology Transfer, 26*(1-2), 37-58.

Zucker, L. G., Darby, M. R., & Armstrong, J. S. (2001). Commercializing knowledge: University science, knowledge capture, and firm performance in biotechnology. National Bureau of Economic Research Working Paper Series. Retrieved from http://www.nber.org/papers/w8499

Zucker, L. G., Darby, M. R., & Torero, M. (2002). Labor mobility from academe to commerce. *NBER Journal of Labor Economics, 20*, 629-660. Retrieved from www.nber.org/papers/w6050

Zucker, L. G., & Darby M. R. (2002). Going public when you can in biotechnology. National Bureau of Economic Research Working Paper Series. Retrieved from http://www.nber.org/papers/w8954

Zucker L. G., & Darby, M. R. (2003). Grilichesian breakthroughs: Inventions of methods of inventing and firm entry in nanotechnology. National Bureau of Economic Research Working Paper Series. Retrieved from http://www.nber.org/papers/w9825.pdf

Zucker, L. G., & Darby, M. R. (2003).Growing by leaps and inches: Creative destruction, real cost reduction, and inching up. *Economic Inquiry, 41*(1), 1-19.

Zucker, L. G., Darby, M. R., & Brewer, M. B. (2003). Intellectual human capital and the birth of US biotechnology enterprises. In J. Hand & B. Lev (Eds.), *Intangible assets* (pp.185-206). Oxford: Oxford University Press.

Zucker, L. G., & Darby, M. R. (2006). Movement of star Scientists and engineers and igh-tech firm entry. National Bureau of Economic Research Working Paper Series. Retrieved from http://www.nber.org/papers/w12172

Appendix A: Questionnaire for Program Managers

Is your program designed to fund high-risk, high-reward research? (yes, no)

Are teams allowed to apply for the program? (yes, no)

Does your program fund goal-oriented research, exploratory research, or both? (if both, estimate fraction of each)

What is your funding type? (Individual grant, Project funding, Contract, Cooperative agreement, other)

Which aspects of the scientist(s) are criteria in making funding decisions? (Please rate each as Necessary, Important, Not Important)

- Career stage
- Grant history
- Track record of high quality research
- Track record of highly innovative research
- Interdisciplinarity of PI/team
- Group leadership qualification/Project management skills
- Research Environment
- Other (list)

Which aspects of the research problem are criteria in making funding decisions? (Please rate each as Necessary, Important, or Not Important)

- Scientific Merit
- Riskiness of proposal
- Interdisciplinarity of problem
- Importance of problem to field
- Potential payback to society
- New direction from PI's previous research
- Not suitable for funding through other mechanisms/sources
- Relevance to agency/organization mission
- Other (list)

Does your program have working definitions of "innovative," "transformative," "high-risk, high- reward"? If so, what are they?

What do you believe are the strengths of your program? (open)

What do you believe are the weaknesses of your program? (open)

We are seeking to understand 'best practices' of high-risk, high-reward research programs. Has an evaluation of your program been performed? If not, have you developed metrics for measuring program success?

Appendix B: Evolution of the Pioneer Program

Table B1.1

NDPA Process Changes in Detail: Candidate Recruitment Emphasis

Aspect of NDPA	FY 2004	FY 2005	FY 2006	FY 2007	FY 2008
Emphasis given in Notice or RFA	"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."	"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."	"Women and members of groups underrepresented in biomedical or behavioral research are especially encouraged to apply. Investigators at all career levels who are currently engaged in research are eligible to apply."
Definition of "pioneering" and "award" given in RFA	Not specifically defined	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	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.

Source: Lal et al. (2010)

Table B1.2

NDPA Process Changes in Detail: Selection Process – Phase Mechanics and Candidate/Evaluator Participation

	FY 2004	**FY 2005**	**FY 2006**	**FY 2007**	**FY 2008**
Process	**5 Phases**	**5 Phases**	**4 Phases**	**3 Phases**	**3 Phases**
Phase 1 (Nominees)	Phase 1: 1,331 nomination packages (self-nominees and individuals nominated by someone else) – screened for responsiveness by 29 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.	No nomination phase(s)	No nomination phase(s)
Phase 2 (Responsive Nominees)	Phase 2: 936 responsive nomination packages were reviewed by a first group of 49 external evaluators (yes/no vote)	Phase 2: 567 nominees were deemed responsive and were reviewed by 47 external evaluators (yes/no vote)	No initial (yes/no) screening by external evaluators in FY 2006		
Phase 3 (Applicants)	Phase 3: 245 individuals invited to submit a full application package to be reviewed by a second group of 29 external evaluators – scored on a 7-point scale	Phase 3: 283 individuals invited to submit a full application package to be reviewed by a second group of 37 external evaluators – scored on a 5-point scale; "top 4" votes assigned	Phase 2: 406 responsive individuals were reviewed by a group of 80 external evaluators – scored on a 5-point scale; "top 4" votes assigned	Phase 1: 449 individuals submitted a full application and were reviewed by a group of 69 external evaluators – scored on a 5-point scale; "top 4" votes assigned	Phase 1: 440 individuals submitted a full application and were reviewed by a group of 74 external evaluators – scored on a 5-point scale; "top 4" votes and "ideal candidate" designations assigned
Phase 4 (Interviewees)	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	Phase 2: 25 individuals invited for an interview on July 9-11th, 2007 with a panel of 14 experts	Phase 2: 25 individuals invited for an interview on July 9-11th, 2008 with a panel of 14 experts
Phase 5 (Awardees)	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	Phase 3: 12 awards made on September 19, 2007	Phase 3: 16 awards made on September 22, 2008

Table B1.3

NDPA Process Changes in Detail: Selection Criteria

	FY 2004	FY 2005	FY 2006	FY 2007	FY 2008
Criterion 1	Innovation/creativity—Examples: 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, multidisciplinary 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.	The scientific problem to be addressed: The biomedical or behavioral significance/importance of the problem, the likelihood that, if successful, the project will have a significant impact on this problem, and the innovativeness of the project.	The scientific problem to be addressed: The biomedical or behavioral significance/importance of the problem; the likelihood that, if successful, the project will have a significant impact on this problem; and the innovativeness of the project.
Criterion 2	Intrinsic motivation/ enthusiasm/intellectual energy—Examples: Is the applicant willing to take scientific risks and show persistence in the face of adversity? Is the applicant comfortable with uncertainty (i.e., able to see gray areas as opportunities for new insights)? 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 – "the extent of innovative activities relative to the applicant's career stage"); 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	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	The investigator: Evidence for the investigator's claim of innovativeness/creativity (innovation density) and the demonstrated ability of the investigator to devote at least 51% of his/her effort to activities supported by the Pioneer Award

(table continues)

	FY 2004	FY 2005	FY 2006	FY 2007	FY 2008
Criterion 3	**Potential for or actual scientific leadership; evidence of, or potential for, effective communication/educator skills**—Examples: 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; distinctness 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	**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	**The suitability for Pioneer Award mechanism:** Evidence that the proposed project is of sufficient risk/impact to make it more suitable for a Pioneer Award than for the traditional NIH grant mechanism and that it is distinct from other research previously

Source: Lal et al. (2010)

Appendix C: Pioneer Research Characteristics

Table C1.1

Summary of Pioneer Research

Pioneer Code	Research Area[a]	What did the awardees propose to do with the NDPA funds in their applications?[b]	How did the proposal differ from the research conducted by Pioneers before receiving the award?[b]	How did the actual NDPA research differ from the NDPA proposals?[b]	What have the NDPA funds allowed awardees to do that would not be possible with traditional funding sources? (Pioneer)[c]	What are the applications of awardee research to the diagnosis and treatment of disease? (Pioneer)[c]	In what ways has the NDPA played a role in changing the awardees' research fields over the past five years? (Experts)[d]
1	Quantitative and mathematical biology	Test a specific hypothesis or set of hypotheses	Broaden the focus of their research to a grander systems level	Broad research goals were met or continue to progress	• Follow a natural research trajectory • Take a long term view • Spend more time on lab research	Awareness of potential long-term applications	• NDPA work has influenced other researchers • Too early to tell • Connected formerly disparate research fields
2	Molecular and cellular biology	Pursue and open-ended research objective	Apply previous research methods and ideas to new biomedical issues	Original plan evolved into the research conducted under the NDPA	• Undertake resource-intensive projects • Take a long term view	Awareness of potential long-term applications	• Changed prevailing wisdom/ provided novel perspective • Major contributor
3	Molecular and cellular biology	Develop a new technology or approach to research	Broaden the focus of their research to a grander systems level	Original plan evolved into the research conducted under the NDPA	• Follow a natural research trajectory • Take a long term view	Studies with implications for disease treatment and diagnosis underway	• No significant contributions • NDPA work has influenced other researchers

(table continues)

Pioneer Code	Research Area[a]	What did the awardees propose to do with the NDPA funds in their applications?[b]	How did the proposal differ from the research conducted by Pioneers before receiving the award?[b]	How did the actual NDPA research differ from the NDPA proposals?[b]	What have the NDPA funds allowed awardees to do that would not be possible with traditional funding sources? (Pioneer)[c]	What are the applications of awardee research to the diagnosis and treatment of disease? (Pioneer)[c]	In what ways has the NDPA played a role in changing the awardees' research fields over the past five years? (Experts)[d]
4	Behavioral and social sciences	Test a specific hypothesis or set of hypotheses	Conduct new experiments that support their existing hypotheses	Broad research goals were met or continue to progress	• Follow a natural research trajectory • Take a long term view • Spend more time on lab research	Research already having an impact	• Increased the research field's visibility • Connected formerly disparate research fields
5	Molecular and cellular biology	Develop a new technology or approach to research	Remain in the same field but proposed a project with a distinctly different, lateral (in scope), focus	Original plan evolved into the research conducted under the NDPA	• Follow a natural research trajectory • Take a long term view • Spend more time on lab research	Discoveries of health-related applications within 10 year timeframe	• Major contributor
6	Molecular and cellular biology	Develop a new technology or approach to research	Broaden the focus of their research to a grander systems level	Original plan evolved into the research conducted under the NDPA	• Undertake resource-intensive projects • Follow a natural research trajectory	Awareness of potential long-term applications	• Major contributor
7	Physiological and integrative systems	Develop a new technology or approach to research	Apply previous research methods and ideas to new biomedical issues	Broad research goals were met or continue to progress	• Follow a natural research trajectory • Take a long term view	Studies with implications for disease treatment and diagnosis underway	• Major contributor • Developed new techniques
8	Instrumentation and engineering	Develop a new technology or approach to research	Apply previous research methods and ideas to new biomedical issues	Broad research goals were met or continue to progress	• Follow a natural research trajectory • Spend more time on lab research	Discoveries of health-related applications within 10 year timeframe	• Too early to tell • No significant contributions

Pioneer Code	Research Area[a]	What did the awardees propose to do with the NDPA funds in their applications?[b]	How did the proposal differ from the research conducted by Pioneers before receiving the award?[b]	How did the actual NDPA research differ from the NDPA proposals?[b]	What have the NDPA funds allowed awardees to do that would not be possible with traditional funding sources? (Pioneer)[c]	What are the applications of awardee research to the diagnosis and treatment of disease? (Pioneer)[c]	In what ways has the NDPA played a role in changing the awardees' research fields over the past five years? (Experts)[d]
9	Quantitative and mathematical biology	Develop a new technology or approach to research	Apply previous research methods and ideas to new biomedical issues	Original plan evolved into the research conducted under the NDPA	• Follow a natural research trajectory	Awareness of potential long-term applications	• No significant contributions *(table continues)*
10	Behavioral and social sciences	Test a specific hypothesis or set of hypotheses	Conduct new experiments that support their existing hypotheses	Broad research goals were met or continue to progress	• Follow a natural research trajectory • Take a long term view • Spend more time on lab research • Improve their labs	Awareness of potential long-term applications	• Connected formerly disparate research fields • Major contributor
11	Pathogenesis and epidemiology	Test a specific hypothesis or set of hypotheses	Conduct new experiments that support their existing hypotheses	Broad research goals were met or continue to progress	• Undertake resource-intensive projects • Follow a natural research trajectory	Studies with implications for disease treatment and diagnosis underway	• Changed prevailing wisdom/ provided novel perspective
12	Molecular and cellular biology	Pursue and open-ended research objective	Broaden the focus of their research to a grander systems level	Broad research goals were met or continue to progress	• Undertake resource-intensive projects • Follow a natural research trajectory • Improve their labs	Discoveries of health-related applications within 10 year timeframe	• Major contributor • Increased the research field's visibility

Pioneer Code	Research Area[a]	What did the awardees propose to do with the NDPA funds in their applications?[b]	How did the proposal differ from the research conducted by Pioneers before receiving the award?[b]	How did the actual NDPA research differ from the NDPA proposals?[b]	What have the NDPA funds allowed awardees to do that would not be possible with traditional funding sources? (Pioneer)[c]	What are the applications of awardee research to the diagnosis and treatment of disease? (Pioneer)[c]	In what ways has the NDPA played a role in changing the awardees' research fields over the past five years? (Experts)[d]
13	Other	Develop a new technology or approach to research	Apply previous research methods and ideas to new biomedical issues	Broad research goals were met or continue to progress	• Follow a natural research trajectory	Studies with implications for disease treatment and diagnosis underway	• No significant contributions • Major contributor
14	Quantitative and mathematical biology	Pursue and open-ended research objective	Apply previous research methods and ideas to new biomedical issues	Broad research goals were met or continue to progress	• Take a long term view • Spend more time on lab research • Improve their labs	Awareness of potential long-term applications	• Connected formerly disparate research fields
15	Instrumentation and engineering	Develop a new technology or approach to research	Apply previous research methods and ideas to new biomedical issues	Broad research goals were met or continue to progress	• N/A	N/A	• Major contributor *(table continues)*
16	Molecular and cellular biology	Pursue and open-ended research objective	Broaden the focus of their research to a grander systems level	Original plan evolved into the research conducted under the NDPA	• Undertake resource-intensive projects • Follow a natural research trajectory • Take a long term view • Spend more time on lab research	Studies with implications for disease treatment and diagnosis underway	• NDPA work has influenced other researchers

Pioneer Code	Research Area[a]	What did the awardees propose to do with the NDPA funds in their applications?[b]	How did the proposal differ from the research conducted by Pioneers before receiving the award?[b]	How did the actual NDPA research differ from the NDPA proposals?[b]	What have the NDPA funds allowed awardees to do that would not be possible with traditional funding sources? (Pioneer)[c]	What are the applications of awardee research to the diagnosis and treatment of disease? (Pioneer)[c]	In what ways has the NDPA played a role in changing the awardees' research fields over the past five years? (Experts)[d]
17	Quantitative and mathematical biology	Develop a new technology or approach to research	Apply previous research methods and ideas to new biomedical issues	Broad research goals were met or continue to progress	• Follow a natural research trajectory • Take a long term view • Spend more time on lab research • Improve their labs	Research already having an impact	• Major contributor • No significant contributions
18	Physiological and integrative systems	Test a specific hypothesis or set of hypotheses	Conduct new experiments that support their existing hypotheses	Broad research goals were met or continue to progress	• Take a long term view • Spend more time on lab research • Improve their labs	Discoveries of health-related applications within 10 year timeframe	• Major contributor • Developed new techniques
19	Quantitative and mathematical biology	Pursue and open-ended research objective	Broaden the focus of their research to a grander systems level	Original plan evolved into the research conducted under the NDPA	• Follow a natural research trajectory • Improve their labs	N/A	• N/A *(table continues)*
20	Molecular and cellular biology	Pursue and open-ended research objective	Apply previous research methods and ideas to new biomedical issues	Broad research goals were met or continue to progress	• Follow a natural research trajectory • Spend more time on lab research	Research already having an impact	• NDPA work has influenced other researchers • Too early to tell • Connected formerly disparate research fields

Pioneer Code	Research Area[a]	What did the awardees propose to do with the NDPA funds in their applications?[b]	How did the proposal differ from the research conducted by Pioneers before receiving the award?[b]	How did the actual NDPA research differ from the NDPA proposals?[b]	What have the NDPA funds allowed awardees to do that would not be possible with traditional funding sources? (Pioneer)[c]	What are the applications of awardee research to the diagnosis and treatment of disease? (Pioneer)[c]	In what ways has the NDPA played a role in changing the awardees' research fields over the past five years? (Experts)[d]
21	Instrumentation and engineering	Develop a new technology or approach to research	Remain in the same field but proposed a project with a distinctly different, lateral (in scope), focus	Broad research goals were met or continue to progress	• Follow a natural research trajectory • Take a long term view • Spend more time on lab research	Discoveries of health-related applications within 10 year timeframe	• Major contributor
22	Molecular and cellular biology	Test a specific hypothesis or set of hypotheses	Remain in the same field but proposed a project with a distinctly different, lateral (in scope), focus	Original plan evolved into the research conducted under the NDPA	• Follow a natural research trajectory • Spend more time on lab research	Studies with implications for disease treatment and diagnosis underway	• No significant contributions • NDPA work has influenced other researchers

Source: Lal et al. (2011)

Note:

[a] Awardee applications to the NDPA.

[b] Awardee applications to the NDPA, publications in Web of Science, Pioneer interviews, Expert review.

[c] Pioneer interviews.

[d] Expert review.