

# Disciplinary differences in research integrity: Social sciences

Jesper W. Schneider

Centre for Studies in Research and Research Policy,

Aarhus University

[jws@ps.au.dk](mailto:jws@ps.au.dk)

# Research Ethics and Integrity for Social Scientists

Mark Israel

2<sup>nd</sup> Edition



# Danish code of conduct

## **I. Principles of Research Integrity**

This chapter presents three central principles that should pervade all phases of research:

1. Honesty
2. Transparency
3. Accountability

## **II. Responsible Conduct of Research**

This chapter presents six basic standards for conducting responsible research – from the planning phase to the dissemination of results:

1. Research planning and conduct
2. Data management
3. Publication and communication
4. Authorship
5. Collaborative research
6. Conflicts of interest


# Example of “responsibilities”

- Researchers demonstrate respect for research participants, animals used in research and the environment
- Conduct research safely
- Make and keep complete, clear, accurate records of all research
- Share findings and data openly and promptly, as soon as they have established priority and ownership claims.
- Publish and communicate research honestly and accurately
- Employ appropriate methods and use a high level of rigour and objectivity in research activities
- Appropriately cite and, where applicable, obtain permission for the use of all published and unpublished work
- Acknowledge in research outputs all contributors and contributions to the research described in the research output
- Be listed as an author of a research output only when they have made a significant intellectual or scholarly contribution that they are willing to be accountable for and agree to be listed as an author.
- Give fair, prompt and rigorous evaluations, and respect confidentiality when participating in peer review.
- Disclose financial and other conflicts of interest that could compromise the trustworthiness of their work
- ...

# Example of “breaches”

- Breaches of research integrity occurs when there is a deviation from accepted practice
- Breaches of research integrity, where there is a serious deviation from accepted practice that is intentional, reckless, or negligent, are also described as research misconduct

# The challenge

- “Thinking” and guidelines on RI and RCR derives from STEM-fields
  - The social sciences are not a uniform entity
    - Widely different meta-theoretical positions which are often in conflict with each other
    - Hence, widely different knowledge production models,
    - Different methodological perceptions, relevance and quality criteria
  - “Employ appropriate methods and use a high level of rigour and objectivity in research activities”
  - “... deviation from accepted practice”
- 
- Positivistic food for thought!

# Some cases

- Current survey case from SDU, seemingly a violation of “guaranteed” anonymity
  - If so, the case seems straightforward
- Bjørn Lomborg’s and Helmuth Nyborg’s cases at UVVU are not so straightforward
  - Political/ideological
  - Some of the breeches seem to be very common albeit going under the radar
- The recent case of Brian Wansink is exemplary as he is probably the first to be fired from his prestigious university position due to seemingly prolonged and systematic use of questionable research practices (QRPs)

# A major challenge for certain empirical knowledge production models in the social sciences and beyond: QRPs related to the use of significance tests and theory testing more generally

Open access, freely available online

## Why Most Published Research Findings Are False

John P. A. Ioannidis

**Summary**

There is increasing concern that most current published research findings are false. This probability that a research claim is true may depend on study power and bias, the number of other studies on the same question, and, to a lesser extent, the ratio of true to no relationships among the relationships posited in each scientific field. In this framework, research finding is less likely to be true when the studies conducted in a field are smaller, when effects have an earlier, when there is a greater number and lower production of related publications, when there is greater freedom to conduct, disseminate, and analyze studies, when there is greater financial and other interest and publication, and when more tests are conducted in a scientific field. In those of statistical significance. Simulations show that for most study designs and settings, it is more likely for research claim to be false than true. However, for many common scientific fields, claimed research findings may be more accurate measures of the prevailing bias in the study (discuss the implications of these problems for the research and interpretation of research).

**It can be proven that most claimed research findings are false.**

Published research findings are sometimes refuted by subsequent evidence, with ensuing confusion and disappointments. Refutation and controversy is seen across the range of research designs, from clinical trials and traditional epidemiological studies (1-3) to the most modern molecular research (4,5). There is increasing concern that in modern research, false findings may be the majority or even the vast majority of published research claims (6-8). However, this should not be surprising. It can be proven that most claimed research findings are false. Here I will examine the key factors that influence this problem and some consider them.

**Modeling the Framework for False Positive Findings**

Several methodologists have pointed out (9-11) that the high rate of nonreproducibility (lack of confirmation) of research discoveries is a consequence of the conventional, well-founded strategy of claiming conclusive research findings solely on the basis of a single study assessed by formal statistical significance, typically for a p-value less than 0.05. Research is not more appropriately represented and summarized by results, but, unfortunately, there is a widespread custom that medical research articles should be interpreted based only on p-values. Research findings are defined here as any relationship resulting from formal statistical significance, e.g., effective interventions, informative predictors, risk factors, or associations. "Negative" research is also very useful. However, here we will target relationships that investigators claim exist, rather than null findings.

As has been shown previously, the probability that a research finding is indeed true depends on the prior probability of it being true (before doing the study), the statistical power of the study, and the level of statistical significance (10,11). Consider a 2x2 table in which research findings are compared against the gold standard of true relationships in a scientific field. In a research field both true and false hypotheses can be made about the presence of relationships. Let  $R$  be the ratio of the number of "true relationships" to "no relationships" among those tested in the field.  $R$  is characteristic of the field and can vary a lot depending on whether the field targets highly likely relationships or searches for only one or a few true relationships among thousands and millions of hypotheses that may be postulated. Let us also consider, for computational simplicity, circumscribed fields where either there is only one true relationship (among many that can be hypothesized) or the power is similar to find any of the several existing true relationships. The pre-study probability of a relationship being true is  $R/(R+1)$ . The probability of a study finding a true relationship reflects the power  $1-\beta$  (one minus the Type II error rate). The probability of claiming a relationship when one truly exists reflects the Type I error rate,  $\alpha$ . Assuming that relationships are being tested in the field, the expected values of the 2x2 table are given in Table 1. After a research finding has been claimed based on achieving formal statistical significance, the post-study probability that it is true is the positive predictive value, PPV. The PPV is also the complementary probability of what Wacholder et al. have called the false positive report probability (10). According to the 2x2 table, one gets  $PPV = (1-\beta)R / (R-\beta R + \alpha)$ . A research finding is thus

**Open Access** Ioannidis JP (2005) Why most published research findings are false. *PLoS Med* 2(8):e124

**Copyright** © 2005 John P. A. Ioannidis. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

**Abbreviation** PPV positive predictive value

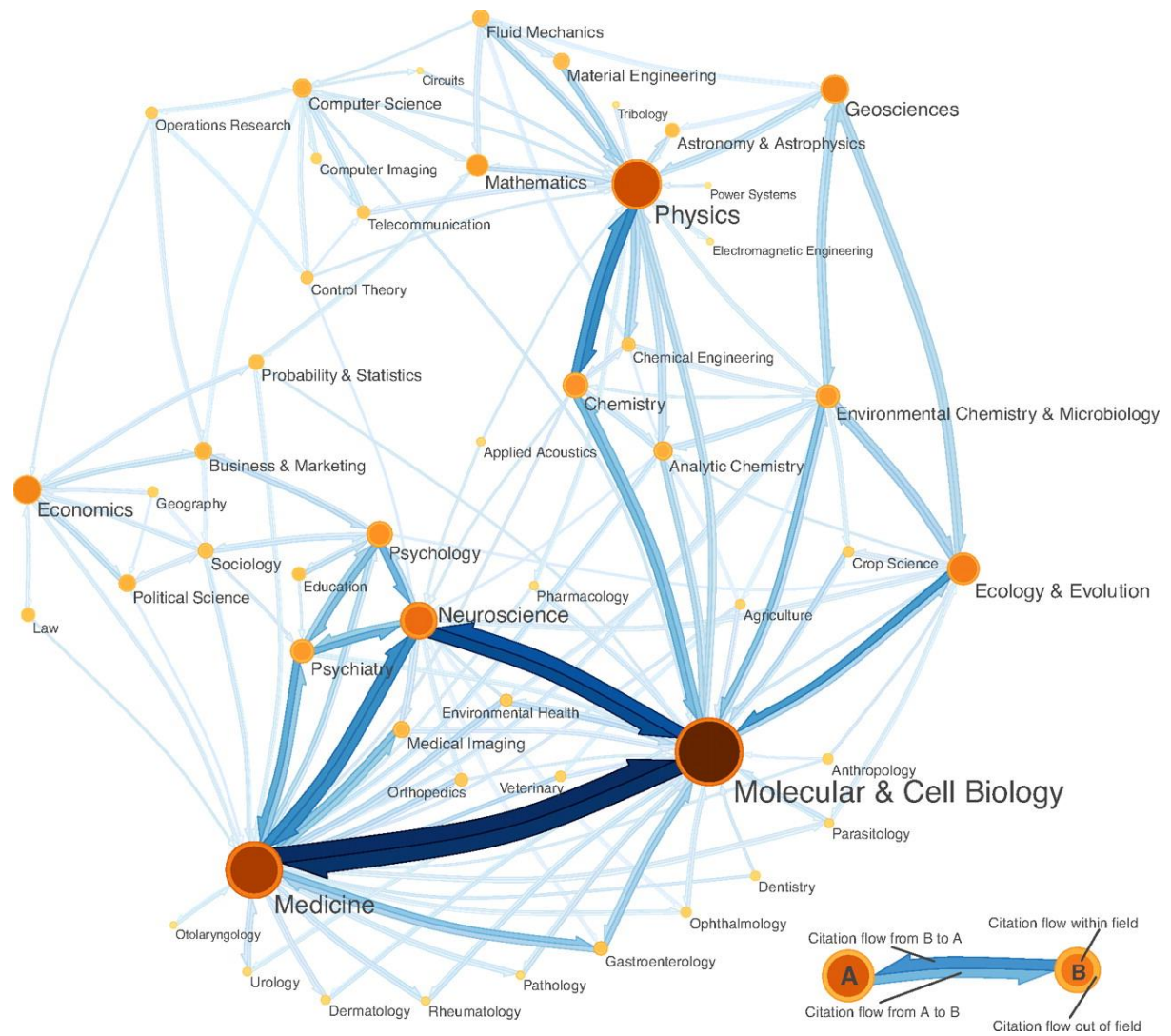
**John P. A. Ioannidis** is in the Department of Organismal and Evolutionary Biology, University of Colorado, School of Medicine, Aurora, Colorado and the Institute for Complex Systems, Tufts-New England Medical Center, Tufts University School of Medicine, Boston, Massachusetts, United States of America. E-mail: ioannidis@hsph.harvard.edu

**Competing interests:** The author has declared that no competing interests exist.

**DOI:** 10.1371/journal.pmed.0020124

August 2005 | Volume 2 | Issue 8 | e124

PLoS Medicine | www.plosmedicine.org 966



Rosvall, M., & Bergstrom, C. T. (2008). Maps of random walks on complex networks reveal community structure. *Proceedings of the National Academy of Sciences*, 105(4), 1118-1123.

# A major challenge for certain empirical knowledge production models in the social sciences and beyond: QRPs related to the use of significance tests and theory testing more generally

Open access, freely available online

## Why Most Published Research Findings Are False

John P. A. Ioannidis

**Summary**

There is increasing concern that most current published research findings are false. This probability that a research claim is true may depend on study power and bias, the number of other studies on the same question, and, to a lesser extent, the ratio of true to no relationships among the relationships posited in each scientific field. In this framework, research finding is less likely to be true when the studies conducted in a field are smaller, when effects have an unclear relation to a specific number and lower production of typical relationships, when there is greater freedom to conduct different outcomes, and analytical models, when there is greater freedom and other errors and publication bias, and when more tests are conducted in a scientific field. In those of statistical significance. Simulations show that for most study designs and settings, it is more likely for research claim to be false than true. Moreover, for many current scientific fields, claimed research findings may often be very accurate estimates of the prevailing bias in the study (bias), the implications of these problems for the research and interpretation of research.

**It can be proven that most claimed research findings are false.**

Published research findings are sometimes refuted by subsequent evidence, with ensuing confusion and disappointments. Refutation and controversy is seen across the range of research designs, from clinical trials and traditional epidemiological studies (1-3) to the most modern molecular research (4,5). There is increasing concern that in modern research, false findings may be the majority or even the vast majority of published research claims (6-8). However, this should not be surprising. It can be proven that most claimed research findings are false. Here I will examine the key factors that influence this problem and some consider them.

**Modeling the Framework for False Positive Findings**

Several methodologists have pointed out (9-11) that the high rate of nonreplication (lack of confirmation) of research discoveries is a consequence of the conventional, well-founded strategy of claiming conclusive research findings solely on the basis of a single study assessed by formal statistical significance, typically for a p-value less than 0.05. Research is not more appropriately represented and summarized by p-values, but, unfortunately, there is a widespread notion that medical research articles should be interpreted based only on p-values. Research findings are defined here as any relationship resembling formal statistical significance, e.g., effective interventions, informative predictors, risk factors, or associations. "Negative" research is also very useful. However, here we will target relationships that investigators claim exist, rather than null findings.

As has been shown previously, the probability that a research finding is indeed true depends on the prior probability of it being true (before doing the study), the statistical power of the study, and the level of statistical significance (10,11). Consider a 2x2 table in which research findings are compared against the gold standard of true relationships in a scientific field. In a research field both true and false hypotheses can be made about the presence of relationships. Let  $R$  be the ratio of the number of "true relationships" to "no relationships" among those tested in the field. It is characteristic of the field and can vary a lot depending on whether the field targets highly likely relationships or searches for only one or a few true relationships among thousands and millions of hypotheses that may be postulated. Let us also consider, for computational simplicity, circumscribed fields where either there is only one true relationship (among many that can be hypothesized) or the power is similar to find any of the several existing true relationships. The pre-study probability of a relationship being true is  $R/(R+1)$ . The probability of a study finding a true relationship reflects the power  $1-\beta$  (one minus the Type II error rate). The probability of claiming a relationship when one truly exists reflects the Type I error rate,  $\alpha$ . Assuming that relationships are being tested in the field, the expected values of the 2x2 table are given in Table 1. After a research finding has been claimed based on achieving formal statistical significance, the post-study probability that it is true is the positive predictive value, PPV. The PPV is also the complementary probability of what Washburner et al. have called the false positive report probability (10). According to the 2x2 table, one gets  $PPV = (1-\beta)R / (1-\beta)R + \alpha$ . A research finding is thus

**Copyright** © 2005 John P. A. Ioannidis. This is an Open Access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

**Abbreviation** PPV, positive predictive value

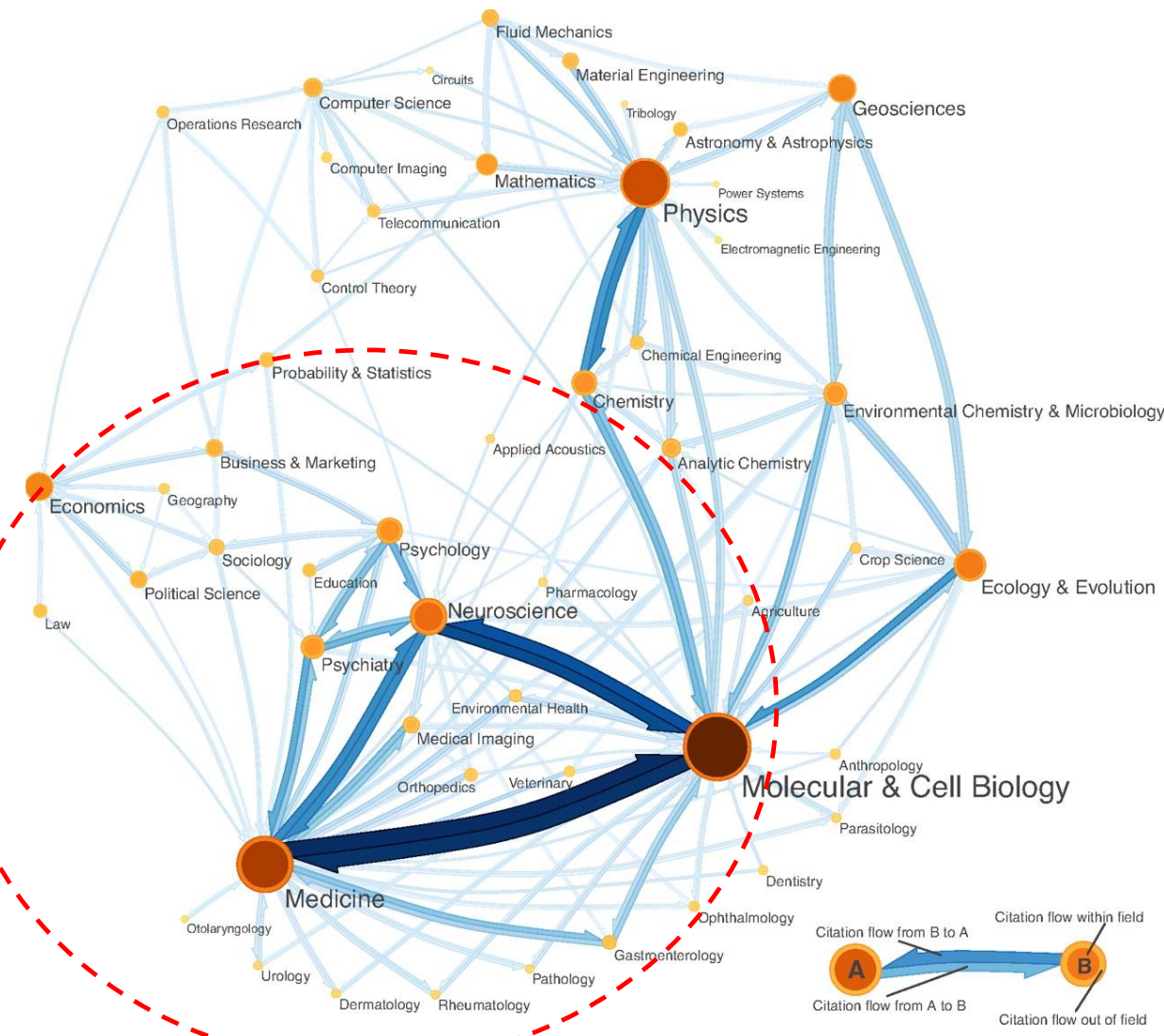
**John P. A. Ioannidis** is in the Department of Organizational Epidemiology, University of Ioannina, School of Medicine, Ioannina, Greece, and the Institute for Clinical Evidence and Health Services Research, University of Medicine, Tufts-New England Medical Center, Tufts University School of Medicine, Boston, Massachusetts, United States of America. E-mail: ioannidis@hsph.harvard.edu

**Competing interests** The author has declared no competing interests exist.

DOI: 10.1186/journal.pmed.0020024

August 2005 | Volume 2 | Issue 3 | e124

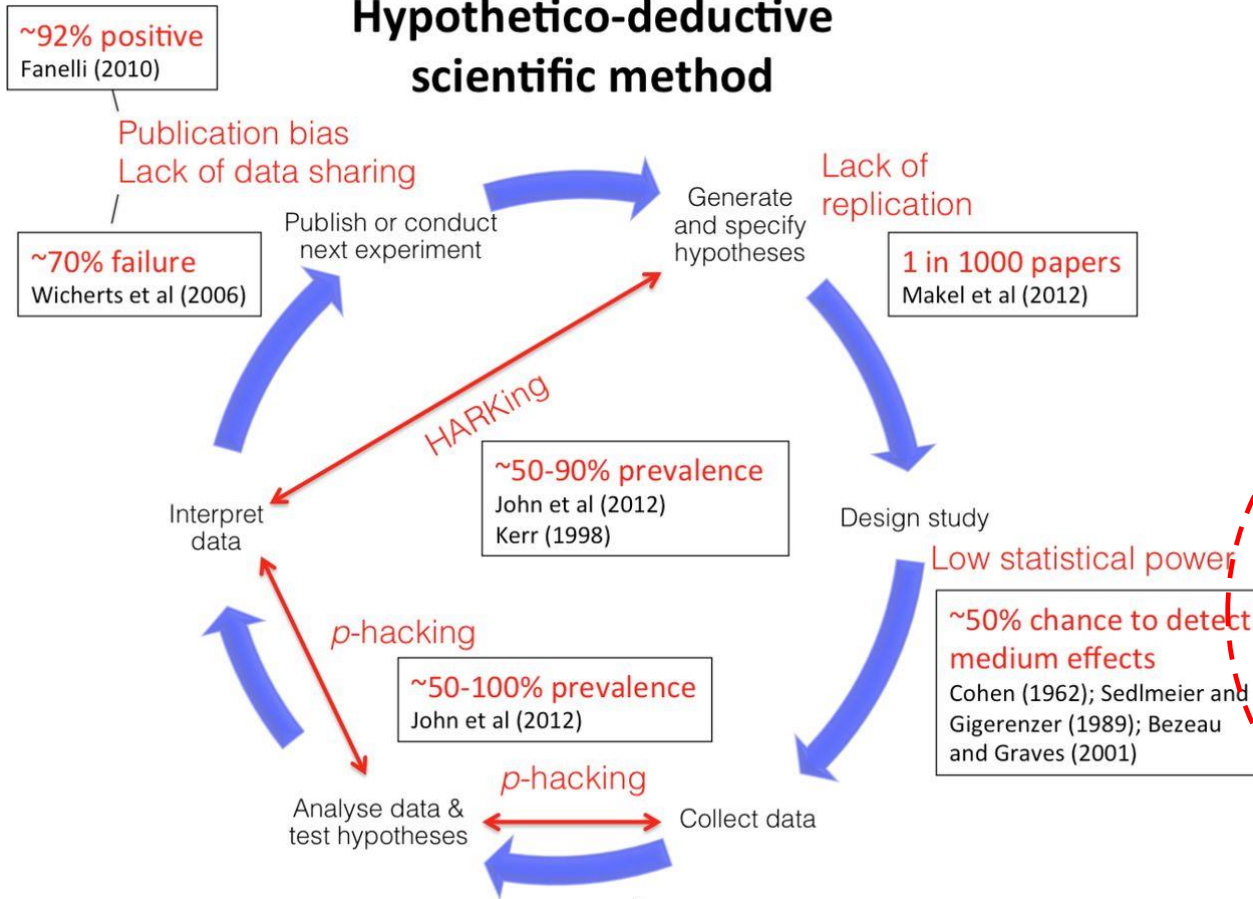
PLoS Medicine | www.plosmedicine.org 966



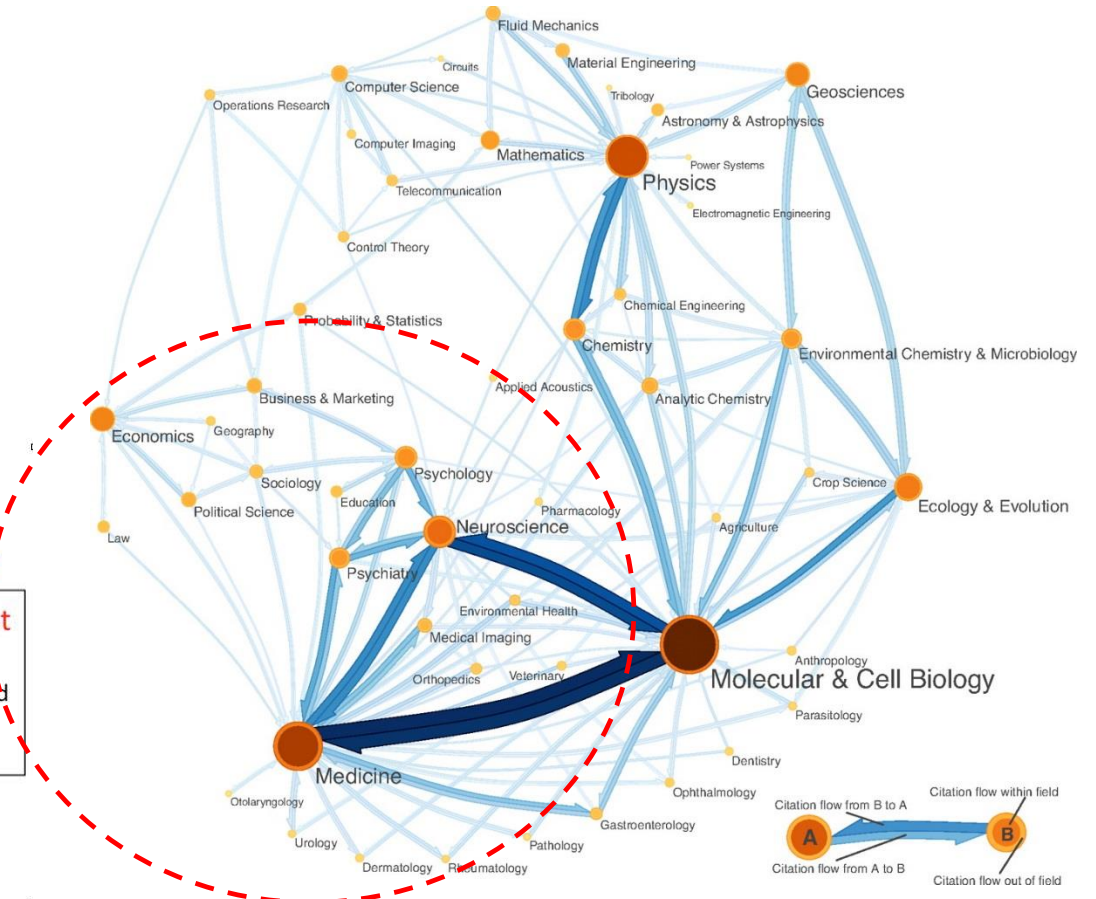
“soft areas”

# Weak-theory testing

## Hypothetico-deductive scientific method



<https://osf.io/8mpji/files/>



Rosvall, M., & Bergstrom, C. T. (2008). Maps of random walks on complex networks reveal community structure. *Proceedings of the National Academy of Sciences*, 105(4), 1118-1123.

# Research integrity in the social sciences

- Broader issues, touching on ethics, seem to have more general consensus
- Starting with authorship and collaboration, things get more complicated
- Coming to methodological practices things get more than complicated, sometimes opposite

## **I. Principles of Research Integrity**

This chapter presents three central principles that should pervade all phases of research:

1. Honesty
2. Transparency
3. Accountability

## **II. Responsible Conduct of Research**

This chapter presents six basic standards for conducting responsible research – from the planning phase to the dissemination of results:

1. Research planning and conduct
2. Data management
3. Publication and communication
4. Authorship
5. Collaborative research
6. Conflicts of interest

Thank you for your attention