

Big Data in Evaluating Transformative Scientific Research: Concepts and a Case Study

Bhavya Lal
Vanessa Peña

IDA Science and Technology Policy Institute (STPI)

Understanding Federal R&D Impact

Big Data: Measuring the Impact of the Government's Research and Development Investments

National Press Club

March 19, 2013

About Us:

Science and Technology Policy Institute

- Federally funded research and development center chartered by Congress
- Provide rigorous and objective analysis of science and technology policy issues for the White House Office of Science and Technology Policy and other offices within the executive branch of the U.S. government and federal agencies
- Conduct science and technology analysis to inform policy decision-makers

Today's Presentation

- Challenges in evaluating non-traditional research
 - Conceptual
 - Data related
- Selected case study

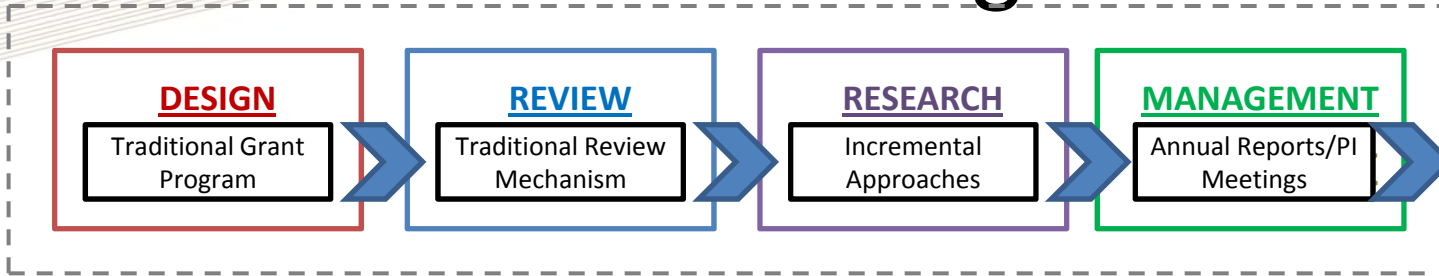
Background

- Belief that the research enterprise is becoming more conservative
- Programs have been created to support transformative research

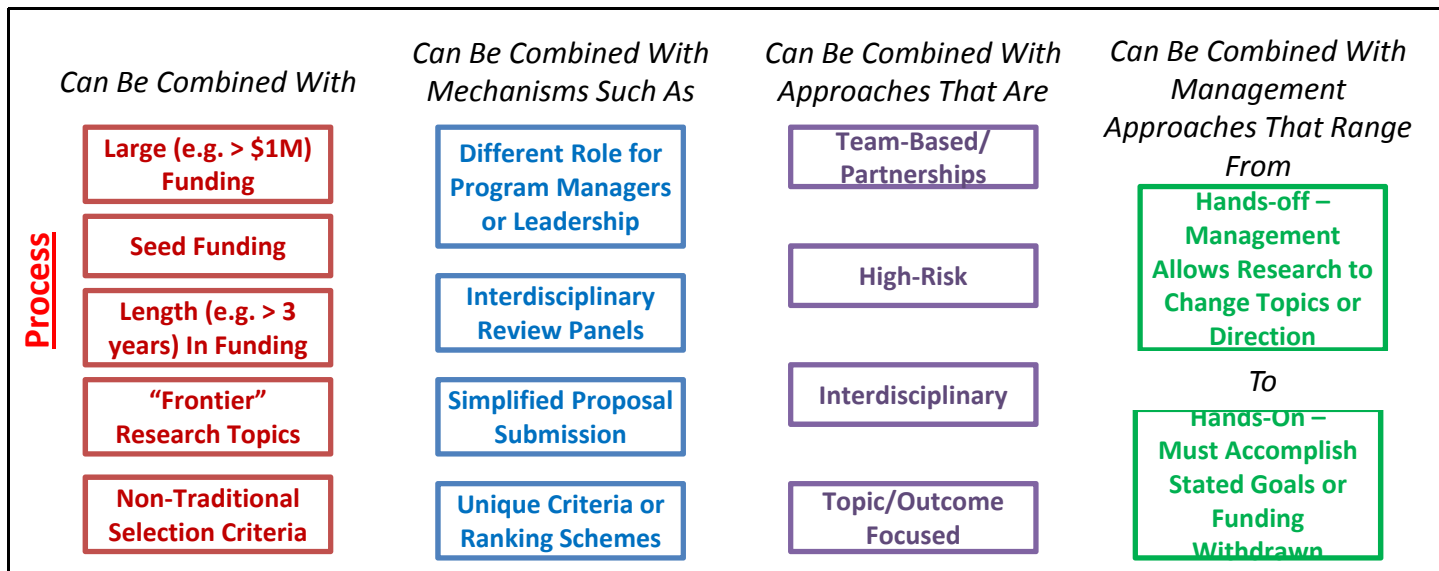
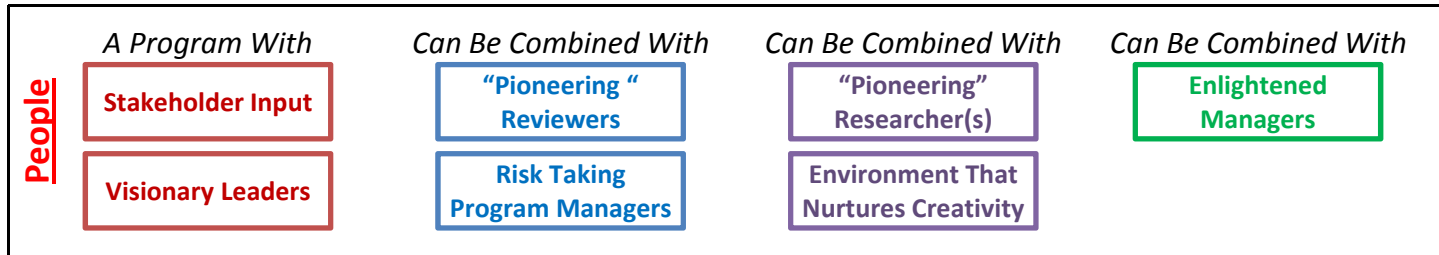


A Smorgasbord of Strategies in Federal Programs

Traditional Programs



Non-Traditional Programs



NIH Director's Pioneer Award Program

Non-Traditional Programs

DESIGN

REVIEW

RESEARCH

MANAGEMENT

OUTCOME
Will Achieve

Transformative Outcomes

And/Or

Incremental Outcomes

With

Spillovers:

- Create Community
- Institutional Change
- Train Students

<u>People</u>	A Program With	Can Be Combined "Pioneering" Reviewers	Can Be Combined "Pioneering" Researcher(s)	Can Be Combined With
---------------	----------------	--	--	-------------------------

<u>Process</u>	Can Be Combined With Large (e.g. > \$1M) Funding	Can Be Combined With Mechanisms Such As Simplified Proposal Submission	Can Be Combined With Approaches That Are High-Risk	Can Be Combined With Management Approaches That Range From Hands-off – Management Allows Research to Change Topics or Direction
	Length (e.g. > 3 years) In Funding			
	Non-Traditional Selection Criteria			

NSF Emerging Frontiers in Research & Innovation

DESIGN

REVIEW

RESEARCH

MANAGEMENT

OUTCOME

Will Achieve

Transformative Outcomes

And/Or

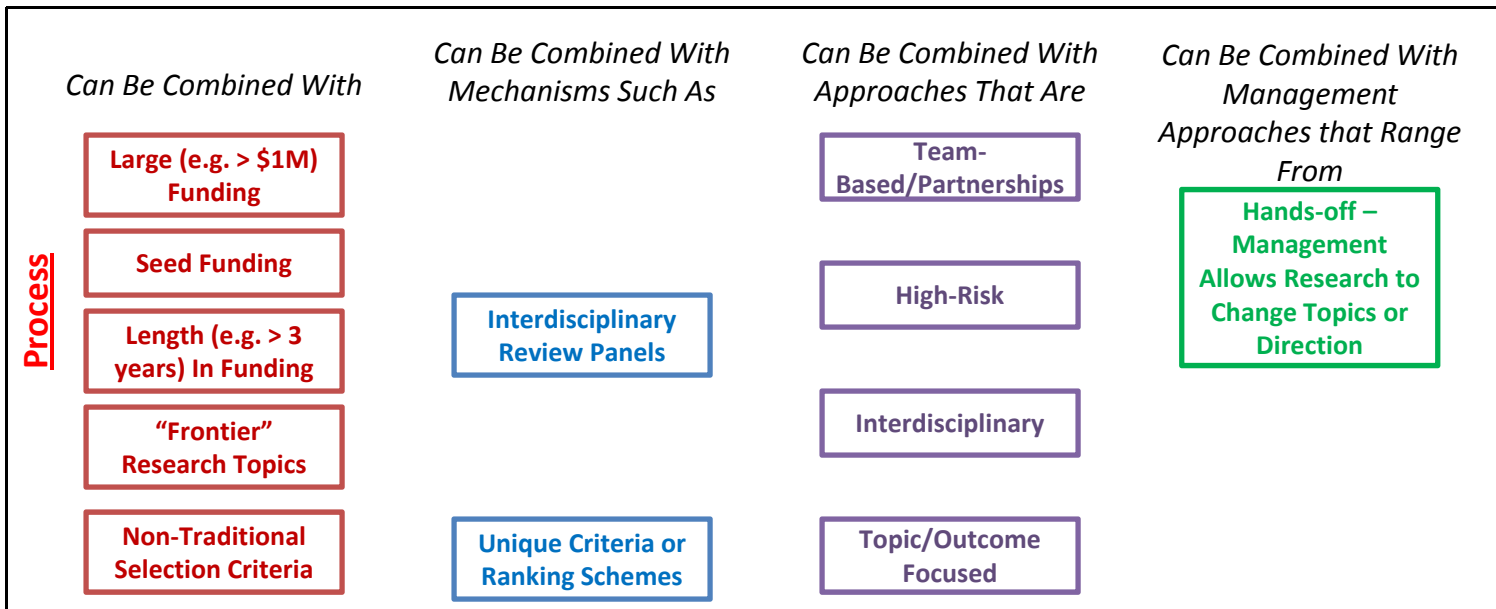
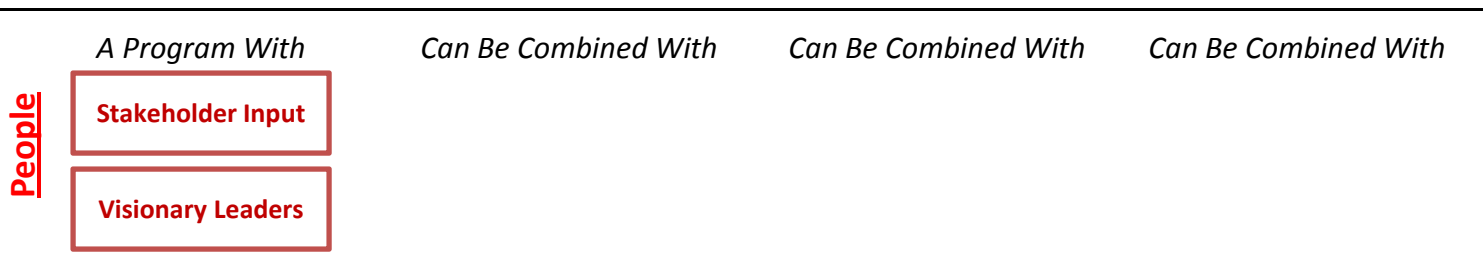
Incremental Outcomes

With

Spillovers:

- Create Community
- Institutional Change
- Train Students

Non-Traditional Programs



Policy Question

- Set-aside transformative research programs are viewed as taking resources from traditional programs
- Do they offer more transformative outcomes in return?

Conceptual Challenge

- How does one operationalize and measure “transformative” research?

Defining Transformative Research

Tolstoyian thesis: although all conventional practitioners in the life sciences may be said to be conventional in the same way, all rebels seem to rebel in their own particular fashion

The dominant paradigm with respect to the definition is
“I know it when I see it”

Working Definition

Transformative means:

- Research that involves ideas, discoveries, or tools that
 - a. **radically change our understanding** of an important existing scientific or engineering concept or educational practice or
 - b. leads to the **creation of a new paradigm or field** of science, engineering, or education
- Characterized not only by exceptional innovation, but also by the **conscious taking of risks** in the choice of its research subjects and methods

Operationalizing Transformative Research is Challenging

- (1) No known metrics for transformative research except **in retrospect**
 - Darwin's theory of natural selection (replacement of Lamarckism as the mechanism for evolution)
 - Development of quantum mechanics (supplanted classical mechanics)
 - Plate tectonics (which combined the hypotheses of continental drift and sea floor spreading into the theory of plate tectonics) replaced the static geosynclinals theory in describing continental drift
- If an evaluation is sought too soon, the transformative nature of the results may not be evident

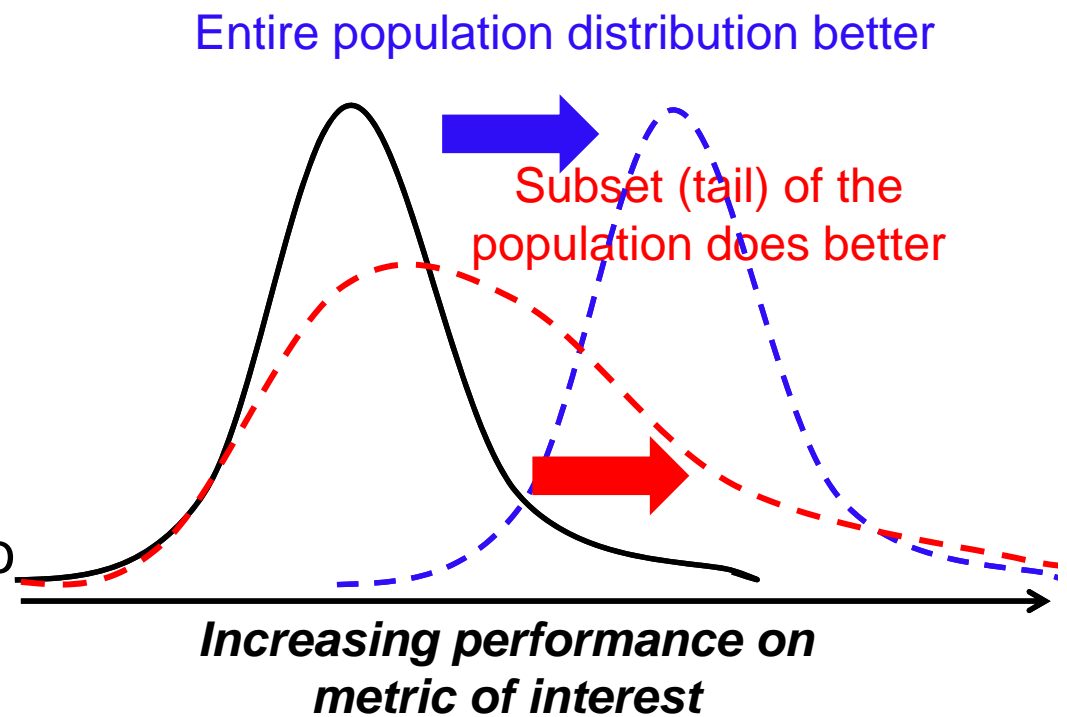
Operationalizing Transformative Research is Challenging (2)

- (2) Results may be controversial - especially if contradicting prevailing wisdom - not even possible to publish them
 - George Akerlof's path-breaking economic theory of asymmetric information and adverse selection, which ultimately garnered him a Nobel Prize, was initially rejected by three major economics journals
 - Barry Marshall and Robin Warren's discovery that peptic ulcers are caused by bacterium *H. pylori* and not stress brought them ridicule in the biomedical community
 - Barbara McClintock on cytogenetics was ignored for decades

Imagine Einstein submitting an application in 1905: "I propose to explore the possibility that time slows down as things speed up."

Operationalizing Transformative Research

- In a traditional research portfolio, we can expect that:
 - Some fraction are breakthroughs
 - Some are incremental
 - If a set-aside program is focused specifically on producing breakthrough research, it should lead to
 - A larger number of research breakthroughs
- And/Or
- At least a few bigger breakthroughs



Evaluation Approach

- Identify comparison group(s)
- Identify indicators for comparison
 - Whether impact is (more/differently) transformative
 - Whether research approaches are (more) unusual
- Compare outcomes

Data Challenge

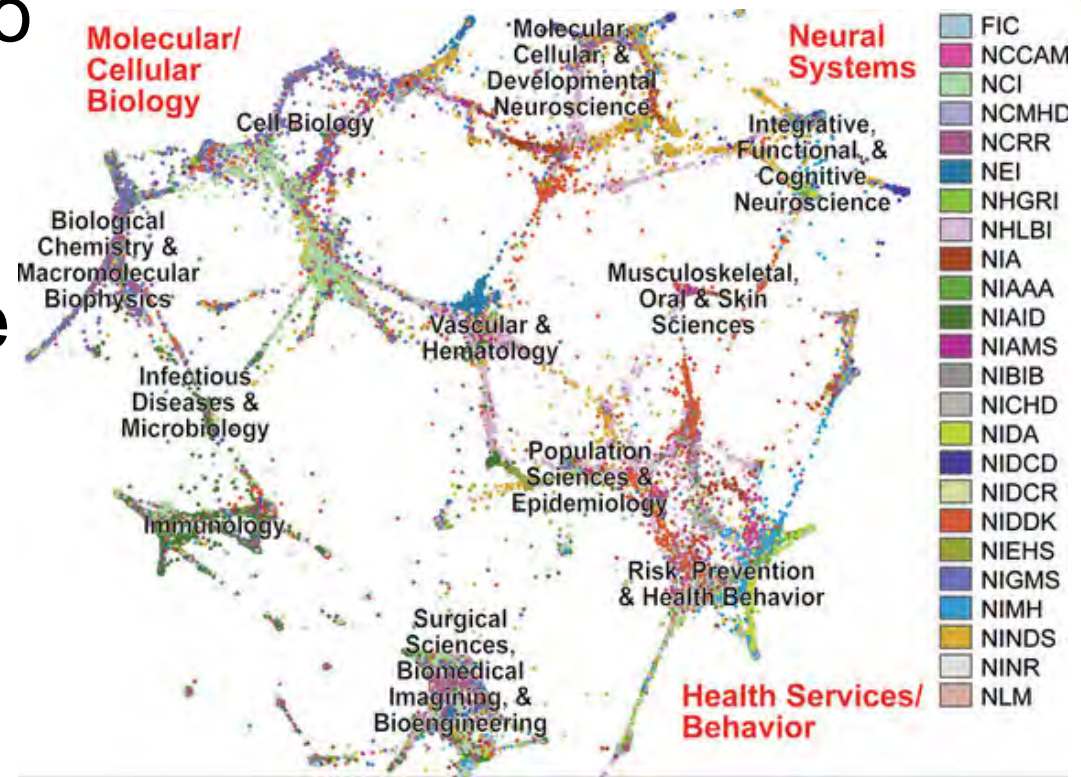
- Comparison research has to be “similar enough” to the transformative research
- Finding comparable comparison groups is an immense challenge

Limitations of Bibliometrics

- Traditional bibliometrics cannot tell the difference between traditional and transformative outcomes

Big Data Helped Us Partly

- “Big data” refers to datasets whose size is beyond the ability of typical database software tools to capture, store, manage, and analyze
- Potential for emerging tools to address data challenges



NIH Maps database: <http://nihmaps.org/index.php>

CASE STUDY: USING BIG DATA IN EVALUATING A TRANSFORMATIVE RESEARCH PROGRAM

Transformative Research Program Outcome Evaluation

- Does the program's research produce **higher impact** as compared to research funded by **other programs**?
- Are the approaches used by the program's researchers **more innovative** as compared to research funded by **other programs**?

Uses of Text Mining

- Comparison Group Selection, selecting researchers working in similar sub-fields
- Data Analysis:
 - Selecting subject-specific experts to assess research
 - Comparing similarity of research topics from test and comparison groups

Evaluation Methods

Test and Comparison Groups

Groups	Source
Test Group – Test Researchers	35 Researchers Transformative Research Program
Group 1 – Similar Researchers	35 Researchers One Program in the Same Funding Agency
Group 2 – Similar Program	39 Researchers High-Prestige Program from a Different Funding Organization
Group 3 – Program Finalists	35 Researchers Applied to Same Funding Agency
Group 4 – Random Researchers	Portfolios on avg. 80 Researchers Many Programs from Same Funding Agency

Selection of Comparison Group 1

- Match researcher and research dimensions:
 - Research area
 - Year of award
 - Years since degree
 - Institutional prestige
 - Prior research program funding
 - Terminal degree(s) received
 - Receipt of early career awards

Comparison Groups	Matched?
Group 1 – Similar Researchers	Yes
Group 2 – Similar Program	By design
Group 3 – Program Finalists	
Group 4 – Random Researchers	No

- Matching helped us select an equal or balanced distribution of variables

Identifying Similar Research Areas

- All other matching indicators are straightforward
 - quantitative
- Research area is a qualitative indicator
 - Accessed and gleaned from award proposal text
 - 40,000 award proposals used in the analysis, typical award proposal title and abstract length ~500 words

Selecting Similar Research Areas

- Topic Modeling and Similarity Index
 - Based on topic modeling using Latent Dirichlet Allocation to assess similarity (Blei 2003, Talley 2011)
 - Topic co-occurrence in unstructured text in NIH award titles and abstracts from 2007 to 2010
 - Values 0 to 1, representing proportion of co-occurrence of topics (1 being identical)
- Each award was matched with 7 - 200 similar awards with their respective index, these values were used in matching
- Selected 35 researchers from more than 12,000 NIH awardees

Data Analyses

Qualitative and Quantitative Analyses

Examination of researchers' publications:

1. Expert Assessments – quantitative / qualitative
2. Text-Mining Analysis – quantitative / qualitative
3. Bibliometric Analysis – quantitative

Test and Comparison Groups	Experts	Text-Mining	Bibliometrics
Test	162	354	3,287
Group 1 – Select Researchers	154	243	3,274
Group 2 – Similar Program	194	117	3,313
Group 3 – Program Finalists	--	--	2,298
Group 4 – Random Researchers	--	11,431	14,352

Expert Assessments

- 94 experts conducted over 1,500 ratings of over 500 publications
- For each test and comparison group researcher, the five most impactful publications were identified (represented the “body of work” for a researcher)
- “Body of work” and individual papers reviewed at least three times

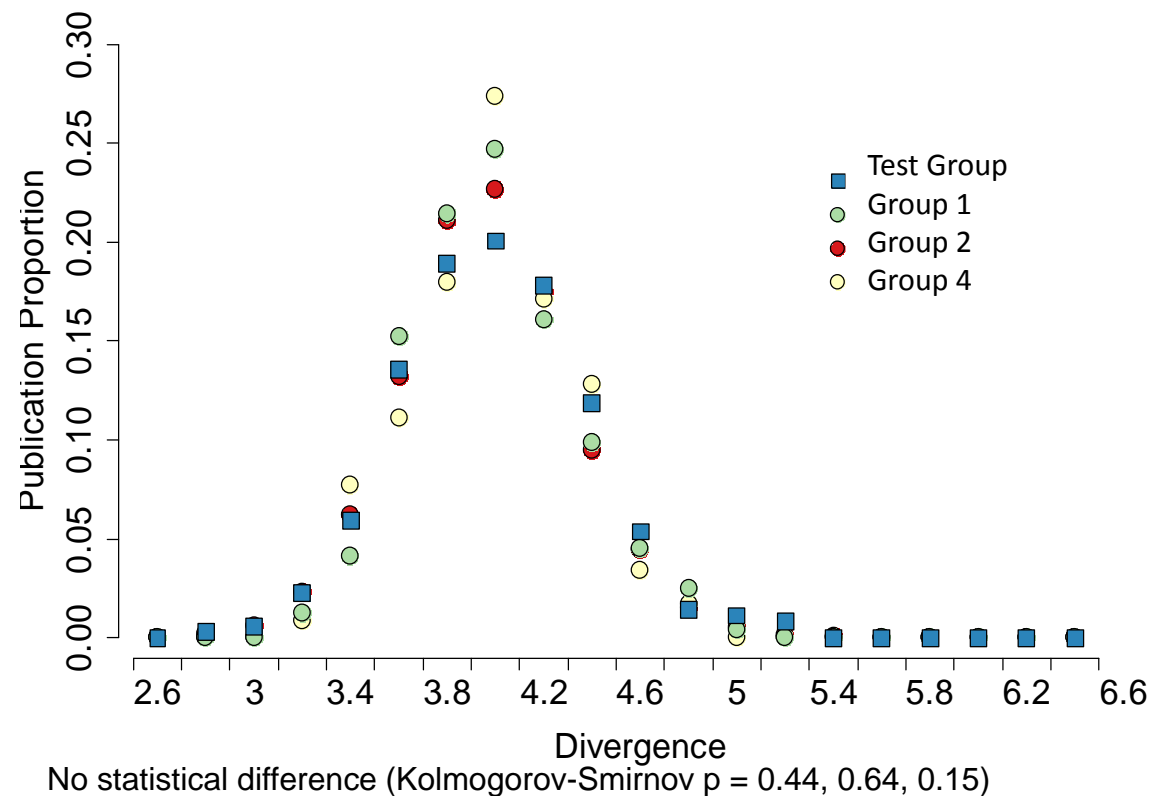
Expert Assessments			
Impact		Innovativeness	
Body of Work	Individual Paper	Body of Work	Individual Paper

Identifying Subject-Specific Experts

- Sought experts from past NIH awardees in similar fields to the test program, used field similarity from topics in researchers' awards (~730)
 - Excluded awards used in comparison groups
 - Selected those with recent early career awards
 - Selected those with previous long-term awards (e.g., NIH/MERIT)
- Supplemented as needed via contact with NIH
- Combined lists and verified suitability

Analyzing Research Topic Divergence

- Investigated how similar research was to broader scientific community
- Used award-attributed publications in PubMed through NIH SPIRES from awards 2007-2010
- Helps support study findings and could be used to identify emerging research



Opportunities for Use of Big Data and Text Mining in Evaluations

- Selection of comparison groups
 - Topic modeling and similarity metrics help match closely related topics
- Selecting experts for assessing research
 - Subject-specific experts can be identified using topic modeling approaches
- Comparing research similarity
 - Validates results and may identify innovative research areas not similar to research in the broader scientific community
 - BUT...usefulness is unclear

Thank you!

For our full evaluation study:

http://commonfund.nih.gov/pdf/IDA_Paper_P_4899.pdf

Acknowledgements: Seth Jonas, Elizabeth Lee, Amy Richards, and Alyson Wilson (STPI), and Ned Talley (NIH)

Bhavya Lal – blal@ida.org

Vanessa Peña – vpena@ida.org

1899 Pennsylvania Avenue NW

Washington, D.C. 20006