

Reproducibility

Is this the 21st Century version of
“Turtles All the Way Down?”



"If everything must have a cause, then God must have a cause. If there can be anything without a cause, it may just as well be the world as God, so that there cannot be any validity in that argument. It is exactly of the same nature as the Hindu's view, that the world rested upon an elephant and the elephant rested upon a tortoise; and when they said, 'How about the tortoise?' the Indian said, 'Suppose we change the subject.'"

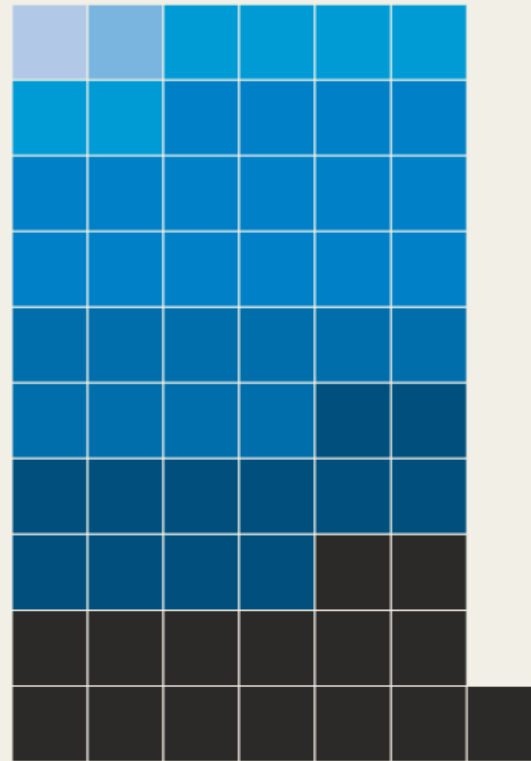
- *Bertrand Russell*

RELIABILITY TEST

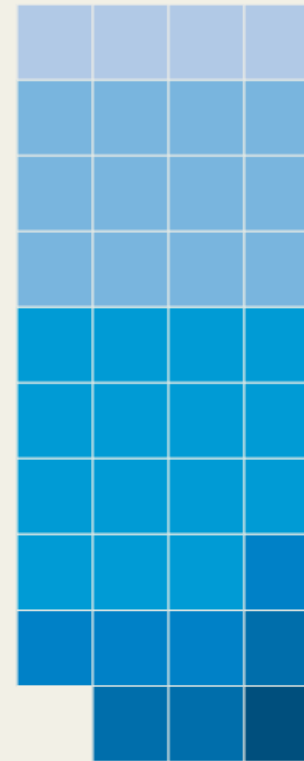
An effort to reproduce 100 psychology findings found that only 39 held up.* But some of the 61 non-replications reported similar findings to those of their original papers.

Did replicate match original's results?

NO: 61



YES: 39



Replicator's opinion: How closely did findings resemble the original study:

- Light Blue: Virtually identical
- Light Blue: Extremely similar
- Light Blue: Very similar
- Light Blue: Moderately similar
- Light Blue: Somewhat similar
- Light Blue: Slightly similar
- Black: Not at all similar

* based on criteria set at the start of each study

According to **the replicators' qualitative** assessments . . . only 39 of the 100 replication attempts were successful. (There were 100 completed replication attempts on the 98 papers, as in two cases replication efforts were duplicated by separate teams.)
But whether a replication attempt is considered successful is not straightforward.

Replication is difficult, as is measuring its success.

On the basis of criteria set at the start of each study, there were 39 replications and 61 non-replications. But replicators were also asked how similar — in their opinion — their overall findings, and the key effect measured in their replication, were to the original study. **This nuanced approach revealed that 24 formal non-replicates had at least moderately similar findings.**

Lots of variables + new criteria = harder still.

Correlational tests suggest that **replication success was better predicted by the strength of original evidence** than by characteristics of the original and replication teams.

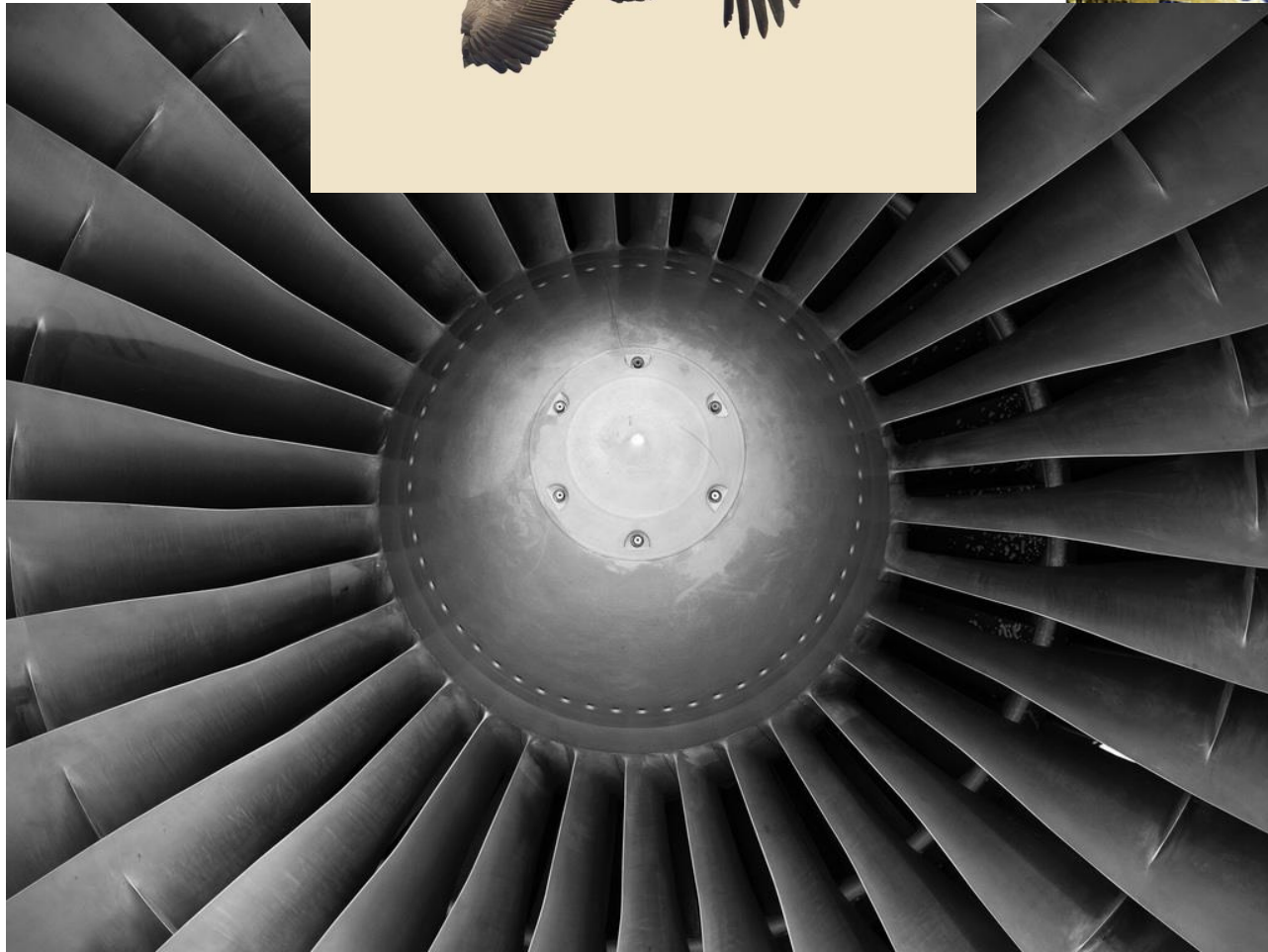
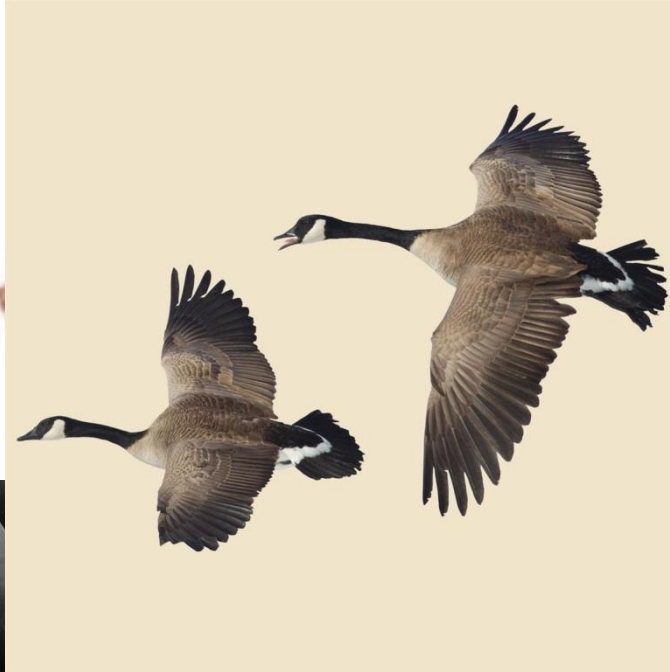
Better evidence is . . . better evidence.

What is Reproducible?

- Earthquakes may be impossible to reproduce
 - Ethical issues, feasibility, “once and done” aspect
- Mathematics are easy to reproduce
- Botanical studies are harder to reproduce
- Human studies are very difficult to reproduce
- Psychology studies are perhaps the most difficult – human studies on soft variables
- First, best, last

Reproducibility Depends on Basic Epistemological Agreement

- What is “poison”?
 - Is our atmosphere just mildly poisonous?
- What is “aging”?
 - Protons age, the universe ages
- What is “human”?
 - Most of our cells do not contain human DNA
- What is “sick”?
 - Bacterial infections are commonplace, but it takes special circumstances for them to make us sick





B944 - 02/06/2014

02/12/2014

02/11/2014

02/10/2014

02/09/2014

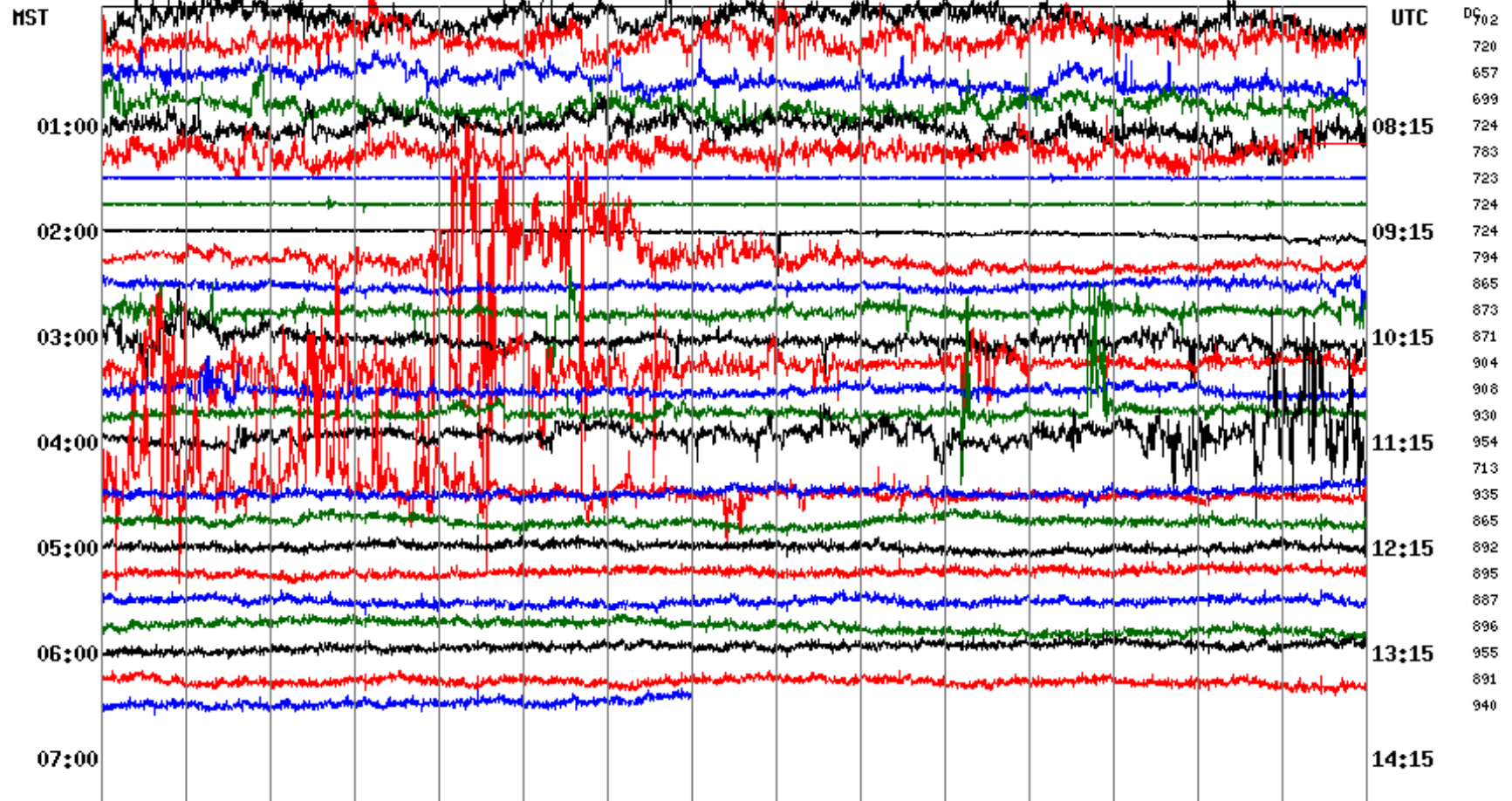
02/08/2014

02/07/2014

Feb 6, 2014

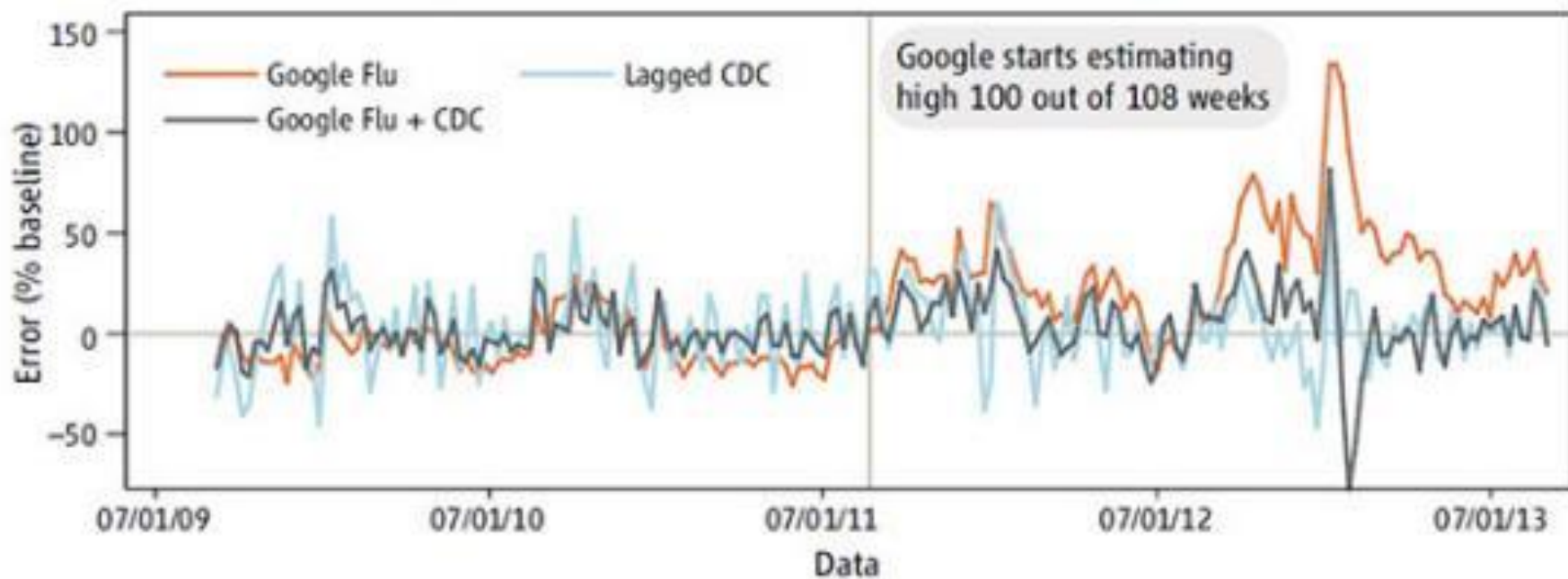
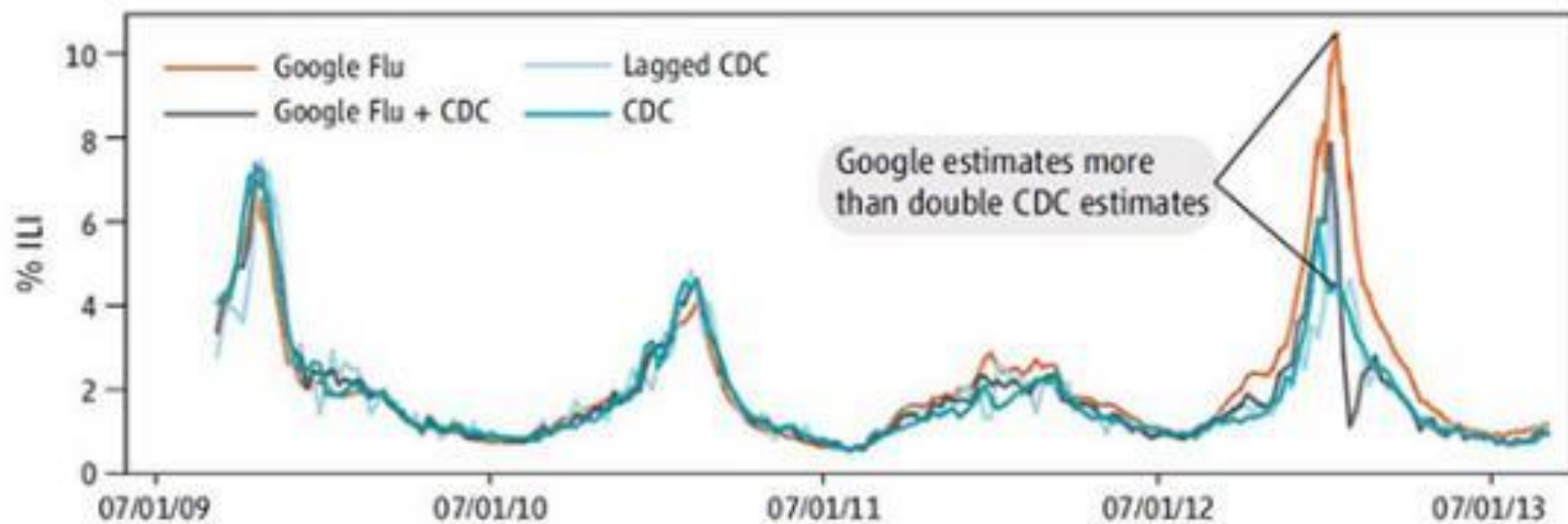
B944 EHZ PB 01

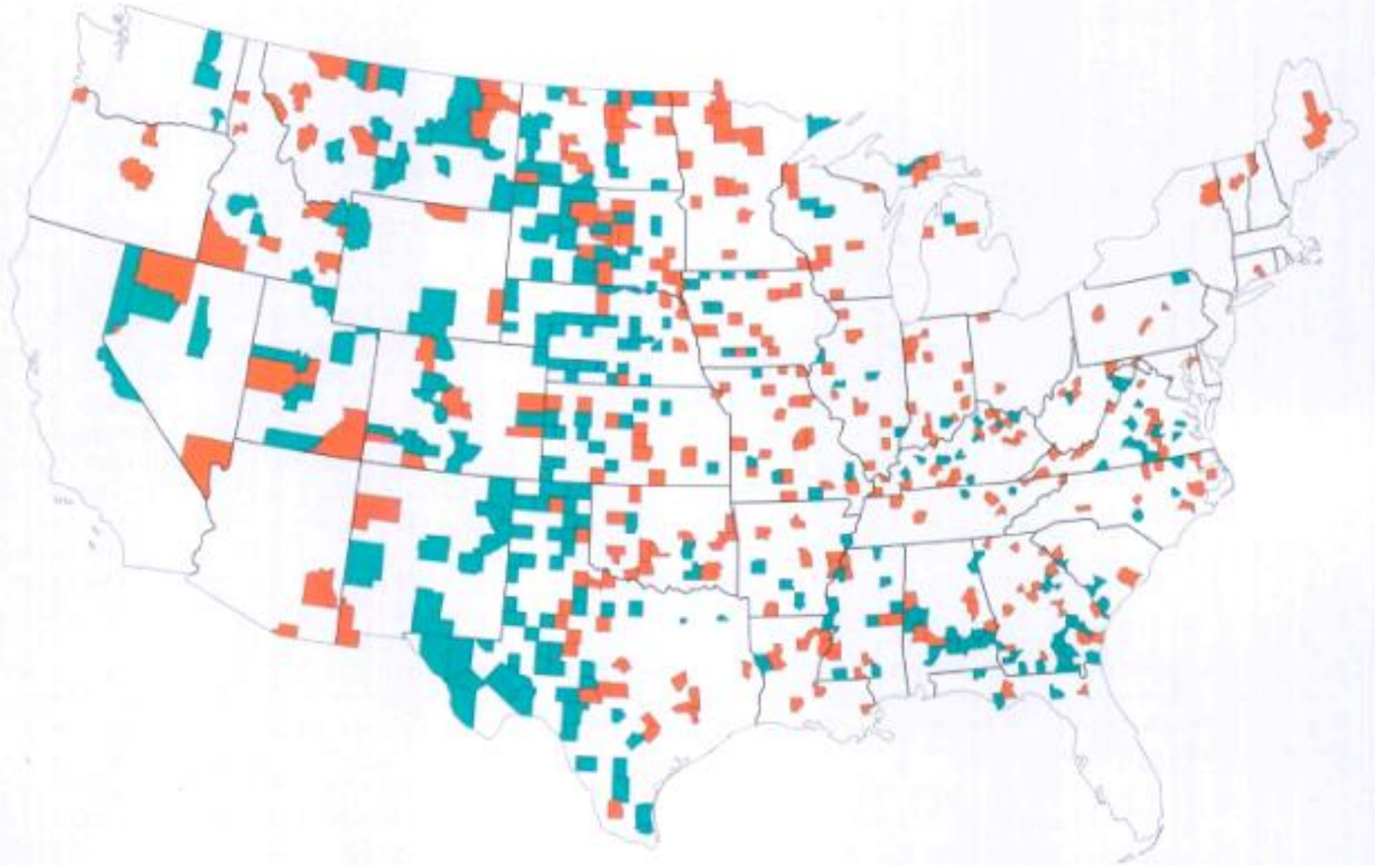
(Borehole, Grant, Yellowstone Park, WY)



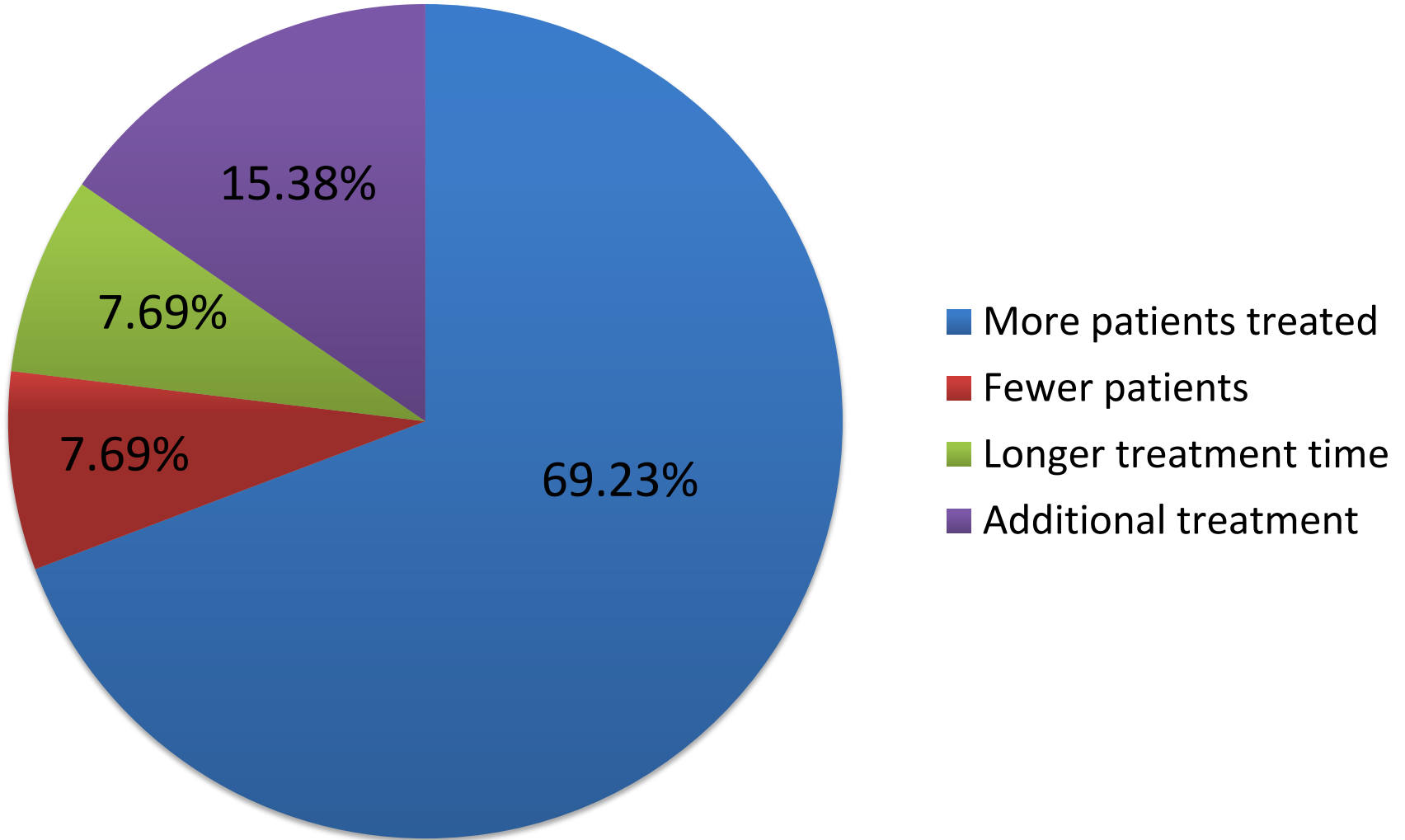
Reproducibility vs. Reanalysis

- Reproducibility requires the same experimental controls and materials
- Reanalysis is limited to the data generated by the initial experiments
 - Is it complete?
- Science = reproducibility
- Reanalysis = science??
 - Motivations are key to reanalysis





Reanalyses of 13 Null Results Medical Studies



PACE Trial Controversy

“I have not given up in my efforts to get the data to demonstrate that this trial did not show that psychotherapy extends the survival of cancer patients, but I am blocked by the unwillingness of authorities to enforce data sharing rules that they espouse.”

- James Coyne, PhD, PLoS Blog



Lee Fang

Nov. 27 2015, 3:54 p.m.

Coal CEO Thanks Lamar Smith, Asks Him to Expand Probe of Climate Scientists

In recent remarks Robert E. Murray, the chief executive officer of Murray Energy, the largest privately-held coal mining company in America, enthusiastically praised Rep. Lamar Smith, R-Tex., the chairman of the House Committee on Science, Space and Technology, for leading an investigation into prominent climate scientists and environmental officials

Murray then declared that the American Meteorology Association and Union of Concerned Scientists, two private nonprofits that serve the scientific community, also “need to be investigated.”

Murray then declared that the American Meteorology Association and Union of Concerned Scientists, two private nonprofits that serve the scientific community, also “need to be investigated.”

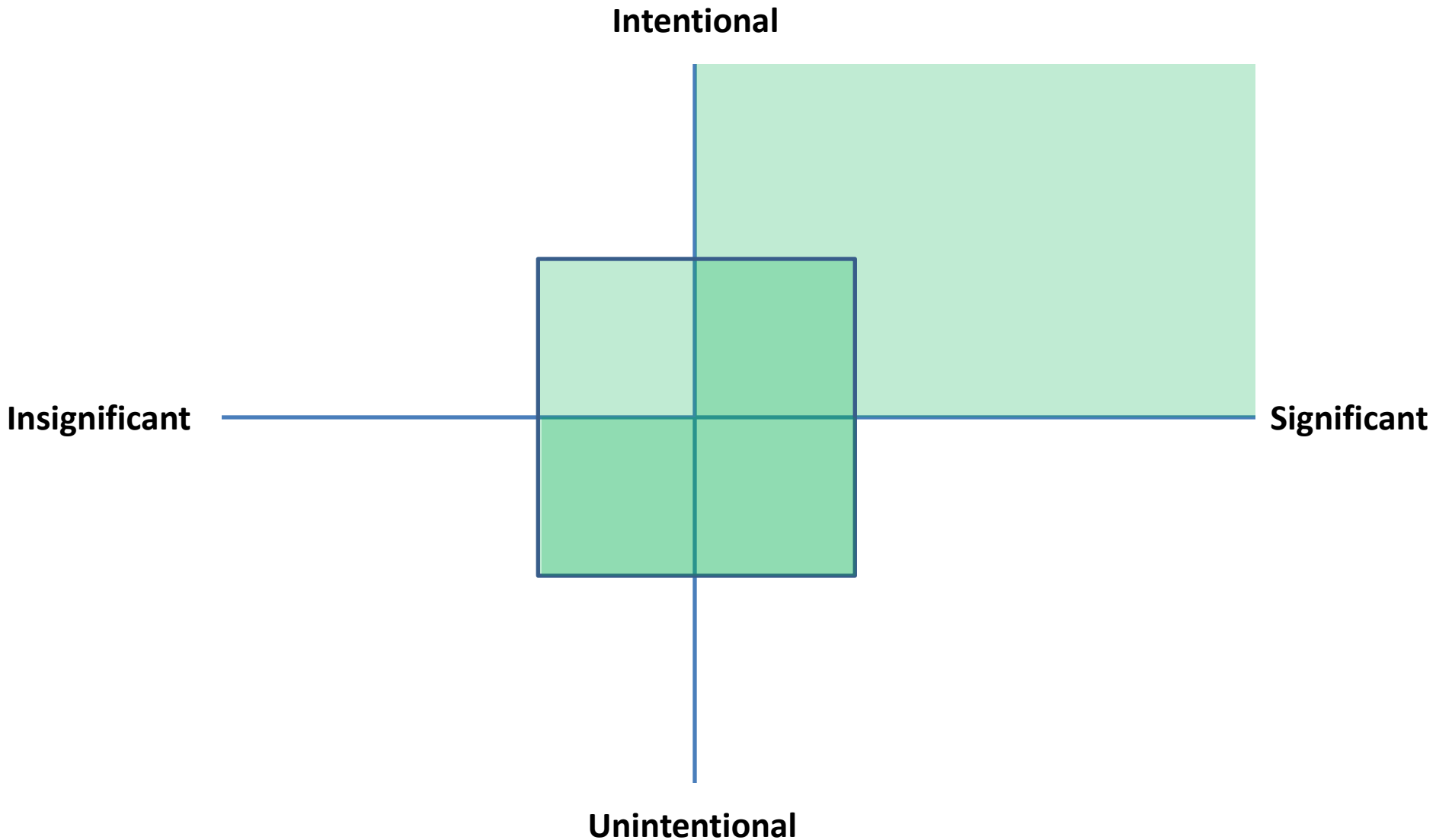
“They’re crony capitalists, they’re making a fortune off of you the taxpayer,” said Murray, who stood up to praise the Texas congressman again on the next day of the conference. After receiving the second round



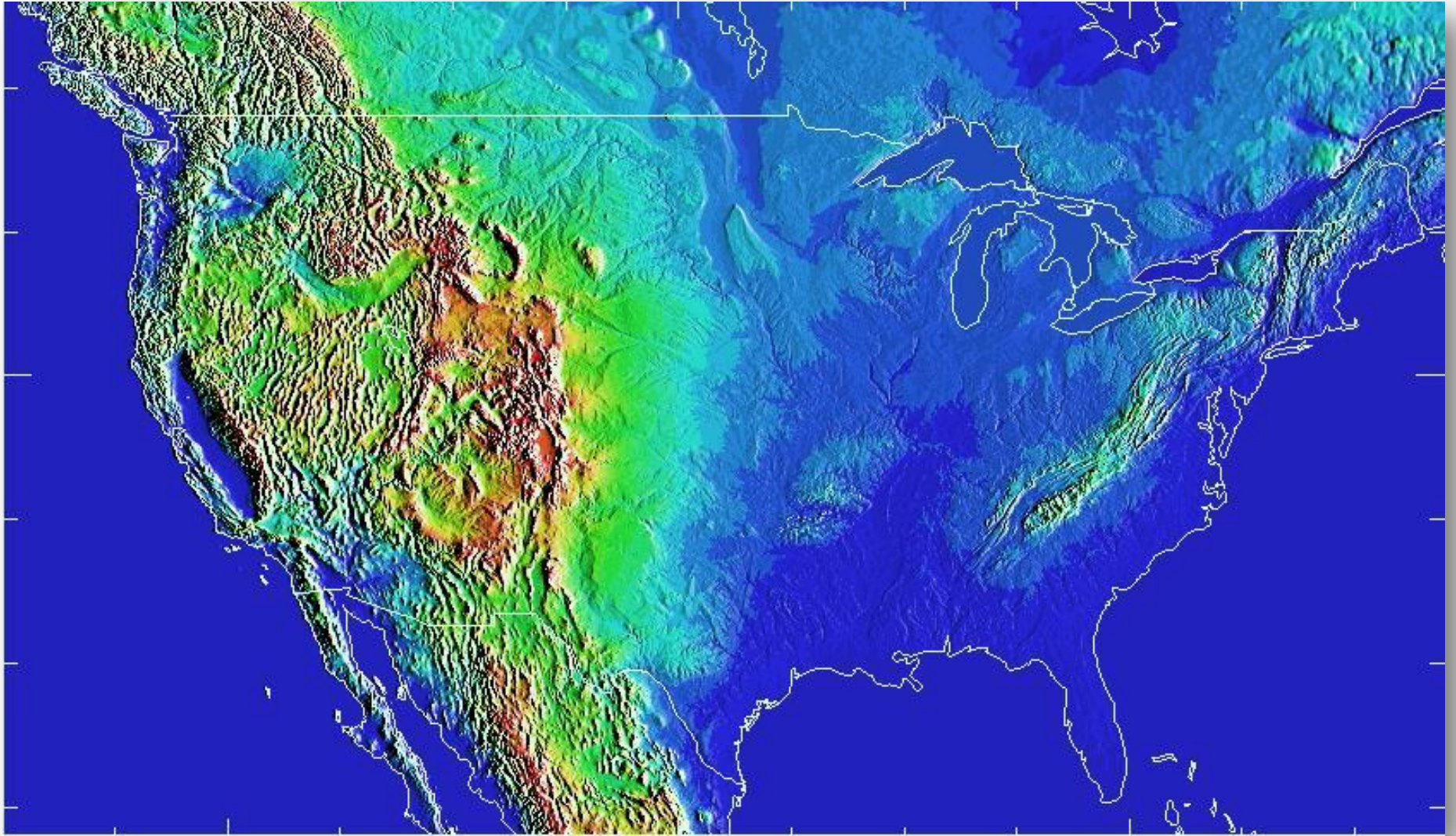
AAAS Leads Coalition to Protest Climate Science Inquiry

[FULL STORY](#)

Types of Inadequate Reporting

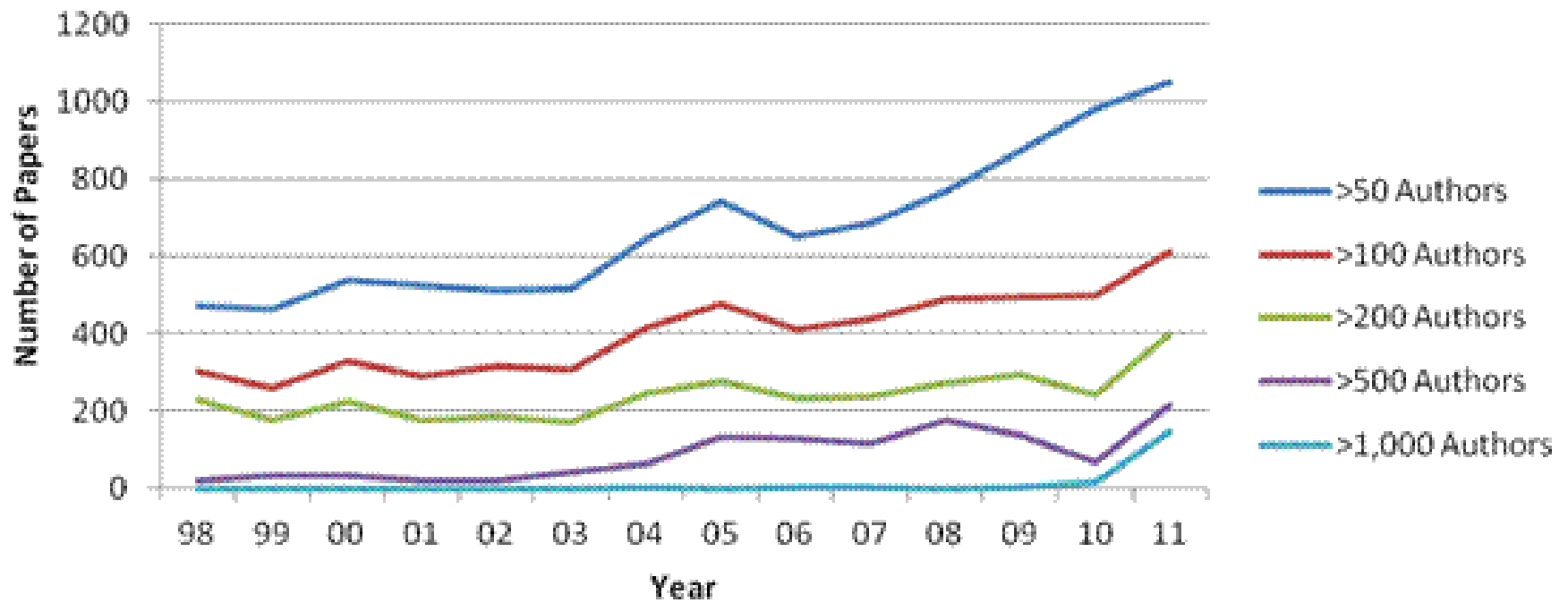








Multiauthor papers, 1998 to 2011





Dorothy Bishop

@deevybee

 Follow

When a manuscript with 20+ authors has grammatical errors, typos and/or no page numbers, you wonder how many authors actually read it.

RETWEETS

16

LIKES

39



8:15 a.m. - 21 Nov 2015



Deirdre Toher

@dtoher

 Follow


@deevybee could be too many making changes for proper version control to be maintained.

9:15 a.m. - 21 Nov 2015



Deirdre Toher

@dtoher

 Follow

@deevybee with that many authors it also means that people assume that the basics are "someone else's responsibility".

9:19 a.m. - 21 Nov 2015

Thank You

Kent R. Anderson
Publisher, *AAAS/Science*

