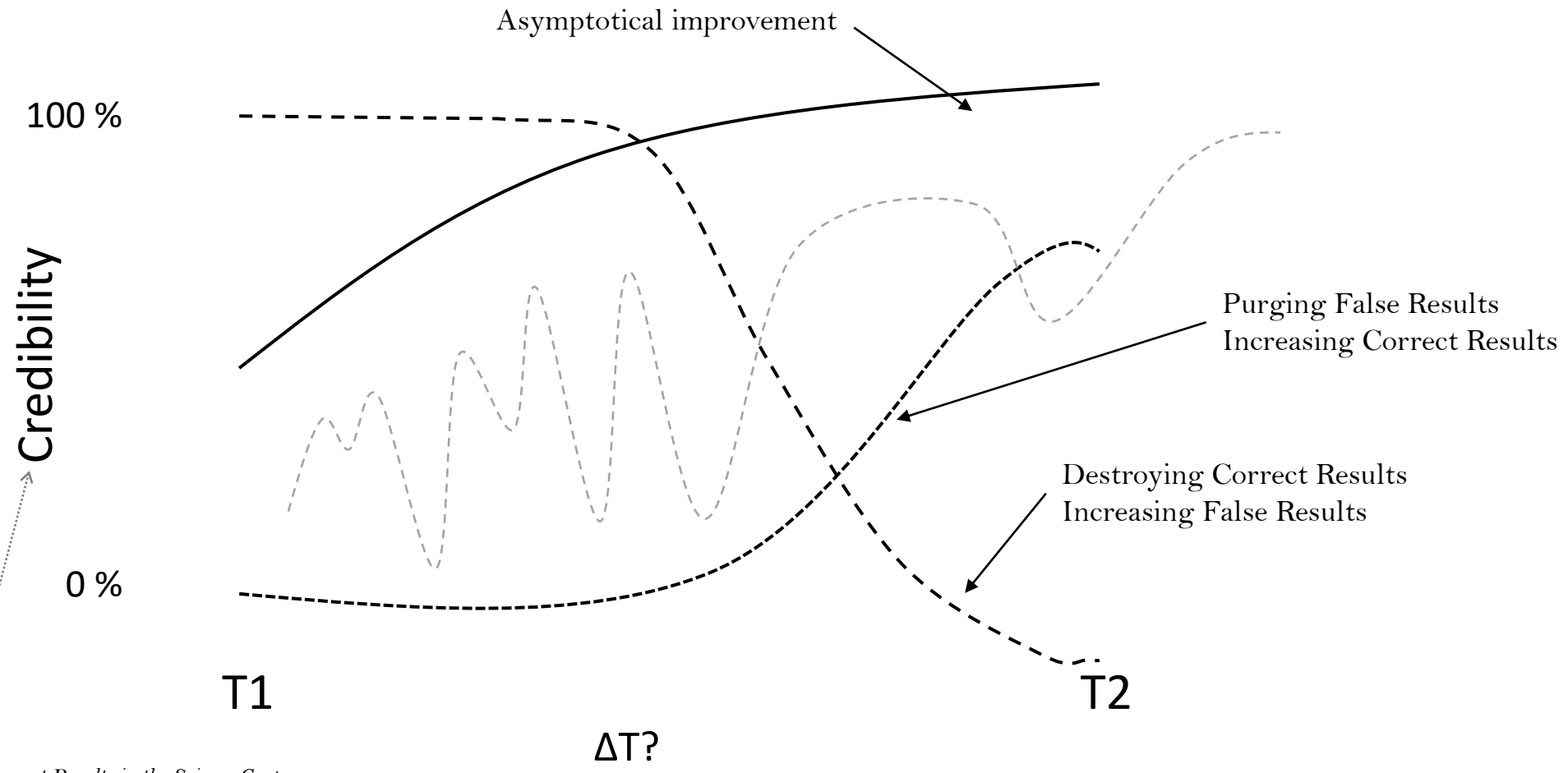
Why Science Is Not Necessarily Self-Correcting

Scientific Credibility

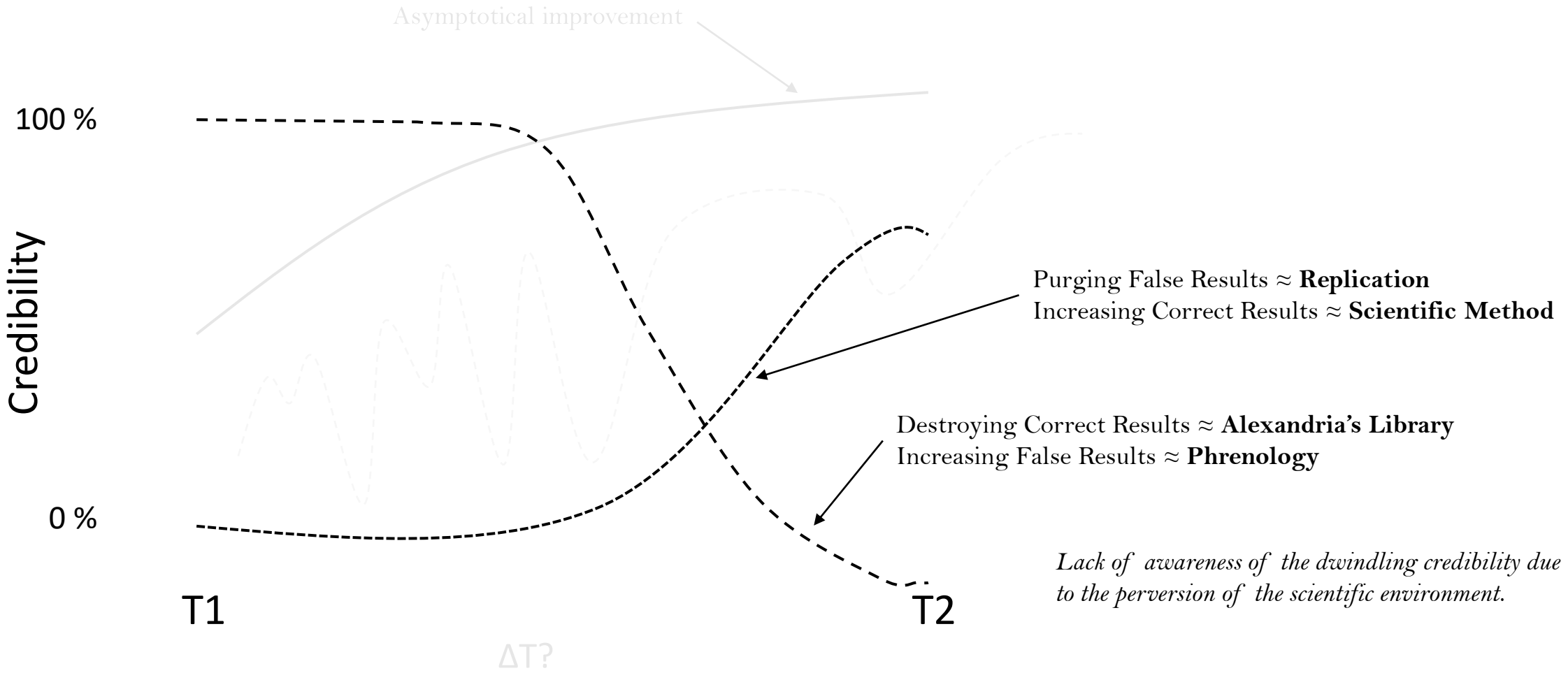
Sooner or later, if something is wrong, a replication effort will show it to be wrong and the scientific record will be corrected.



Heterogeneity
Influence across fields

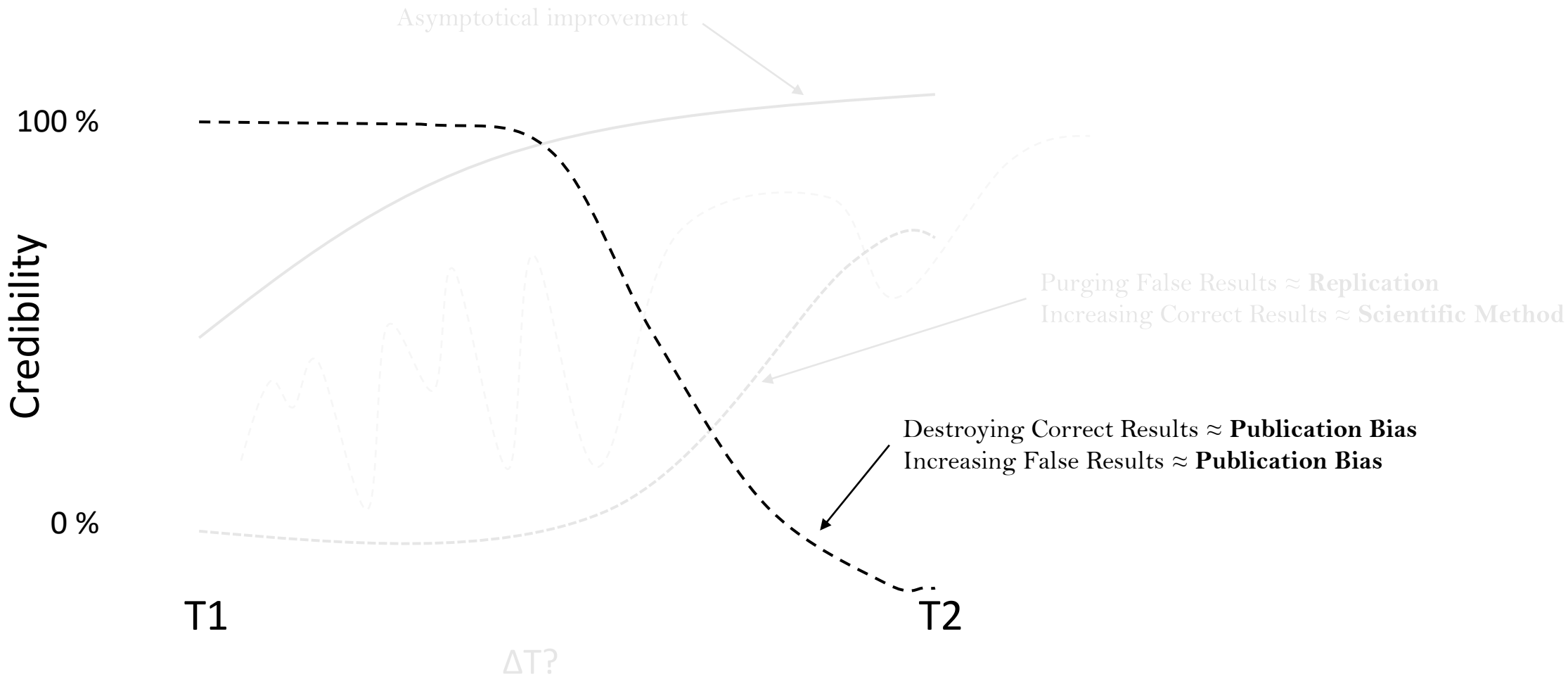
Scientific Credibility

> Sooner or later, if something is wrong, a replication effort will show it to be wrong and the scientific record will be corrected.



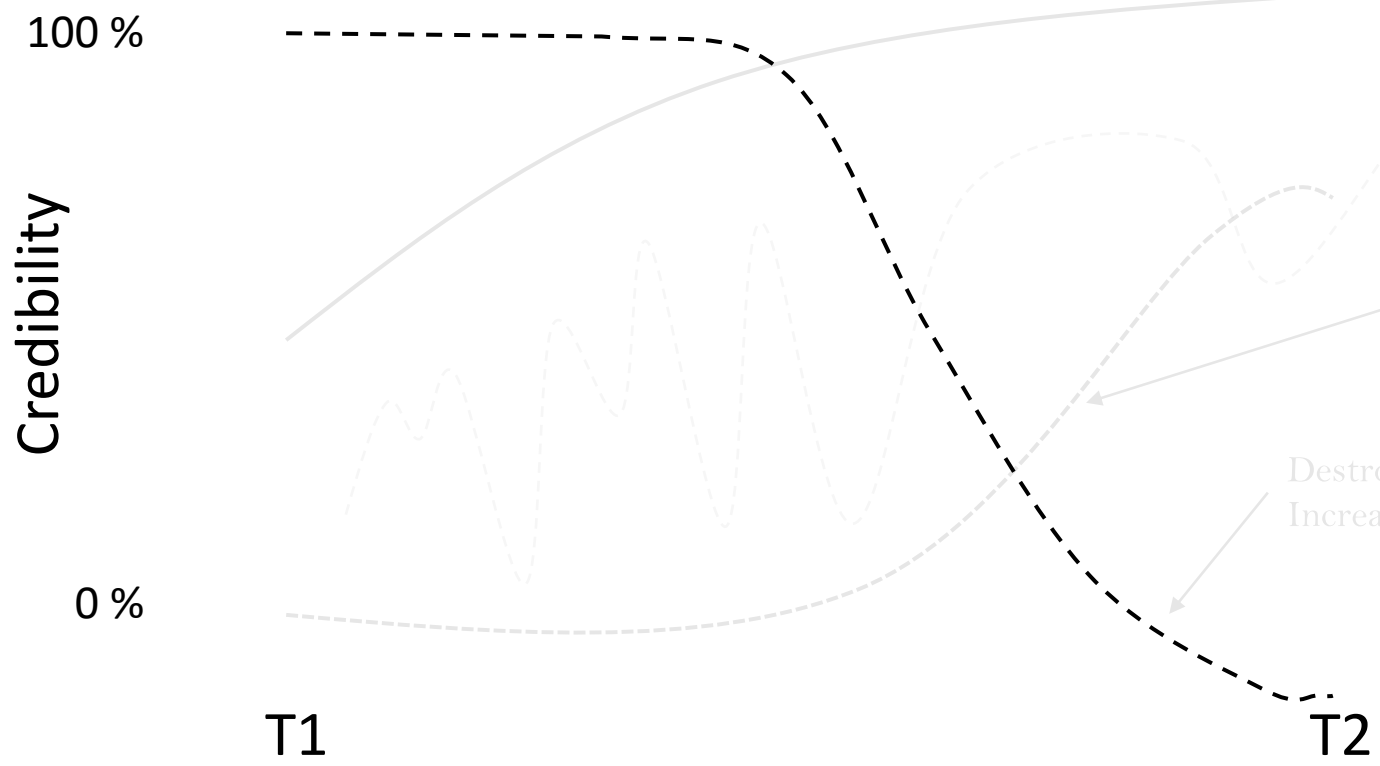
Scientific Credibility

> Sooner or later, if something is wrong, a replication effort will show it to be wrong and the scientific record will be corrected.



Scientific Credibility

> Sooner or later, if something is wrong, a replication effort will show it to be wrong and the scientific record will be corrected.



Raw data and Code is not available even destroyed
Low Reproducibility

Surprising effects and *Statistical Significance*

***Unlikely** effects* and ***Massaged** Statistics*

Low Replicability

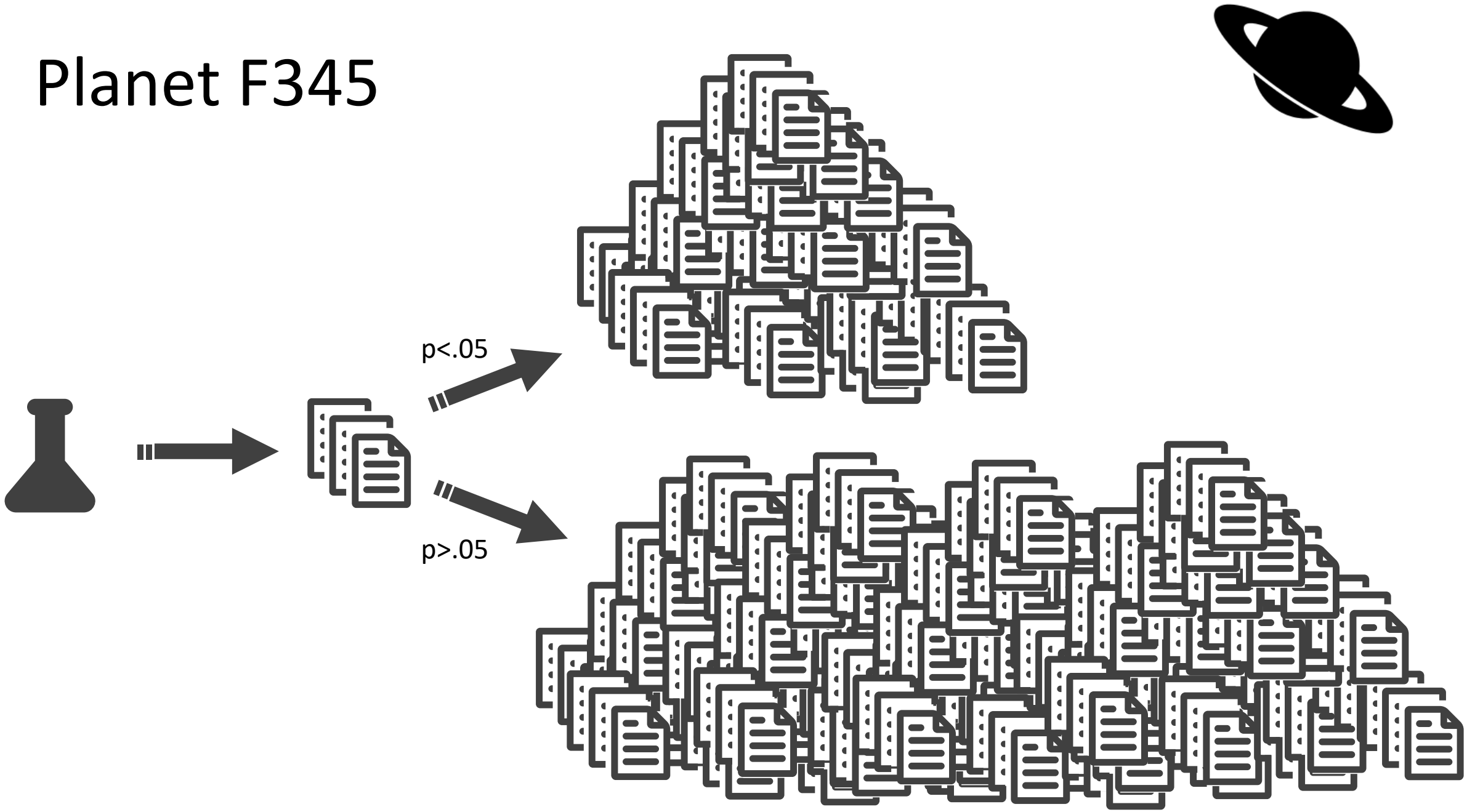
Asymptotical improvement

Purging False Results ≈ **Replication**
Increasing Correct Results ≈ **Scientific Method**

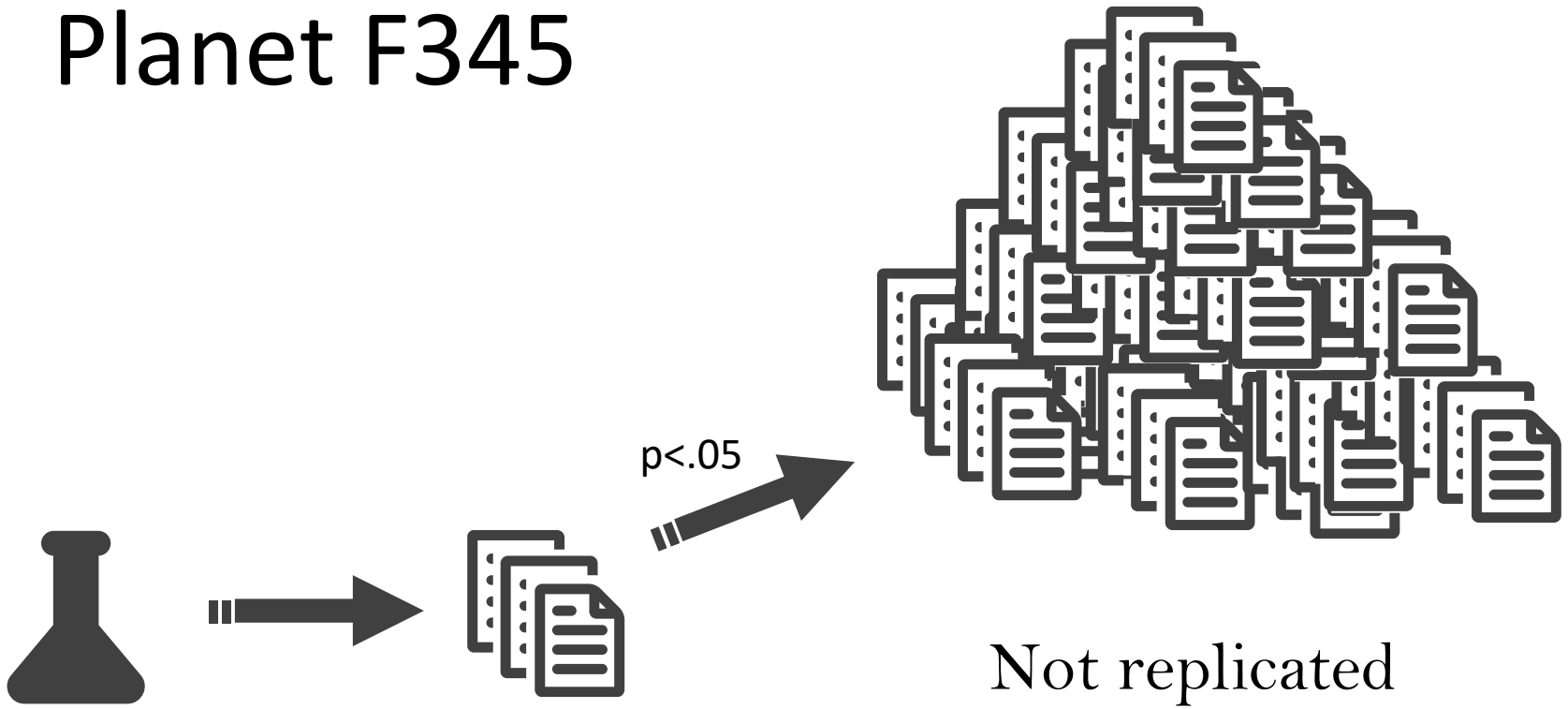
Destroying Correct Results ≈ **Publication Bias**
Increasing False Results ≈ **Publication Bias**

ΔT?

Planet F345



Planet F345



Not replicated

Low credibility

Even worse, we just don't know what is true and what is not!

A short stop on statistical power

A short stop on statistical power

But progress is made!

Maybe we are not doing science in the most efficient way

For sure we are not, inadequate power…

How many **correct** results are there in comparison with incorrect results?

For each real effect we have many incorrect ones.

Maybe we attribute big macro indexes to micro results

Leap of attribution

Poor return of investment

Maybe investing in medicine is not so good as inverting in Climate Change?

Discovery and replication

Researchers use strategies to maximize publications at the cost credibility

Replication '*was*' not rewarded and regarded as futile effort for '*idealess envious*' researchers

Replication is a fundamental piece for self-correction

Change the focus from '*finding*' something novel to discovering the truth

Many samples are better than a big sample when I am looking for a significant result, but not when I am looking for the truth… Questionable research practices.

The Rules of the Game Called Psychological Science
Bakker et al. 2012

Particularly if conceptual replications are *tuned* to replicate the result

> 'Each year, thousands of undergraduate projects are completed as part of the educational experience[…] they provide a good test of the replicability of established findings'

Harnessing the Undiscovered Resource of Student Research Projects
Grahe et al. 2012

Replication effort

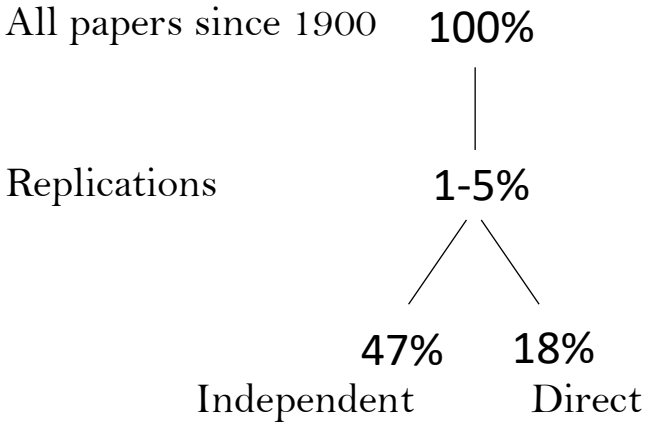


Table 1. Possibilities of Discovery and Replication: Six Possible Paradigms

	Replication results		
Discovery results	Correct	Wrong	Not obtained
Correct (true positive)	Optimal: ≤ 1%*	False nonreplication: <<1%*	Unconfirmed genuine discovery: 43%**
Wrong (false positive)	Self-correcting: ≤1%*	Perpetuated fallacy: 2%*	Unchallenged fallacy: 53%**

Does not include publication biases, according to Ioannidis (2005) it can go up to 95%

Same group replications replications ≈ confirmation bias *(p-hacking)*
Conceptual replications ≈ confirmation bias *(p-hacking)*
Allegiance bias ≈ followers of the theory

Impediments to self-correction

Table 2. A List of Described Impediments to Self-Correction in Science With Reference to Psychological Science

Impediments	Selected references
Publication bias	Ferguson & Brannick (2012); Shadish et al. (1989)
Other selective reporting bias (analysis and outcomes)	
Flexibility in data collection and analysis	Simmons et al. (2011); Wagenmakers et al. (2012);
Misreporting of results	Bakker & Wicherts (2011)
Voodoo correlations	Fiedler (2011)
Fabricated results	Fanneli (2009)
Other questionable research practices	John et al. (2012)
Excess significance bias (may reflect any of the above)	Francis (2012b); Ioannidis & Trikalinos (2007)
Underpowered studies	Maxwell (2004)
No replication work done—especially direct replication by independent investigators	Makel et al. (2012)
Underestimation of the replication crisis	Pashler & Harris (2012)
Editorial bias against replication research	Neuliep & Crandall (1990)
Reviewer bias against replication research	Neuliep & Crandall (1993)
Data, analyses, protocols not publicly available	Alsheikh-Ali et al. (2011); Wicherts, Borsboom, Kats, & Molenaar(2006)

> Classic publication bias considers that there are specific well-delineated studies with clear protocols, data, and analyses that disappear completely in a file drawer. In psychological science, as well as in other scientific fields, a study may be poorly defined and no protocol may exist. Investigators may continue adding and melding data, analyses, and subanalyses until something significant and publishable emerges.

Impediments and problems

Proper replication is insufficient in the presence of publication bias

Lack of open data, code and protocols

Scarce multicenter studies with many participants

Undervalued by community

Individual data is very valuable for meta analysis

Incentives for replication and correcting wrong results

Revalue replication effort

Stop focusing on impact factor and number of publications

Crowdsourcing

Preregistration

Massive replication by students

Problems with incentives

Top to bottom involvement

'Surprising' results will happen less often

Replications are a game played by students

Strictly follow registration/checklist and hide somewhere else

Open data may promote data dredging

Junk papers if everything is published with no review

However, at the end of the day, no matter what changes are made, scientific credibility may not improve unless the pursuit of truth remains our main goal in our work as scientists. This is a most noble mission that needs to be continuously reasserted.