

CERN Colloquium

Crisis? Surely you must be joking

Andrea Saltelli Centre for the Study of the Sciences and the Humanities, University of Bergen, and Open Evidence Research, Open University of Catalonia

Thursday 7 Jun 2018, 16:30 → 17:30 Main Auditorium (CERN)

Where to find this talk: www.andreasaltelli.eu



Crisis in statistics?

Statistics is experiencing a quality control crisis



Effect or no effect?





As debate rumbles on about how and how much poor statistics is to blame for poor reproducibility, Nature asked influential statisticians to recommend one change to improve science. The common theme? The problem is not our maths, but ourselves. CORRESPONDENCE · 16 JANUARY 2018



Fixing statistics is more than a technical issue

Andrea Saltelli 🏧 & Philip Stark

https://www.nature.com/articles/d41586-018-00647-9

CORRESPONDENCE · 16 JANUARY 2018



Integrity must underpin quality of statistics



https://www.nature.com/articles/d41586-018-00648-8

The great paradox of science is that passionate practitioners must carefully produce dispassionate facts (J. Ravetz Scientific Knowledge and its Social Problems Oxford Univ. Press; ¹⁹⁷¹). Meticulous technical and normative judgement, as well as morals and morale, are necessary to navigate the forking paths of the statistical garden (Saltelli and Stark, 2018)

All users of statistical techniques, as well as those in other mathematical fields such as modelling and algorithms, need an effective societal **Commitment** to the maintenance of quality and integrity in their work. If imposed alone. technical or administrative solutions will only breed manipulation and evasion (Ravetz, 2018)

Crisis in science?

There have recently been alarms as to the scientific quality arrangement is several disciplines. The most visible symptom of this possible dysfunction is the so-called reproducibility crisis

The Economist

Economist.com

OCTOBER 197H-257H 2013

Washington's lawyer surplus How to do a nuclear deal with Iran Investment tips from Nobel economists Junk bonds are back The meaning of Sachin Tendulkar



On the radar: October 2013









WRONG.

Essay

Why Most Published Research Findings Are False

John P. A. Ioannidis



John P. A. Ioannides



J. P. A. Ioannidis, Why Most Published Research Findings Are False, PLoS Medicine, August 2005, 2(8), 696-701.

Failed replications, entire subfields going bad, fraudulent peer reviews, predatory publishers, perverse metrics, statistics on trial, …

… misleading science advice, institutions on denial, post-truth, a new breed of science wars ...

The crisis is methodological, epistemological, ethical and metaphysical

Scholars who saw it coming

• • •

and how they were vindicated

In 1963 Derek J. de Solla Price prophesized that Science would reach saturation (and in the worst case senility) under its own weight, victim of its own success and exponential growth (pp 1-32).





Derek J. de Solla Price

de Solla Price, D.J., 1963, Little science big science, Columbia University Press. ~2.2 million articles a year (2016) over ~30,000 journals

newsblog *Nature* brings you breaking news from the world of science

NEWS BLOG

Global scientific output doubles every nine years

07 May 2014 | 16:46 GMT | Posted by Richard Van Noorden | Category: Policy, Publishing

https://www.aje.com/en/arc/scholarly-publishing-trends-2016/

http://blogs.nature.com/news/2014/05/global-scientific-output-doublesevery-nine-years.html p.22: […] The problem of quality control in science is at the centre of the social problems of the industrialized science of the present period."

Ravetz, J., **1971**, Scientific Knowledge and its Social Problems, Oxford University Press.





Jerome R. Ravetz "If [science] fails to resolve this problem […] then the immediate consequences for morale and recruitment will be serious; and those for the survival of science itself, grave"

Ravetz, J., **1971**, Scientific Knowledge and its Social Problems, Oxford University Press.





Jerome R. Ravetz \cdots neoliberal ideologies lead to decreasing state funding of science, which becomes privatized \cdots knowledge as a monetized commodity replaces knowledge as a public good \rightarrow collapse of quality



Philip Mirowski

Mirowski, P. 2011. Science-Mart: Privatizing American Science, Harvard University Press.



p. 179. For it is possible for a field to be diseased […] reforming a diseased field is a task of great delicacy […] not even an apparatus of institutional structures, can do anything to maintain or restore the health of a field in the absence of an essential ethical element operating through the interpersonal channel of communication.

Ravetz, J., **1971**, Scientific Knowledge and its Social Problems, Oxford University Press.





Jerome R. Ravetz



Brow



The Economic Journal, 127 (October), F236-F265. Doi: 10.1111/ecoj.12461 © 2017 Royal Economic Society. Published by John Wiley & Sons, 9600 Garsington Road, Oxford OX4 2DQ, UK and 350 Main Street, Malden, MA 02148, USA.

GOPEN ACCESS June 21, 2017

Why Most Clinical Research Is Not Useful

John P. A. Ioannidis 🖂

ESSAY

Published: June 21, 2016 • https://doi.org/10.1371/journal.pmed.1002049

THE POWER OF BIAS IN ECONOMICS RESEARCH*

John P. A. Ioannidis, T. D. Stanley and Hristos Doucouliagos

October 27, 2017

Rather than isolated instances of corruption now entire fields of research are found diseased

Statistics in the fray

The discipline of statistics has been going through a phase of critique and selfcriticism, due to mounting evidence of poor statistical practice of which misuse and abuse of the P-test is the most visible sign



732 North Washington Street, Alexandria, VA 22314 • (703) 684-1221 • Toll Free: (888) 231-3473 • www.amstat.org • www.twitter.com/AmstatNews

AMERICAN STATISTICAL ASSOCIATION RELEASES STATEMENT ON STATISTICAL SIGNIFICANCE AND P-VALUES

Provides Principles to Improve the Conduct and Interpretation of Quantitative

Science March 7, 2016

+ twenty 'dissenting' commentaries

Wasserstein, R.L. and Lazar, N.A., 2016. 'The ASA's statement on p-values: context, process, and purpose', The American Statistician, DOI:10.1080/00031305.2016.1154108.

See also Christie Aschwanden at http://fivethirtyeight.com/features/not-even-scientists-can-easily-explain-p-values/

P-hacking (fishing for favourable p-values) and HARKing (formulating the research Hypothesis) After the Results are Known); Desire to achieve a sought for – or simply publishable – result leads to fiddling with the data points, the modelling assumptions, or the research hypotheses themselves

Leamer, E. E. Tantalus on the Road to Asymptopia. J. Econ. Perspect. 24, 31-46 (2010).

Kerr, N. L. HARKing: Hypothesizing After the Results are Known. Personal. Soc. Psychol. Rev. 2, 196–217 (1998).

A. Gelman and E. Loken, "The garden of forking paths: Why multiple comparisons can be a problem, even when there is no 'fishing expedition' or 'p-hacking' and the research hypothesis was posited ahead of time," 2013.

An existential crisis?

Most observers have noted that the crisis has technical as well as ethical and behavioural elements which interact with one another – e.g. the 'publish or perish' obsession has an impact on selection bias – the tendency to favour positive over negative results Bad science reproduces better than the good sort Downloaded from http://rsos.royalsocietypublishing.org/ on September 23, 2016

ROYAL SOCIETY OPEN SCIENCE

rsos.royalsocietypublishing.org



Cite this article: Smaldino PE, McElreath R. 2016 The natural selection of bad science. *R. Soc. open sci.* **3**: 160384. http://dx.doi.org/10.1098/rsos.160384

Received: 1 June 2016 Accepted: 17 August 2016

The natural selection of bad science

Paul E. Smaldino¹ and Richard McElreath²

¹Cognitive and Information Sciences, University of California, Merced, CA 95343, USA ²Department of Human Behavior, Ecology, and Culture, Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany

D PES, 0000-0002-7133-5620; RME, 0000-0002-0387-5377

Poor research design and data analysis encourage false-positive findings. Such poor methods persist despite perennial calls for improvement, suggesting that they result from something more than just misunderstanding. The persistence of poor methods results partly from incentives that favour them, leading to the natural selection of bad science. This dynamic requires no conscious strategizing—no deliberate cheating nor loafing by scientists, only that publication is a principal factor for As in the real world, successful labs produce more 'progeny,' such that their methods are more often copied and their students are more likely to start labs of their own. Selection for high output leads to poorer methods and increasingly high false discovery rates.

Improving the quality of

research requires change at the institutional level.

Smaldino PE, McElreath R., 2016 The natural selection of bad science. R. Soc. open sci. 3: 160384. http://dx.doi.org/10.1098/rsos.160384

Bad science is 'sticky'



Article Open Access 💿 🛈

Do rebuttals affect future science?

Jeannette A. Banobi 💌, Trevor A. Branch, Ray Hilborn

First published: 30 March 2011 | https://doi.org/10.1890/ES10-00142.1 | Cited by: 13

"We examined seven high-profile original articles and their rebuttals, finding that original articles were cited 17 times more than rebuttals, and that annual citation numbers were unaffected by rebuttals" For Gigerenzer & Marewski statistics has changed the nature all disciplines …

… Creating a surrogate science based on worshipping p-values

Better to have no beliefs than to embrace falsehoods... (→ F. Bacon's idols)

G. Gigerenzer and J. N. Marewski, "Surrogate Science," J. Manage., vol. 41, no. 2, pp. 421-440, Feb. 2015.

Magnitude-based inference: <u>persistent</u> bad stats in sports research

MBI false positive rate two to six time higher than in NHST (Null hypothesis significance testing)

Christie Aschwanden and Mai Nguyen, How Shoddy Statistics Found A Home In Sports Research, Fivethirtyeight, May 16, 2018, https://fivethirtyeight.com/features/how-shoddystatistics-found-a-home-in-sports-research/

K. L. Sainani, "The Problem with 'Magnitude-Based Inference," Med. Sci. Sport. Exerc., p. 1, Apr. 2018.

Bad science or bad journals?



Prestigious Science Journals Struggle to Reach Even Average Reliability

"...an accumulating body of evidence suggests that methodological quality & reliability of published research works in several fields may be decreasing with increasing journal rank" (20 February, 2018)

Björn Brembs*

Institute of Zoology–Neurogenetics, Universität Regensburg, Regensburg, Germany



Fang FC, Casadevall A and Morrison R (2011) Retracted science and the retraction index. *Infection and Immunity* 79(10): 3855–3859 The 'bad' in bad science is not just methodological
Evidence based medicine hijacked to serve corporate agendas. "Under market pressure, clinical medicine has been transformed to finance-based medicine" (David Sackett interviewed by J.P.A. Ioannidis)

J. P. A. Ioannidis, "Evidence-based medicine has been hijacked: a report to David Sackett," J. Clin. Epidemiol., vol. 73, pp. 82–86, May 2016. Power asymmetries in the framing of issues: those who have the deepest pockets marshal the best evidence → Instrumental use of quantification to obfuscate

A. Saltelli and M. Giampietro, "What is wrong with evidence based policy, and how can it be improved?," Futures, vol. 91, pp. 62–71, Feb. 2017.

A. Saltelli and S. Funtowicz, "What is science's crisis really about?," Futures, vol. 91, pp. 5–11, 2017.

Unintended effects of reforms

Good intentions going bad

Incentive	Intended effect	Actual effect
"Researchers rewarded for increased number of publications."	"Improve research productivity," provide a means of evaluating performance.	"Avalanche of" substandard, "incremental papers"; poor methods and increase in false discovery rates leading to a "natural selection of bad science" (Smaldino and Mcelreath, 2016); reduced quality of peer review
"Researchers rewarded for increased number of citations."	Reward quality work that influences others.	Extended reference lists to inflate citations; reviewers request citation of their work through peer review
"Researchers rewarded for increased grant funding."	"Ensure that research programs are funded, promote growth, generate overhead."	Increased time writing proposals and less time gathering and thinking about data. Overselling positive results and downplay of negative results.
Increase PhD student productivity	Higher school ranking and more prestige of program.	Lower standards and create oversupply of PhDs. Postdocs often required for entry-level academic positions, and PhDs hired for work MS students used to do.
Reduced teaching load for research- active faculty	Necessary to pursue additional competitive grants.	Increased demand for untenured, adjunct faculty to teach classes.
"Teachers rewarded for increased student evaluation scores."	"Improved accountability; ensure customer satisfaction."	Reduced course work, grade inflation.
"Teachers rewarded for increased student test scores."	"Improve teacher effectiveness."	"Teaching to the tests; emphasis on short-term learning."
"Departments rewarded for increasing U.S. News ranking."	"Stronger departments."	Extensive efforts to reverse engineer, game, and cheat rankings.
"Departments rewarded for in- creasing numbers of BS, MS, and PhD degrees granted."	"Promote efficiency; stop students from being trapped in degree programs; impress the state legislature."	"Class sizes increase; entrance requirements" decrease; reduce graduation requirements.
"Departments rewarded for increasing student credit/contact hours (SCH)."	"The university's teaching mission is fulfilled."	"SCH-maximization games are played": duplication of classes, competition for service courses.

TABLE 1. GROWING PERVERSE INCENTIVES IN ACADEMIA

Modified from Regehr (pers. comm., 2015) with permission.

Academic Research in the 21st Century: Maintaining Scientific Integrity in a Climate of Perverse Incentives and Hyper-competition, Marc A. Edwards and Siddhartha Roy, ENVIRONMENTAL ENGINEERING SCIENCE, 34(1), 2017

Incentive

"Researchers rewarded for increased number of publications." Academic Research in the 21st Century: Maintaining Scientific Integrity in a Climate of Perverse Incentives and Hypercompetition, Marc A. Edwards and Siddhartha Roy, ENVIRONMENTAL ENGINEERING SCIENCE, 34(1), 2017

Intended effect

"Improve research productivity," provide a means of evaluating performance. Actual effect

"Avalanche of" substandard, "incremental papers"; poor methods and increase in false discovery rates leading to a "natural selection of bad science" (Smaldino and Mcelreath, 2016); reduced quality of peer review

See also P. Mirowski, "The future(s) of open science," Soc. Stud. Sci., vol. 48, no. 2, pp. 171-203, Apr. 2018.

Not all disciplines the same



OPEN ORCESS Freely available online



"Positive" Results Increase Down the Hierarchy of the Sciences

Daniele Fanelli*

INNOGEN and ISSTI-Institute for the Study of Science, Technology & Innovation, The University of Edinburgh, Edinburgh, United Kingdom

"odds of reporting a positive result ~5 times higher among papers in the disciplines of Psychology and Psychiatry and Economics and Business than Space Science" April 7, 2010

Physics as a model:

R. Horton, "Offline: What is medicine's 5 sigma?," Lancet, vol. 385, p. 1380, 2015.

Saul Perlmutter, an astrophysicist at the University of California, Berkeley. "Science is an ongoing race between our inventing ways to fool ourselves, and our inventing ways to avoid fooling ourselves."

R. Nuzzo, "How scientists fool themselves – and how they can stop," Nature, vol. 526, no. 7572, pp. 182–185, Oct. 2015.

Is modelling 'breaking bad'?

Unlike statistics, mathematical modelling is not a discipline, hence the lack of appropriate internal antibodies to fight a possible infection in the form of quality standards, disciplinary fora and journals and recognized leaders The heterogeneous nature of the modelling and simulation community prevents the emergence of consolidated paradigms \rightarrow

Serification and verification procedures are a rather trial and error business

This is a survey involving 283 responding modellers in J. J. **Padilla**, S. Y. Diallo, C. J. Lynch, and R. Gore, "Observations on the practice and profession of modeling and simulation: A survey approach," Simulation, vol. I14, 2017

Most users unaware of limitations, uncertainties, omissions and subjective choices in models → over-reliance in the quality of model-based inference

Modellers oversimplify or overelaborate, obfuscating model use

A large review of several existing checklists model quality: A. J. **Jakeman**, R. A. Letcher, and J. P. Norton, "Ten iterative steps in development and evaluation of environmental models," Environ. Model. Softw., vol. 21, no. 5, pp. 602–614, 2006.

Padilla et al. call for a more structured, generalized and standardized approach to verification

Jakeman et al. call for a 10 points participatory checklist including NUSAP and J. R. Ravetz's process based approach

For NUSAP: Funtowicz, S.O., Ravetz, J.R., 1990. Uncertainty and Quality in Science and Policy. Kluwer, Dordrecht

J. R. Ravetz, "Integrated Environmental Assessment Forum, developing guidelines for 'good practice', Project ULYSSES.," 1997.http://www.jvds.nl/ulysses/eWP97-1.pdf

Modelling as a craft rather than as a science for Robert Rosen

Modelling as distinct from physical laws which can be falsified for Naomi Oreskes

R. Rosen, Life Itself: A Comprehensive Inquiry Into the Nature, Origin, and Fabrication of Life. Columbia University Press, 1991.

N. Oreskes, K. Shrader-Frechette, and K. Belitz, "Verification, Validation, and Confirmation of Numerical Models in the Earth Sciences," Science, 263, no. 5147, 1994.

N. Oreskes, "Prediction: science, decision making, and the future of nature" in D. Sarewitz, R. A. Pielke, Jr., and R. Byerly, Jr. Eds. in Prediction, Science, Decision Making and the future of Nature, Island Press, 2010. Egregious modelling failure from Pilkey and Pilkey-Jarvis (from AIDS to coastal erosion…)

For John Kay modelling needs as input information which we don't have (The case of WEBTAG and knowing car passengers number decades into futures)

O. H. Pilkey and L. Pilkey-Jarvis, Useless Arithmetic: Why Environmental Scientists Can't Predict the Future. Columbia University Press, 2009.

J. A. Kay, "Knowing when we don't know," 2012, https://www.ifs.org.uk/docs/john_kay_feb2012.pdf

Economics

Paul Romer's Mathiness = use of mathematics to veil normative stances

Erik Reinert: scholastic tendencies in the mathematization of economics

P. M. Romer, "Mathiness in the Theory of Economic Growth," Am. Econ. Rev., vol. 105, no. 5, pp. 89–93, May 2015.

E. S. Reinert, "Full circle: economics from scholasticism through innovation and back into mathematical scholasticism," J. Econ. Stud., vol. 27, no. 4/5, pp. 364–376, Aug. 2000.

Cooping with uncertainty or quantification hubris

The main issue in existing practices of mathematical modelling is in the management of uncertainty in model-based inference. Modelling studies can be seen which tend to overestimate certainty, pretending to produce crisp numbers precise to the third decimal digits even in situation of pervasive uncertainty or ignorance



How uncertainty is downplayed in modelling studies: the case of sensitivity analysis

An engineer's vision of UA, SA



Can one lie with sensitivity analysis as one can lie with statistics?



Saltelli, A., Annoni P., 2010, How to avoid a perfunctory sensitivity analysis, Environmental Modeling and Software, 25, 1508–1517.

In 2014 out of 1000 papers in modelling 12 have a sensitivity analysis and < 1 a global SA; most SA still move one factor at a time



Ferretti, F., Saltelli A., Tarantola, S., 2016, Trends in Sensitivity Analysis practice in the last decade, Science of the Total Environment, http://dx.doi.org/10.1016/j.scitotenv.201 6.02.133

------ TOT_SA/TOT_MOD (%) ------ TOT_GSA/TOT_MOD (%)

OAT in 2 dimensions



Area circle / area square =?

~ 3/4

OAT in 3 dimensions



Volume sphere / volume cube =?

~ 1/2

OAT in 10 dimensions; Volume hypersphere / volume ten dimensional hypercube =? ~ 0.0025





Once a sensitivity analysis is done via OAT there is no guarantee that either uncertainty analysis (UA) or sensitivity analysis (SA) will be any good:

→ UA will be non conservative

SA may miss important factors

Why ethics of quantification?

Just as per the case of statistics, no solution is possible without careful appraisal of the social and cultural dimensions of the problem. We suggest that the situation calls an ethics of quantification to be developed, analogous to what is happening in the field of algorithms and big data.

Symbiotic relationship between quantification and trust

Theodor M. Porter





Porter's story: Quantification needs judgment which in turn needs trust …without trust quantification becomes mechanical, a system, and systems can be gamed





Big data and algorithms

Algorithms decide upon an ever-increasing list of cases, such as recruiting, carriers – including of researchers, prison sentencing, paroling, custody of minors…

Alexander, L. Is an algorithm any less racist than a human? | Technology | The Guardian. Available at https//www.theguardian.com/technology/2016/aug/03/algorithm-racist-human-employers-work (2016) (Accessed: 30th August 2017).

Abraham C. Turmoil rocks Canadian biomedical research community. Statnews, Available at https://www.statnews.com/2016/08/01/cihr-canada-research/ (2016) (Accessed: 30th August 2017).

R. Brauneis and E. P. Goodman, "Algorithmic Transparency for the Smart City," Algorithmic Transpar. Smart City, vol. 20, pp. 103–176, 2018.

Weapons of Math Destruction

O'Neil, C. Weapons of math destruction : how big data increases inequality and threatens democracy. (Crown/Archetype, 2016).

Algorithmic audit in New York city

Dwyer J. Showing the Algorithms Behind New City Services - The New York Times. New Yo Times Aug. 24, (2014).



Mathematical modelling does not make it to the headlines but is possibly in a worse shape Statistical modelling

Mathematical

modelling

Algorithms

Blurring lines:

"what qualities are specific to rankings, or indicators, or models, or algorithms?"

E. Popp Berman and D. Hirschman, The Sociology of Quantification: Where Are We Now?, Contemp. Sociol., vol. in press, 2017.

Ethics of quantification; a new grammar for mathematical modelling?

1. Uncertainty and sensitivity analysis (never execute the model once)

2. Sensitivity auditing and quantitative storytelling (investigate frames and motivations)

Saltelli, A., Guimarães Pereira, Â., Van der Sluijs, J.P. and Funtowicz, S., 2013, 'What do I make of your latinorum? Sensitivity auditing of mathematical modelling', Int. J. Foresight and Innovation Policy, (9), 2/3/4, 213–234.

Saltelli, A., Does Modelling need a reformation? Ideas for a new grammar of modelling, available at https://arxiv.org/abs/1712.06457

3. Replace 'model to predict and control the future' with 'model to help mapping ignorance about the future' ...

... in the process exploiting and making explicit the metaphors embedded in the model

J. R. Ravetz, "Models as metaphors," in Public participation in sustainability science : a handbook, and W. A. B. Kasemir, J. Jäger, C. Jaeger, Gardner Matthew T., Clark William C., Ed. Cambridge University Press, 2003, available at http://www.nusap.net/download.php?op=getit&lid=11

END


Extra slides

Statistics as a garden of forking paths even with no explicit HARKing





Jorge Luis Borges

Andrew Gelman

http://www.stat.columbia.edu/~gelman/research/unpublished/p_hacking.pdf

S|S|S

The future(s) of open science

Social Studies of Science 2018, Vol. 48(2) 171–203 © The Author(s) 2018 Reprints and permissions: sagepub.co.uk/journalsPermissions.nav DOI: 10.1177/0306312718772086 journals.sagepub.com/home/sss





John J. Reilly Center, University of Notre Dame, Notre Dame, IN, USA

Abstract

Almost everyone is enthusiastic that 'open science' is the wave of the future. Yet when one looks seriously at the flaws in modern science that the movement proposes to remedy, the prospect for improvement in at least four areas are unimpressive. This suggests that the agenda is effectively to re-engineer science along the lines of platform capitalism, under the misleading banner of opening up science to the masses.

EC impact assessment guidelines: sensitivity analysis & auditing

European Commission European Commission > Better Regulatio	n > Guidelines	
Home	Better Regulation Guidelines	🐼 Share 🚺 🚺
 Stakeholder consultations Roadmaps / Inception Impact Assessments Impact Assessment Evaluation Regulatory Scrutiny Board Guidelines Better Regulation Guidelines Better Regulation "Toolbox" Key documents 	 These guidelines explain what Better Regulation is and how it should be applied in the day to day practices when preparing new initiatives and proposals or managing existing policies and legislation. They cover the whole policy cycle, from policy preparation and adoption to implementation and application, to evaluation and revision of EU law. For each of these phases there are a number of Better Regulation principles, objectives, tools and procedures to make sure that the EU has the best regulation possible. These relate to planning, impact assessment, stakeholder consultation, implementation and evaluation. The <u>Better Regulation Quidelines</u> are structured into chapters which cover each of the instruments of the law-making process. The corresponding <u>toolbox</u> gives more detailed and technical information. Better Regulation Guidelines are based on the outcomes of public consultation exercises carried out in 2013 and 2014. <u>Public consultation on the revision of the Commission's Impact Assessment Guidelines</u> <u>Stakeholder Consultation Guidelines</u> <u>Stakeholder Consultation Guidelines</u> <u>Consultation on the draft Commission Evaluation Policy Guidelines</u> 	Search Stay connected Facebook Stay connected Facebook Subtract

http://ec.europa.eu/smart-regulation/guidelines/docs/br_toolbox_en.pdf

August Comte (1798–1857)

Hierarchy of Sciences



Philip Mirowski devotes a full chapter in Never Let a Serious Crisis Go to Waste to disparage the OVer-reliance on DSGE (Dynamic Stochastic General Equilibrium) models

P. Mirowski, Never Let a Serious Crisis Go to Waste: How Neoliberalism Survived the Financial Meltdown. Verso, 2013.

Rules for sensitivity analysis

1. Never run a model just once

 Sensitivity analysis is not "run" on a model but on a model once applied to a question

3. Sensitivity analysis should not be used to hide assumptions

4. If SA shows that a question cannot be answered change either the question or the model (don't shave the uncertainties)

5. SA shows that there is always one more bug! (Lubarsky's Law of Cybernetic Entomology)

6. Never run a SA where each factors has a 5% uncertainty range

The rules of sensitivity auditing

- Check against rhetorical use of mathematical modelling;
- 2. Adopt an "assumption hunting" attitude; focus on unearthing possibly implicit assumptions;
- 3. Check if uncertainty been instrumentally inflated or deflated.

4. Find sensitive assumptions before these find you; do your SA before publishing;

5. Aim for transparency; Show all the data;

6. Do the right sums, not just the sums right; frames; → quantitative storytelling

7. Perform a proper global sensitivity analysis.

The importance of frames Quantitative storytelling Frames; The expression 'tax relief' is apparently innocuous but it suggests that tax is a burden, as opposed to what pays for road, hospitals, education ...

George Lakoff



Lakoff, G., 2010, Why it Matters How We Frame the Environment, Environmental Communication: A Journal of Nature and Culture, 4:1, 70-81.

Lakoff, G., 2004–2014, Don't think of an elephant: know your values and frame the debate, Chelsea Green Publishing.





For Akerlof and Shiller against what the 'invisible hand' would contend - economic actors have no choice but to exploit frames to 'phish' people into practices which benefit the actors not the subject phished



George Akerlof



Robert R. Shiller

Quantitative storytelling tests frames/narratives for:

- Internal contradictions
- Feasibility (outside human control);
- Viability (under human control); and
- Desirability (normative; plurality of actors)

An example: Sensitivity auditing of the OECD PISA study



Q. Search analysis, research, academics



Arts + Culture Business + Economy Cities Education Environment + Energy FactCheck Health + Medicine Politics + Society Science + Technology

L. Araujo, A. Saltelli, and S. V. Schnepf, "Do PISA data justify PISA-based education policy?," Int. J. Comp. Educ. Dev., vol. 19, no. 1, pp. 20-34, 2017.

Saltelli, A., International PISA tests show how evidence-based policy can go wrong, The Conversation, June 12, 2017

With PISA the OECD gained the centre-stage in the international arena on education policies, which led to important controversies

http://www.theguardian.com/education/2014/may/06/oecdpisa-tests-damaging-education-academics

theguardian

OECD and Pisa tests are damaging education worldwide - academics

In this letter to Dr Andreas Schleicher, director of the OECD's Programme for International Student Assessment, academics from around the world express deep concern about the impact of Pisa tests and call for a halt to the next round of testing



School children in Sichuan province in China. Academics say the OECD should develop alternatives to league tables and find more meaningflu ways of reporting assessment, taking account of different cultures. Photograph: James Zeng Huang/Corbis Sygma

Critical remarks by 80 signatories of the letter:

- Flattening of curricula (exclusion of subjects)
- Short-termism (teaching to the test)
- Promoting "life skills to function in knowledge societies"
- Stressing the student
- ••• \rightarrow Stop the test!
- Ask for more participation in design

Figure 1 Present value of Scenario I (improve student performance in each country by 25 points on the PISA scale) in billion USD (PPP)



Note: Discounted value of future increases in GDP until 2090 due to reforms that improve student performance in each

http://www.oecd.org/edu/school/programmeforinternationalstudentassessmentpisa/t hehighcostofloweducationalperformance.htm

"If every EU Member State achieved an improvement of 25 points in its PISA score as Germany and Poland did over the last decade, the GDP of the whole EU would increase by between 4% and 6% by 2090; such a 6% increase would correspond to 35 trillion Euro"

Woessmann, L. (2014), "The economic case for education", EENEE Analytical Report 20, European Expert Network on Economics of Education (EENEE), Institute and University of Munich.

We find both technical and normative issues:

1) Non response bias (which students are excluded) PISA non-response for England: the bias turned out to be twice the size of the OECD declared standard error in 2003

2) Non open data, which makes SA impossible

3) Flattening curricula (do all countries wish to prosper by becoming knowledge societies?)

4) Power implications: OECD (unelected officers and scholars) becoming a global superministry of education

Initiatives to solve the crisis in science

Fixing science?









Brian Nosek, the John Ioannidis, Meta-Reproducibility research innovationProject. centre at Stanford

Ben Goldacre, alltrials.net

Gary Taubes, The case against sugar

From science's crisis to science wars?

A science war against trump, predicted in 2016



January 27, 2017

To tackle the post-truth world, science must reform itself

Andrea Saltelli, University of Bergen and Silvio Oscar Funtowicz, University of Bergen Scientists must bear some responsibility for the post-truth era and the current crisis in democracy.



November 16, 2016

Science wars in the age of Donald Trump

Andrea Saltelli, University of Bergen and Silvio Oscar Funtowicz, University of Bergen Is the election of Donald Trump going to reignite a futile war between science and anti-science?

What the present science war looks like:

Opinion: Is science really facing a reproducibility crisis, and do we need it to?

Daniele Fanelli

PNAS March 12, 2018. 201708272; published ahead of print March 12, 2018. https://doi.org/10.1073 /pnas.1708272114



"The new "science is in crisis" narrative is not only empirically unsupported, but also quite obviously counterproductive. Instead of inspiring younger generations to do more and better science, it might foster in them cynicism and indifference. Instead of inviting greater respect for and investment in research, it risks discrediting the value of evidence and feeding antiscientific agendas."

What the present science war looks like:

Crisis or self-correction: Rethinking media narratives about the well-being of science

Kathleen Hall Jamieson

PNAS March 13, 2018. 115 (11) 2620-2627; published ahead of print March 12, 2018. https://doi.org /10.1073/pnas.1708276114

1	Submit /	About	Contact	Journal Club	Subscribe
1	PN	A	S a	Proceedings of the National Academy o of the United States	f Sciences of America
	Home	Art	ic l es	Front Matter	News

"Because those whose work is prominently cited to certify that science is broken … are spearheading efforts to solve identified problems, their work is evidence of the resilience of science."

THE IRREPRODUCIBILITY CRISIS OF MODERN SCIENCE

Causes, Consequences, and the Road to Reform



DAVID RANDALL AND CHRISTOPHER WELSE NATIONAL ASSOCIATION OF SCHOLARS APRIL 2018 ISBN: 978-0-9986635-5-5



On the other (conservatives, corporations) side:

"31. Congress should pass an expanded Secret Science Reform Act to prevent government agencies from making regulations based on irreproducible research.

National Association of Scholars

THE GLOBAL WARMING POLICY FOUNDATION

Director: Dr Benny Peiser

Common Sense on Climate Change

GWPF

Date: 27/10/16 | Global Warming Policy Foundation

PEER REVIEW Why skepticism is essential

Donna Laframboise

On the other (conservatives, corporations) side:

"If half of published, peer-reviewed papers 'may simply be untrue', half of the papers cited by the IPCC may also be untrue..."