Finnish Society of Forest Science
Open Forest Science seminar, 7 May 2018

Transparent writing

Pekka Nygren, Associate Editor, Silva Fennica
https://silvafennica.fi

e-mail: pekka.nygren@metsatiede.org

Suomen Metsätieteellinen Seura

Quality of science

- Three main building blocks:
  - Objectivity (honesty) of the scientists
  - Peer-review
  - Reproducibility
- Openness (Transparency)

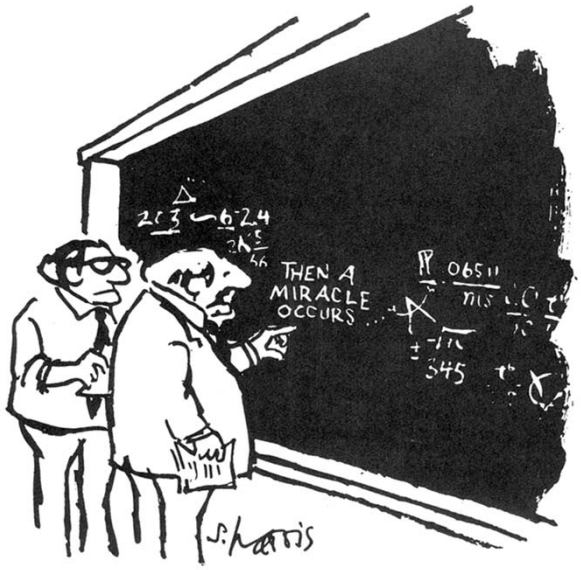
Objectivity?

- Serious scientific misconduct: fabrication, falsification and plagiarism:
  - about 2 % of scientists admit misconduct during the past three years (5 % in economics).
- Sloppy science – 50 shades of grey between misconduct and good scientific practice:
  - about 34 % of scientist admit sloppy science (94 % in economics).

Prof. Lex Bouter, Free University of Amsterdam, EASE Conference 11 June 2016

Sloppy science in practice

- Several analyses are done but only those supporting the hyptheses are published.
- ”Answering” to questions, for which the methods are not suitable.
- Unpleasant results are not published.



"I think you should be more explicit here in step two."

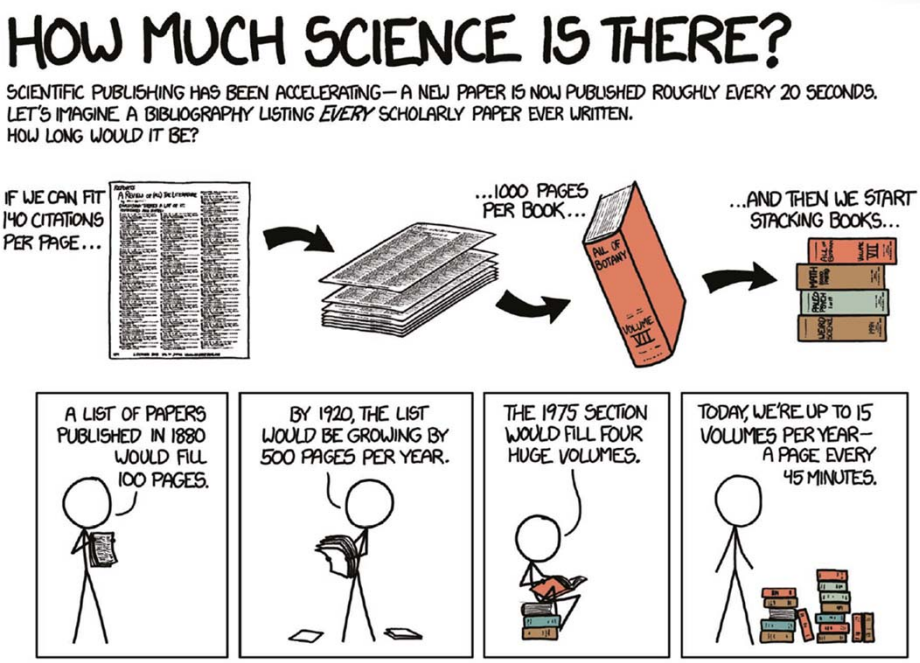
Peer-review

- Peer-review is considered to guarentee the quality of scientific publications.
- Peer-review system is saturating because of the increasing publication volume: there are less qualified reviewers than authors;
  - e.g. in Silva Fennica, (rejection-% 65) six reviews are needed for each accepted article.

➢ increasingly difficult to recruit reviewers,

➢ the reviewers do not have time to do thorough reviews,

➢ in the fragmented science, the best reviewers probably know the authors, or even collaborate with them.



Fortunately, reproducibility helps to detect the problems with sloppy writing and peer-review – doesn't it?

EDITORIAL

Reproducibility

Marcia McNutt is Editor-in-Chief of *Science*.

SCIENCE ADVANCES ON A FOUNDATION OF TRUSTED DISCOVERIES. REPRODUCING AN EXPERIMENT is one important approach that scientists use to gain confidence in their conclusions. Recently, the scientific community was shaken by reports that a troubling proportion of

THE CANCER TEST

A nonprofit's effort to replicate 50 top cancer papers is shaking up labs

By Jocelyn Kaiser

BIOMEDICAL RESEARCH

Calling all failed replication experiments

New journal will publish methods, data, and results

By Jocelyn Kaiser

to show data and let people decide for

tool molecule to test bexarotene's purp

REPRODUCIBILITY

Many psychology papers fail replication test

An effort to repeat 100 studies yields sobering results, but many researchers are positive about the process

By John Bohannon

The largest effort yet to replicate psychology studies has yielded both

experimental footing. The strategy was to replicate a sample of published studies using an approach that Nosek has popularized through the Center for Open Science, a

SCIENTIFIC INTEGRITY

Survey fraud test sparks battle

Pew Research Center challenges statistical test

By John Bohannon

positives. The organization has gone so far as to request Robbins and Kuriakose de-

What is reproducibility?

- Repetition of analyses
- Direct repetition of an experiment
- Conceptual repetition: repetition of a study under different conditions

- Possible only if
  - reporting is transparent,
  - data are open,
  - analyses are open,
  - source code is open.

Forest research is often inherently non-reproducible

SILVA FENNICA

Silva Fennica vol. 49 no. 4 article id 1408
Category: research article
www.silvafennica.fi
ISSN-L 0037-5330 | ISSN 2242-4075 (Online)
The Finnish Society of Forest Science
Natural Resources Institute Finland

Anna K. Franke[1], Pasi Aatsinki[1], Ville Hallikainen[1], Esa Huhta[1], Mikko Hyppönen[1], Vesa Juntunen[2], Kari Mikkola[3], Seppo Neuvonen[4] and Pasi Rautio[5]

Quantifying changes of the coniferous forest line in Finnish Lapland during 1983–2009

The study by Franke et al. cannot be reproduced

- another 25 yrs of research for verifying the results?
  - global climate change;
  - forest succession;
  - biotic factors like reindeer grazing or pest outbreaks;
  - etc.
- Only way for reproducibility is transparent writing and opening of data and analyses for the research community;
  - cf. guidelines of the Center for Open Science, https://cos.io

Suggestions for reporting forest research

- Adopted from:
https://www.nih.gov/research-training/rigor-reproducibility/principles-guidelines-reporting-preclinical-research

1. Use community-based standards (such as nomenclature and reporting standards), where applicable.
   - e.g. either FAO or USDA soil taxonomy
2. Provide enough information to uniquely identify biological materials
   - unique accession number in a repository
   - seed source, lot number, provenance, date of collection, etc.

Reporting forest research II

3. Describe environmental conditions with sufficient detail so that the readers understand where you worked.
   - e.g. provide weather data summary and a link to a stable data base (e.g., a national meteorological institute).
4. Report how often each experiment was performed and whether the results were substantiated by repetition under a range of conditions.
   - distinguish between independent biological data points and technical replicates
   - identify pseudoreplicates as such and use appropriate statistical methods

Reporting forest research III

5. Statistics must be fully reported in the article:
   - the statistical test used
   - exact number of observations
   - definition of centre, dispersion and precision measures (e.g., mean, median, standard deviation, standard error of mean, confidence intervals)
   - distribution tests
- Statistics must always be justified by the research problem and data!

Reporting forest research IV

6. Justify the sampling method based on your research problem
   - if samples were randomised, specify the method of randomisation
   - how sample size was estimated and how an appropriate sample size was computed when designing the study
7. Clearly state the criteria that were used for exclusion of any data or subjects
   - data exclusion is the most important single cause for non-reproducibility!

A warning example

- In a recent article, it was claimed that in an arid site, only 3% of grass biomass is in roots. However,
  - in dry environments, it is expected that plants allocate resources to water acquisition
  - other studies show that a considerable proportion of grass biomass is underground (several 10s of percents)
- How the sampling was done?

> Whole plants were manually removed from the soil, split into shoot and root samples and placed into separate bags.

Publication in a JUFO Class 3 journal!

Open the data and code

- The topic of the next presentation – stay tuned!
- Necessary for evaluating ”inherently non-reproducible” forest research.
  - open code makes it possible to repeat all analyses on the original data
  - other methods may open new viewpoints to the published data (provided that they are suitable for the problem)
  - useless if the article reporting the research is written in a sloppy way

Transparent article

- Transparency: Study materials, methods, and analyses are well described, and special attention is paid on data exclusion and inclusion criteria.
- Accuracy: biological materials and environmental conditions are reported following appropriate standards when available.
- Openness: Data and software are made publicly available when possible.

- Even after a careful peer-review, the authors are responsible on the accuracy and reproducibility of the reported work.