COGNITIVE SCIENCE
IN SEARCH OF UNITY

Replicability of Computational Models: Achilles Heel of Neuroscience

Marcin Miłkowski, Witold Hensel & Mateusz Hohol

IACAP 2018

Behavioral and Brain Sciences

MAKING REPLICATION MAINSTREAM

Rolf A. Zwaan [a1], Alexander Etz [a1], Richard E. Lucas [a1] and M. Brent Donnellan [a1]

https://doi.org/10.1017/S0140525X17001972 Published online: 25 October 2017

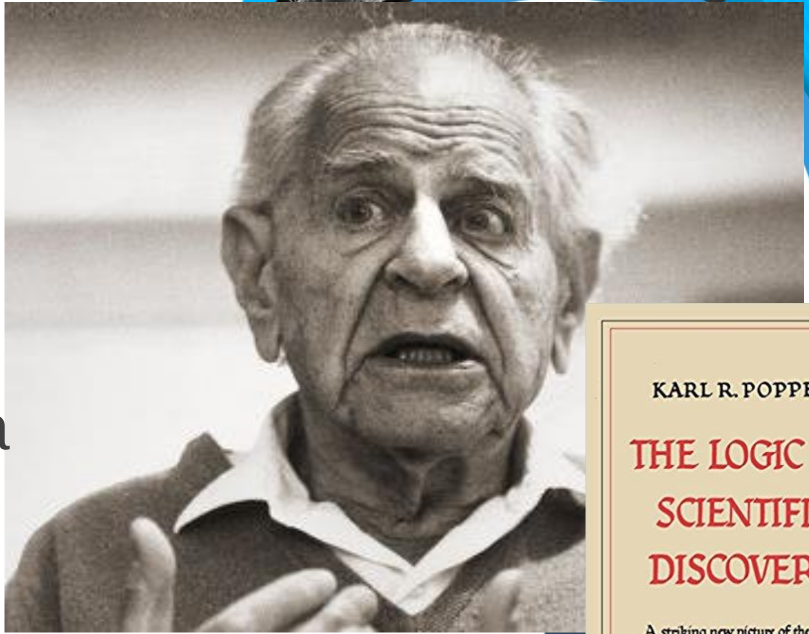
Many philosophers of science and methodologists have argued that the ability to repeat studies and obtain similar results is an essential component of science. A finding is elevated from single observation to scientific evidence when the procedures that were used to obtain it can be reproduced and the finding itself can be replicated. Recent replication attempts show that some high profile results---most notably in psychology, but in many other disciplines as well---cannot be replicated consistently. These replication attempts have generated a considerable amount of controversy and the issue of whether direct replications have value has, in particular, proven to be contentious. However, much of this discussion has occurred in published commentaries and social media outlets, resulting in a fragmented discourse. To address the need for an integrative summary, we review various types of replication studies and then discuss the most commonly voiced concerns about direct replication. We provide detailed responses to these concerns and consider different statistical ways to evaluate replications. We conclude there are no theoretical or statistical obstacles to making direct replication a routine aspect of psychological science.

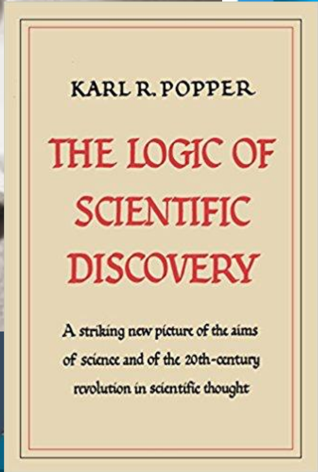
Plan of the talk

- Discussion of the general requirement of repeatability
- Various types and functions of replication studies in psychology
- Replicability of computational models
- Functions of replications in computational modeling
- The principle of completeness in computational modeling

Repeatability of empirical studies

“Any controversy over the question whether events which are in principle unrepeatable and unique ever do occur cannot be decided by science: it would be a metaphysical controversy”
(Popper, 1959)





Functions of replication studies in psychology (Schmidt, 2009)

- discovering false positives,
- controlling for artifacts,
- addressing researcher fraud,
- attempting to generalize a result to a different population,
- trying to confirm a previously supported hypothesis using a different experimental procedures…

Types of replication in psychology

(Schmidt 2009, Stroebe & Strack 2014, Zwaan et al. 2017)

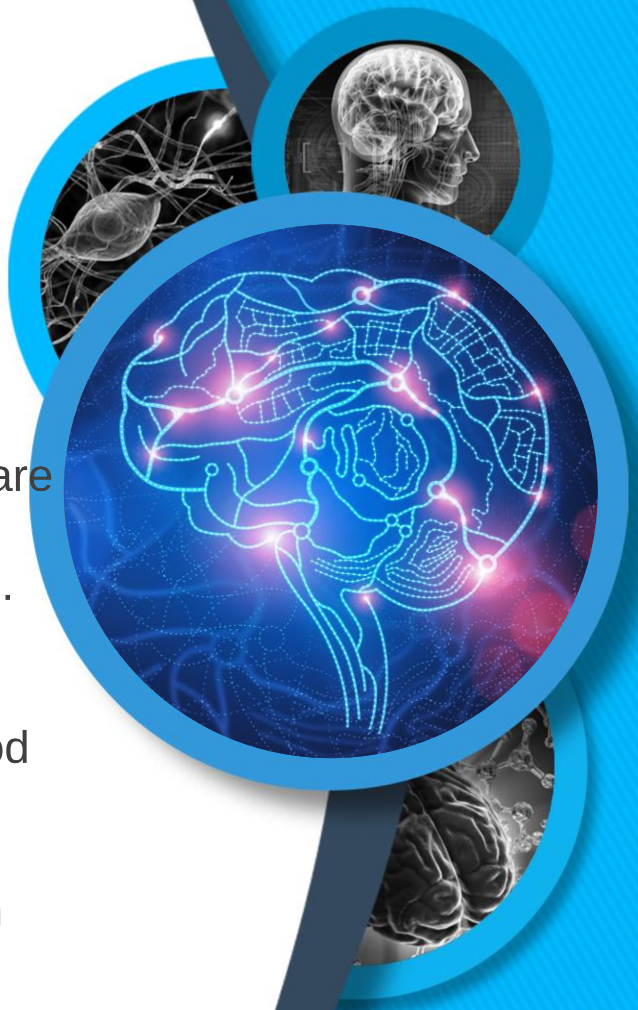
- Direct replications - recreation an original study (its samples, measures, procedures, etc.) according to the current understanding of what is needed to produce the phenomenon under investigation.

- Conceptual replications - modifying the critical elements of an original procedure in order to test the robustness of a phenomenon or the generality of a theoretical claim.

Types of direct replications (Hüffmeier et al., 2016)

- Exact replications - replications performed by the same group of researchers, is to protect the scientific community against false positives, which are likely to occur when the first study is statistically underpowered; exact replications are strongly recommended when initial findings are either unexpected or loosely based on current theoretical models.

- Close replications - replications performed by an independent team of researchers, also reduce the likelihood of false positives, especially those stemming from experimenter effects and tacit knowledge; they provide information needed to establish the size of an effect, which the original investigators are prone to overestimate.



The replication crisis in computational modeling

- A number of authors are drawing attention to problems with research reproducibility (Hutson 2018, Peng 2011, Rougier et al. 2017, Sandve et al. 2013).
- Computational model replication is rarely performed because successful replications do not seem to deliver novel scientific results and causes of failed replication may be difficult to discern (Legéndi et al. 2013).
- Instead of replicating a model with new data, researchers tend to compare new models with previous work.

Reproducible Science is good. Replicated Science is better.

ReScience is a peer-reviewed journal that targets computational research and encourages the explicit replication of already published research, promoting new and open-source implementations in order to ensure that the original research is reproducible.

To achieve this goal, the whole publishing chain is radically different from other traditional scientific journals. ReScience lives on GitHub where each new implementation of a computational study is made available together with comments, explanations and tests. Each submission takes the form of a pull request that is publicly reviewed and tested in order to guarantee that any researcher can re-use it. If you ever replicated computational results from the literature in your research, ReScience is the perfect place to publish your new implementation.

ReScience is collaborative by design. Everything can be forked and modified. Don't hesitate to write a submission, join us and to become a reviewer.

Current activity

The ReScience Journal

Published articles: 19
Submitted articles awaiting review: 0
Articles currently under review: 4 (#39, #41, #45, #47)
Articles awaiting publication: 0
Replication rate: 100%

frontiers in Neuroinformatics

ORIGINAL RESEARCH
published: 21 February 2017
doi: 10.3389/fninf.2017.00011

Reproducibility and Comparability of Computational Models for Astrocyte Calcium Excitability

Tiina Manninen, Riikka Havela and Marja-Leena Linne *

Computational Neuroscience Group, Faculty of Biomedical Sciences and Engineering and BioMediTech Institute, Tampere University of Technology, Tampere, Finland

The replication crisis in computational modeling

- Two of the models deal with spontaneous Ca^{2+} excitability in single astrocytes (Lavrentovich and Hemkin 2008, Riera et al. 2011); the other two simulate the neurotransmitter-evoked excitability of this element (De Pittà et al. 2009, Dupont et al. 2011).
- It is impossible to reimplement three of the models (Riera et al. De Pittà et al., Dupont et al.), due to insufficient information in published papers.
- They were able to reproduce the outcomes of only one model of astrocyte activity (by Lavrentovich and Hemkin)
- They found serious mistakes in the mathematical formalisms presented in two original papers (Riera et al., and Dupont et al.), which made exact reproduction impossible
- The models target the same phenomenon, but their performance differs significantly

Missing data hinder replication of artificial intelligence studies

By Matthew Hutson | Feb. 15, 2018 , 12:30 PM

Science

Code break

In a survey of 400 artificial intelligence papers presented at major conferences, just 6% included code for the papers' algorithms. Some 30% included test data, whereas 54% included pseudocode, a limited summary of an algorithm.

CREDITS: (GRAPHIC) E. HAND/*SCIENCE*; (DATA) GUNDERSEN AND KJENSMO, ASSOCIATION FOR THE ADVANCEMENT OF ARTIFICIAL INTELLIGENCE 2018

- Call for sharing all code is largely ignored as only 6% of the 400 algorithms presented at two top AI conferences in the past few years contained the code and only a third had pseudocode, or simplified summaries of the code.
- It may not be possible to recreate a published result even when the code is available (Crook et al., 2013)

Technical rules to ensure replicability (Sandve et al., 2013)

- tracking how results were produced,
- avoiding the manual manipulation of data sets,
- archiving the exact versions of external programs used,
- using version control to store custom scripts,
- recording intermediate results (preferably in standardized formats),
- noting random seeds for randomized analyses,
- storing raw data behind plots, connecting textual statements to underlying results and finally providing public access to scripts, runs and results

- but, before adopting a methodological rule, researchers should understand its purpose and know its possible side-effects

Types and purposes of computational replications

- Reproduction versus replication (Claerbout and Karrenbach 1992):
  - Reproduction - the procedure of obtaining the same outputs by running the same software on the same inputs (it corresponds with the notion of direct replication)
  - Replication - obtaining sufficiently similar results by designing and running new code based on a published description of a model (it corresponds with conceptual replication)
- the Association for Computing Machinery (Delling et al. 2016) has recently recommended the following usage:
  - repeatability involves a researcher being able to reliably repeat her calculations, replicability consists in a group of researchers being able to obtain the same results using an original author’s artifacts;
  - reproducibility means that an independent group of researchers can obtain the same results using artifacts which they develop completely independently.

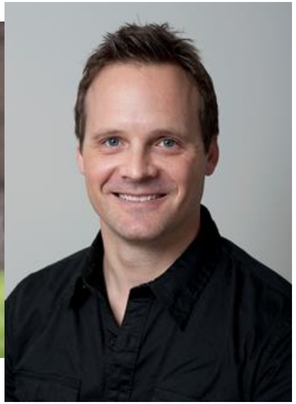
Kinds of model evaluation

- Replicating a model:
  - contributes to model verification or validation
  - demonstrates that the implementation of a model follows the official specification
  - detects type I errors in scientific papers, such as typographical mistakes in numerical values in figures or tables.

- Reproducing a model:
  - contributes to model validation,
  - discover its hidden assumptions,
  - bugs,
  - unexpected interactions

- Model reproducibility is essential to long-term scientific progress.

The British Journal for the Philosophy of Science



ACCEPTED MANUSCRIPT

Are More Details Better? On the Norms of Completeness for Mechanistic Explanations

Carl Craver, David M Kaplan

The British Journal for the Philosophy of Science, axy015,
https://doi.org/10.1093/bjps/axy015
Published: 17 January 2018

Abstract

Completeness is an important but misunderstood norm of explanation. It has recently been argued that mechanistic accounts of scientific explanation are committed to the thesis that models are complete only if they describe everything about a mechanism and, as a corollary, that incomplete models are always improved by adding more details. If so, mechanistic accounts are at odds with the obvious and important role of abstraction in scientific modelling. We respond to this characterization of the mechanist's views about abstraction and articulate norms of completeness for mechanistic explanations that have no such unwanted implications.

The principle of completeness in computational modeling

- Effective scientific communication requires that all and only relevant information is shared

- A paper should above all contain information needed to reproduce a model

- Information necessary to replicate a model should be deposited in open repositories

Thanks for your attention ! ☺

Contact

Website : http://cognitive.ifispan.pl

Email :
marcin.milkowski@gmail.com