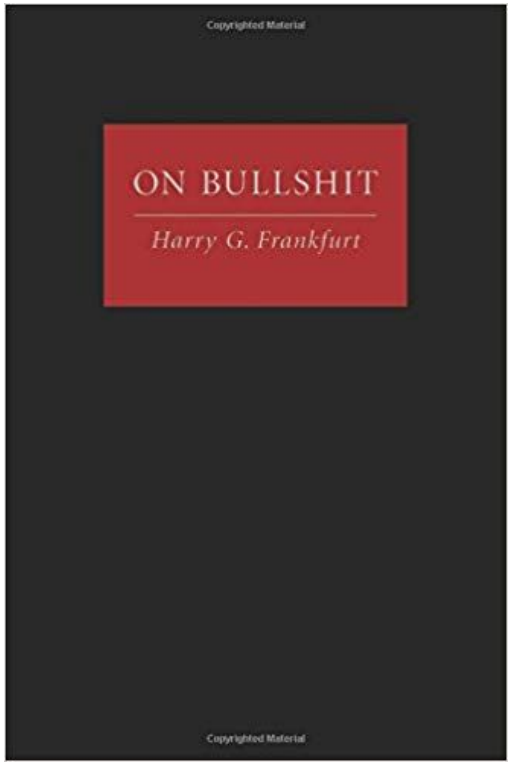
Reproducibility Fallacies

towards a better kind of scientific reproducibility

A followup on “Spin vs Bullshit”



Shashi had a slide on “**spin**” - in reality there is also bullshit to contend with.

On Bullshit (2005). A bestseller on moral philosophy.

- A **liar** knows what the truth is. They lie.
- A **spinner** knows what the truth is - they spin it their way
- A **bullshitter** does not care of what the truth is. They just want to benefit themselves.

Bullshit is more dangerous to the truth.

Learn to recognize bullshit - an important skill!

Theranos - 2018

Pure bullshit - how come they were able to raise $700 million?

Read more about it, understand their story -
you will understand and recognize bullshit.

How to tell bullshit?

They never enable you how to do something. Always lacks specifity. It is always about making you “**feel**” a certain way.

Ask yourself the following. Having listened to what someone states:

- What can I do that I was not able to do before?
- What is it that I do differently after reading/studying this?
- Look for simple, actionable information.

A reproducibility case study

2015 vs 2019

Contrast the reproducibility of two publications

Both are scientifically correct

1. A well-meaning non technical author driven paper:
   - Genomic surveillance elucidates Ebola virus origin and transmission during the 2014 outbreak

2. A paper by the “best” bioinformaticians on the planet:
   - A synthetic-diploid benchmark for accurate variant-calling evaluation

Heng Li

From Wikipedia, the free encyclopedia

Heng Li is a Chinese bioinformatics scientist. He is an Assistant Professor at the department of Biomedical Informatics of Harvard Medical School and the department of Biostatistics & Computational Biology of Dana-Farber Cancer Institute.[3][4][5] He was previously a research scientist working at the Broad Institute in Cambridge, Massachusetts with David Reich and David Altshuler.[6] Li's work has made several important contributions in the field of next generation sequencing.

Daniel MacArthur

Daniel is a group leader within the Analytic and Translational Genetics Unit (ATGU) at Massachusetts General Hospital. He is also Assistant Professor at Harvard Medical School, and the Co-Director of Medical and Population Genetics at the Broad Institute of Harvard and MIT.

@dgmacarthur

macarthurlab.org

Ebola paper - 2015

Assembly of full-length EBOV genomes

EBOV reads were extracted from the demultiplexed Fastq files using Lastal against a custom-made database containing all full-length EBOV genomes. The reads were then *de novo* assembled using Trinity and contigs were oriented, merged and cleaned using a custom-made pipeline. Contigs were indexed and all sequencing reads from each individual sample were aligned back to its own EBOV consensus sequence using Novoalign v3 with the following parameters: -k -l 40 -g 40 -x 20 -t 160. Duplicates were removed using Picard v1.4 and alignment files were realigned using GATK v2. Consensus sequences were called from the EBOV-aligned reads using GATK v2. All generated genomes were annotated as well as manually inspected for accuracy, such as the presence of intact ORFs, using Geneious v7. Regions where depth of coverage was less < 3x were called as 'N'. Eight patients in our data set had sequences for multiple time points of collection. There were no differences in their consensus assemblies across time. Therefore only one consensus sequence per patient was reported.

SYN-DIP paper - 2018

```
# Download and install evaluation suite (Linux only)
curl -L https://github.com/lh3/CHM-eval/releases/download/v0.4/CHM-evalkit-20180221.tar \
    | tar xf -
# Call CHM1-CHM13 variants in the GRCh37 coordinate (will take a while...)
wget -qO- ftp://ftp.sra.ebi.ac.uk/vol1/ERA596/ERA596361/bam/CHM1_CHM13_2.bam \
    | freebayes -f hs37.fa - > CHM1_CHM13_2.raw.vcf
# Filter (use your own filters if you like)
CHM-eval.kit/run-flt -o CHM1_CHM13_2.flt CHM1_CHM13_2.raw.vcf
# Distance-based evaluation
CHM-eval.kit/run-eval -g 37 CHM1_CHM13_2.flt.vcf.gz | sh
more CHM1_CHM13_2.flt.summary
# Evaluating allele and genotype accuracy (Java required)
CHM-eval.kit/rtg format -o hs37.sdf hs37.fa   # if you haven't done this before
CHM-eval.kit/run-eval -g 37 -s hs37.sdf CHM1_CHM13_2.flt.vcf.gz | sh
more CHM1_CHM13_2.flt.rtg.summary
```

Evaluation

The paper written by expert bioinformaticians today is “less” reproducible than that written by less experienced researchers four years ago.

WHY?

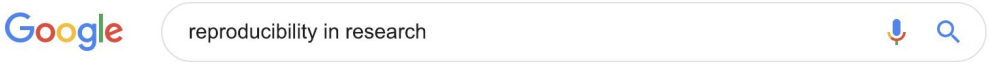
The unexpected challenges of defining reproducibility

Even champions of reproducibility routinely mix different definitions,

use the wrong words and incorrect terminology

with that lead us down the wrong way

Suggestions already start to pop up before you even begin searching



About 14,000,000 results (0.51 seconds)

Scholarly articles for **reproducibility in research**
… to improve the validity and **reproducibility** of **research** … - Springate - Cited by 79
Reproducibility of **research** and preclinical validation: … - Pusztai - Cited by 44
Is there a **reproducibility** crisis? A Nature survey lifts the … - Baker - Cited by 205

What does research reproducibility mean? | Science Translational ...
stm.sciencemag.org/content/8/341/341ps12
by SN Goodman - 2016 - Cited by 240 - Related articles
Jun 1, 2016 - Abstract. The language and conceptual framework of "**research reproducibility**" are nonstandard and unsettled across the sciences. In this ...

Challenges in irreproducible research - Nature
https://www.nature.com/collections/prbfkwmwvz
Oct 18, 2018 - Science moves forward by corroboration – when **researchers** verify others' results. ... Explanations include increased levels of scrutiny, complexity of experiments and statistics, and pressures on **researchers**. Journals, scientists, institutions and funders all have a part in tackling **reproducibility**.

Reproducibility - Wikipedia
https://en.wikipedia.org/wiki/Reproducibility
Jump to **Reproducible research** - The term **reproducible research** refers to the idea that the ultimate product of academic **research** is the paper along with the laboratory notebooks and full computational environment used to produce the results in the paper such as the code, data, etc. that can be used to reproduce the results and create new work based ...
About · History · Reproducible data · Noteworthy irreproducible ...

Study raises new questions about reproducibility of research
https://www.insidehighered.com/.../study-raises-new-questions-about-reproducibility-r...
Aug 30, 2018 - Academics are easily able to predict whether an experiment's findings will be **reproducible**, according to a **study** that raises more questions ...

SHARE

PERSPECTIVE SCIENTIFIC INTEGRITY

What does research reproducibility mean?

INTRODUCTION

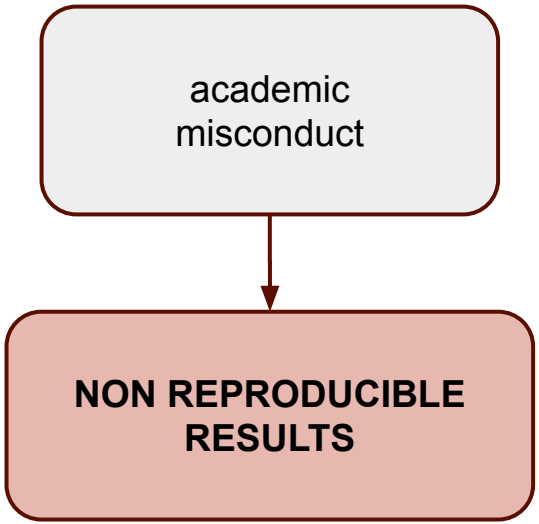
Concern about the reproducibility of scientific research has been steadily rising recently with reports that the results of experiments in numerous domains of science could not be replicated (*1*, *2*). Whereas problems in

“scientific integrity”

is being equated with

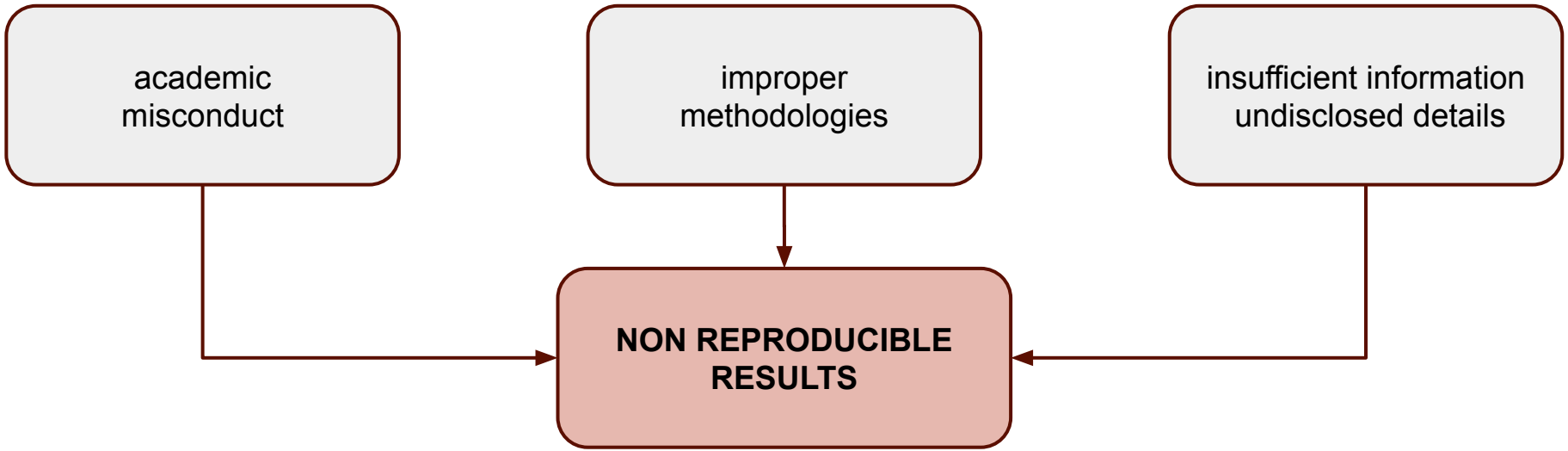
“scientific reproducibility”

What are the causes for non-reproducibility?



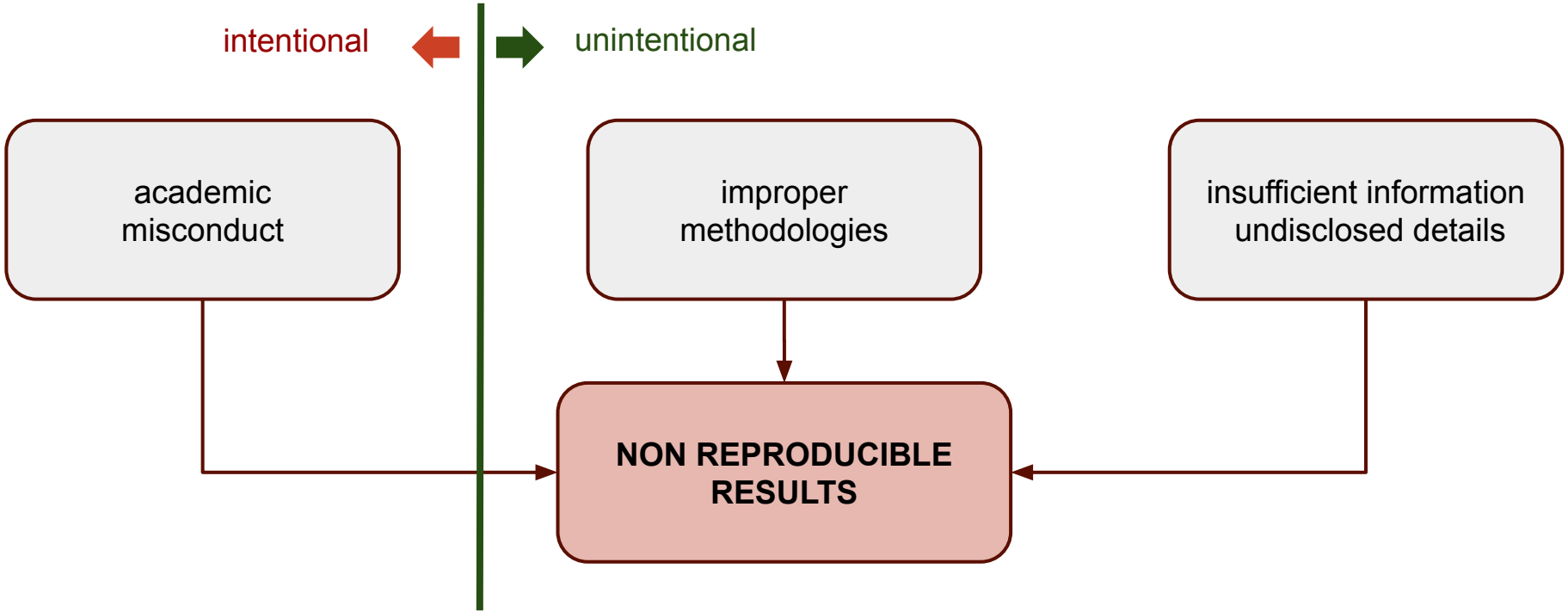
What are the causes for non-reproducibility?

What are the causes for non-reproducibility?



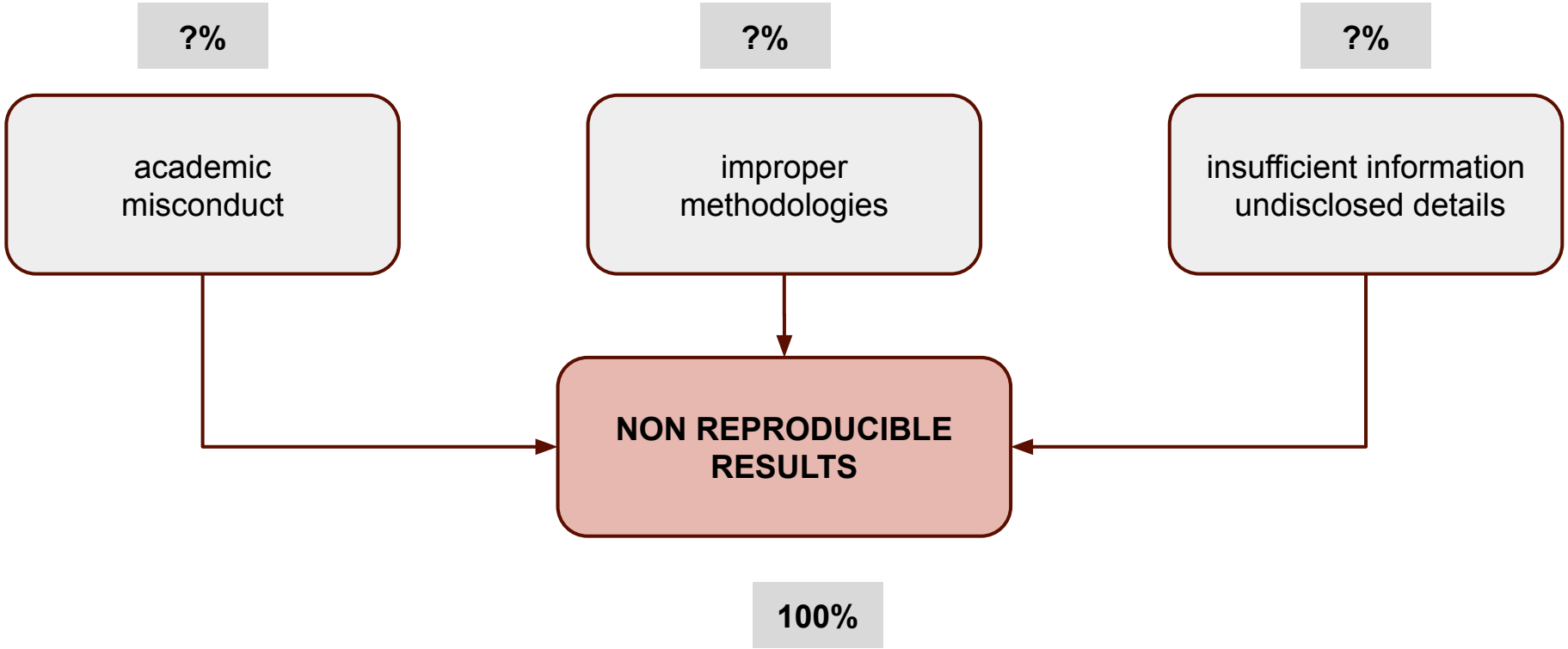
These are radically different problems with radically different solutions

What are the causes for non-reproducibility?

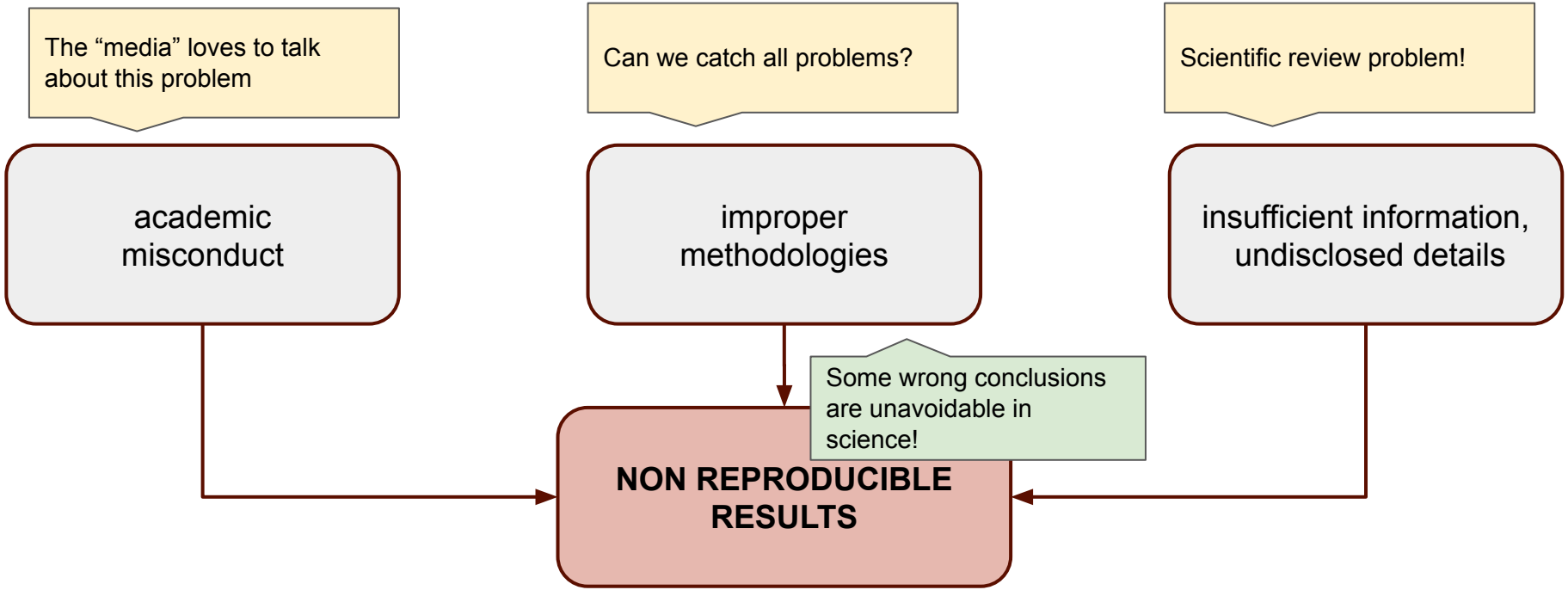


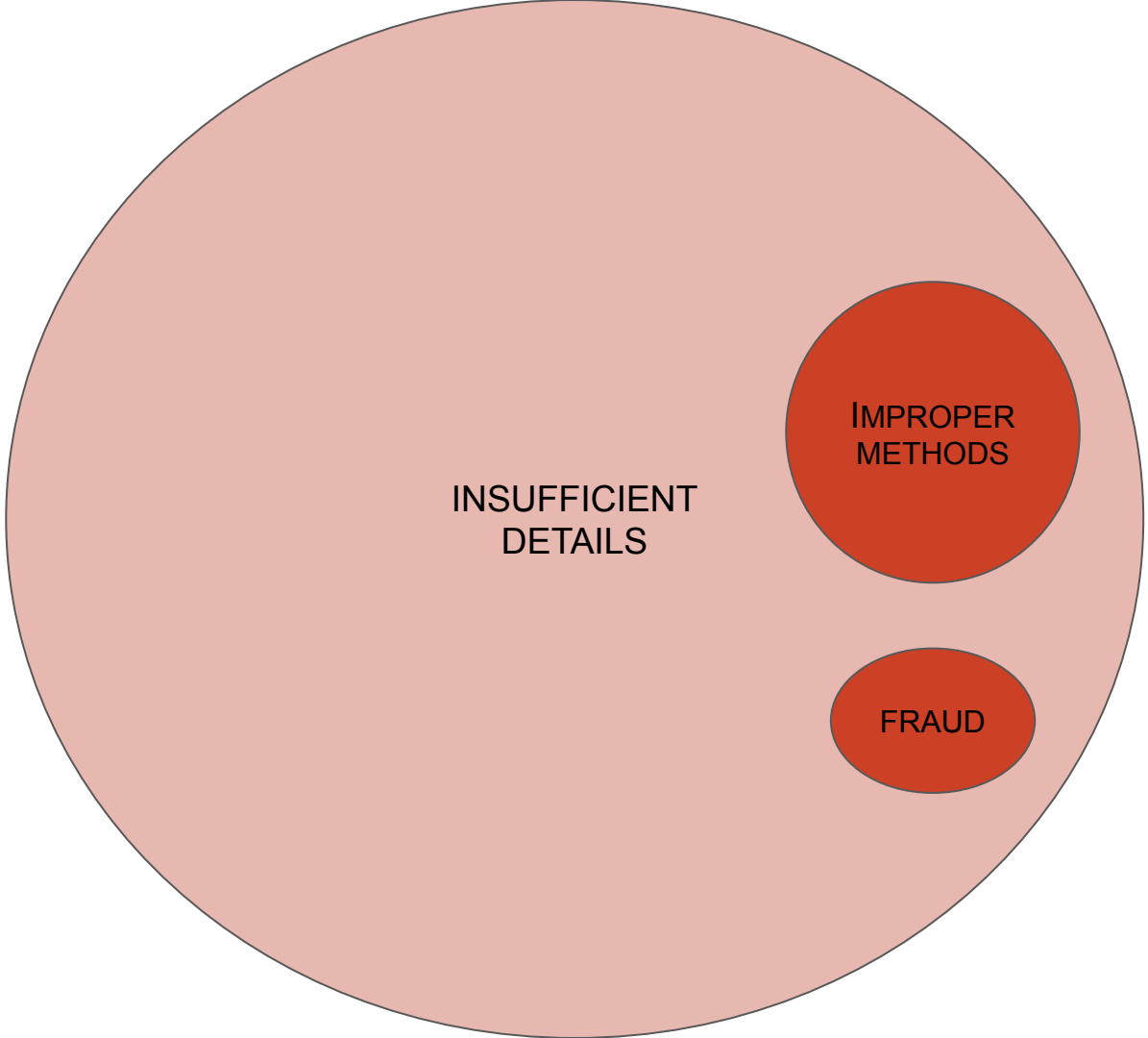
These are radically different problems with radically different solutions

What are the causes for non-reproducibility?



What are the causes for non-reproducibility?




INSUFFICIENT
DETAILS
IMPROPER
METHODS
FRAUD

Reproducibility Fallacies and how to avoid them

Fallacy 1

Reproducibility is associated with scientific integrity.

Fallacy 1

Reproducibility is associated with scientific integrity.

NO!

The most counter-productive definition

Irreproducible Results -> cheating

Reproducible Results -> not cheating

Fallacy 2

There is **reproducible** and there is **irreproducible** research.

Fallacy 2

There is **reproducible** and there is **irreproducible** research.

NO!

Reproducibility is not a binary quality.

Yes or **No**

Reproducibility is a scale - how well do we understand what took place?

Reproducibility is a scale

Reproducibility starts with quantifying the effort it takes to **understand** how to "reproduce" results

imagine that the scale is in years

Low Effort
Reproducible Research

Unreasonably high effort
Irreproducible Results

Fallacy 3

Irreproducible results means bad science!

Fallacy 3

Irreproducible results means bad science!

NO!

Some irreproducibility is part of the science.

Reproducibility is not a statement on the “correctness” of a results.

My definition of reproducibility

The primary purpose of “reproducibility” is to educate one another about the decision making that went into a discovery.

Your work is reproducible if you are able to teach someone else how to do it.

We could catch 95% of reproducibility problems if we evaluated the “educational” potential of the paper.

MENU nature International journal of science

WORLD VIEW · 24 MAY 2018

Before reproducibility must come preproducibility

Instead of arguing about whether results hold up, let's push to provide enough information for others to repeat the experiments, says Philip Stark.

Philip B. Stark

Please no! We don't need a new word!

We have a perfectly good word for it already. **REPRODUCIBILITY**

Reproducibility *in my opinion* is the “educational” component of the research publication.

If you can teach someone
how to perform the same actions
you have created reproducible results