Bad science, and how to avoid it

Reproducibility crisis

* http://dx.doi.org/10.1371/journal.pmed.0020124 *
http://matthew-brett.github.io/teaching/ioannidis_2005.html

Begley and Ellis 2012

Scientists at Amgen (a drug company) tried to reproduce findings from 53 "landmark" studies.

> *. . . when findings could not be reproduced, an attempt was made to contact the original authors, discuss the discrepant findings, exchange reagents and repeat experiments under the authors' direction, occasionally even in the laboratory of the original investigator.*

Of 53 studies, only 6 replicated (11%).

Glenn Begley and Lee Ellis (2012) "Raise standards for preclinical cancer research" Nature 483 pp. 531–533

Cargo cult science

> *In the South Seas there is a cargo cult of people. During the war they saw airplanes land with lots of good materials, and they want the same thing to happen now. So they've arranged to imitate things like runways, to put fires along the sides of the runways, to make a wooden hut for a man to sit in, with two wooden pieces on his head like headphones and bars of bamboo sticking out like antennas —— he's the controller – and they wait for the airplanes to land. They're doing everything right. The form is perfect [. . .]. But it doesn't work.*

Richard Feynman (1974) Cargo Cult Science.

The SCHAT principles

- Sceptical,
- Careful,
- Honest And
- Transparent.

Sceptical

(by kladcat under CC BY 2.0, via Wikimedia Commons)

Sceptical

> *Science alone of all the subjects contains within itself the lesson of the danger of belief in the infallibility of the greatest teachers in the preceding generation. . . Learn from science that you must doubt the experts . . . Science is the belief in the ignorance of experts*

Richard Feynman, What is Science? (1969)

Careful

> *The scientific method's central motivation is the ubiquity of error - the awareness that mistakes and self-delusion can creep in absolutely anywhere and that the scientist's effort is primarily expended in recognizing and rooting out error."*

Donoho, David L, et al. 2009. Reproducible research in computational harmonic analysis. *Computing in Science & Engineering* 11, 8–18.

Careful

> *The first principle is that you must not fool yourself—and you are the easiest person to fool. So you have to be very careful about that. After you've not fooled yourself, it's easy not to fool other scientists. You just have to be honest in a conventional way after that.*

Richard Feynman, Cargo Cult Science (1974)

Honest

> *It's a kind of scientific integrity, a principle of scientific thought that corresponds to a kind of utter honesty — a kind of leaning over backwards.*

Richard Feynman, Cargo Cult Science (1974)

Transparent

For example, if you're doing an experiment, you should report everything that you think might make it invalid–not only what you think is right about it: other causes that could possibly explain your results; and things you thought of that you've eliminated by some other experiment, and how they worked — to make sure the other fellow can tell they have been eliminated . . . In summary, the idea is to try to give all of the information to help others to judge the value of your contribution; not just the information that leads to judgment in one particular direction or another.

Richard Feynman, Cargo Cult Science (1974)

The difference between true and false

> *In studies for which findings could be reproduced, authors had paid close attention to controls, reagents, investigator bias and describing the complete data set. For results that could not be reproduced, however, data were not routinely analysed by investigators blinded to the experimental versus control groups. Investigators frequently presented the results of one experiment, such as a single Western-blot analysis. They sometimes said they presented specific experiments that supported their underlying hypothesis, but that were not reflective of the entire data set.*

Begley and Ellis (2012).

Wrong tools, more errors

	B	C	I	J	K	L	M
2			Real GDP growth				
3			Debt/GDP				
4	Country	Coverage	30 or less	30 to 60	60 to 90	90 or above	30 or less
26			3.7	3.0	3.5	1.7	5.5
27	Minimum		1.6	0.3	1.3	-1.8	0.8
28	Maximum		5.4	4.9	10.2	3.6	13.3
29							
30	US	1946-2009	n.a.	3.4	3.3	-2.0	n.a.
31	UK	1946-2009	n.a.	2.4	2.5	2.4	n.a.
32	Sweden	1946-2009	3.6	2.9	2.7	n.a.	6.3
33	Spain	1946-2009	1.5	3.4	4.2	n.a.	9.9
34	Portugal	1952-2009	4.8	2.5	0.3	n.a.	7.9
35	New Zealand	1948-2009	2.5	2.9	3.9	-7.9	2.6
36	Netherlands	1956-2009	4.1	2.7	1.1	n.a.	6.4
37	Norway	1947-2009	3.4	5.1	n.a.	n.a.	5.4
38	Japan	1946-2009	7.0	4.0	1.0	0.7	7.0
39	Italy	1951-2009	5.4	2.1	1.8	1.0	5.6
40	Ireland	1948-2009	4.4	4.5	4.0	2.4	2.9
41	Greece	1970-2009	4.0	0.3	2.7	2.9	13.3
42	Germany	1946-2009	3.9	0.9	n.a.	n.a.	3.2
43	France	1949-2009	4.9	2.7	3.0	n.a.	5.2
44	Finland	1946-2009	3.8	2.4	5.5	n.a.	7.0
45	Denmark	1950-2009	3.5	1.7	2.4	n.a.	5.6
46	Canada	1951-2009	1.9	3.6	4.1	n.a.	2.2
47	Belgium	1947-2009	n.a.	4.2	3.1	2.6	n.a.
48	Austria	1948-2009	5.2	3.3	-3.8	n.a.	5.7
49	Australia	1951-2009	3.2	4.9	4.0	n.a.	5.9
50							
51			4.1	2.8	2.8	=AVERAGE(L30:L44)	

Reinhart & Rogoff (2010) American Economic Review. 100: 573–78.

Herndon, Ash & Pollin (2014). Cambridge Journal of Economics. 38: 257–279

Picketty (2013) Capital in the Twenty-First Century.

Understanding by building

"What I cannot create, I do not understand"

Richard Feynman - found on his blackboard after his death.

Computers are confusing

- if you can, work with someone else;
- always get someone else to check what you did;
- force yourself to get up and walk away;
- always have a pencil and paper next to you.

The data analysis landscape

- Python
- R

Plan from here

- Introduction to the Jupyter Notebook;
- Some questions from some data;
- Thinking about chance;
- You do the logic, the computer does the work.