Social networks as a potential source of bias in peer review

Sándor Soós
Zsófia Viktória Vida

Beatriz Barros
Ricardo Conejo
Richard Walker

The organizational background

- Hungarian Academy of Sciences (HAS)
  - Linbrary and Information Centre, Dept. Science Policy and Scientometrics

- Professional research evaluation and research monitoring for HAS and national research policy
- Science policy studies (Academic Career Research Programmes)
- Impact assessment of policy measures (institutional level)
- „Translational scientometrics"
- Basic research in scientometrics

Scientific background and activity
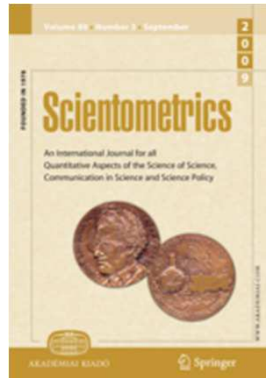
- TTO: institutional successor of the „Budapest School of Scientometrics" (Tibor Braun, Glänzel Wolfgang, András Schubert)
  - Founder and editorial group of **Scientometrics** (Springer-Akadémiai)

- **European Projects (FP7)**
  - Science in Society Observatorium (SISOB)
  - Evaluating the Impact and Outcomes of SSH research (IMPACT-EV)
  - Open Access Policy Alignment Strategies for European Union Research (PA

- **COST Actions**
  - Analyzing the dynamics of information and knowledge landscapes
  - PEERE

Peer review and the SISOB project

- Observatorium for Science in Society:
  - Big data-driven discovery of actual social factors in S&T communities affecting outcomes
  - Related to three specific areas: (1) Mobility (2) Knowledge sharing (3) **Peer review**

- SISOB workbench

Peer review and the SISOB project

- Toolkit (both methodological and technological) to identify:

- Biases (favoritism/discrimination) based on reviewer vs. author characteristics
  - Walker, R., Barros, B., Conejo, R., Neumann, K., & Telefont, M. (2015). Bias in peer review: a case study. *F1000Research*, *4*.
- Biases based on social organization (relations) of the scientific community
  - Social networks as a potential source of bias in peer review (in preprint version)

      DATA provided by publishing house

      Frontiers In

Biases by characteristics

F1000Research
Open for Science

Articles Channels For Authors For Referees About Blog

F1000Research » Articles

RESEARCH ARTICLE

Bias in peer review: a case study [v1; ref status: approved with reservations 2, http://f1000r.es/4yq]

Richard Walker[1], Beatriz Barros[2], Ricardo Conejo[2], Konrad Neumann[3], Martin Telefont[4]

+ Author affiliations

+ Grant information

Abstract

Peer review is the "gold standard" for evaluating journal and conference papers, research proposals, on-going projects and university departments. However, it is widely believed that current systems are expensive, conservative and prone to various forms of bias. One form of bias identified in the literature is "social bias" linked to the personal attributes of authors and reviewers. To quantify the importance of this form of bias in modern peer review, we analyze three datasets providing information on the attributes of authors and reviewers and review outcomes: one from Frontiers - an open access publishing house with a novel interactive review process, and two from Spanish and international computer science conferences, which use traditional peer review. We use a random intercept model in which review outcome is the dependent variable, author and reviewer attributes

1826

Problem with traditional perspective: „the more central the **author** is, the more awarded in peer review" (Scores to papers, not authors)

av. Score =10

av. Score =10

- Solution to the difficulty above: turn it upside down!

- Paper centrality (instead of author ~):

  - for each paper P with authors {A1, ... , An} and author centralities AC = { C(A1), ..., C(An)}, the maximum value of AC was obtained along each measures.

- New question, operationalized: whether reviewer scores for papers reflect the authors include high centrality ones.

- Links and scores made independent, empirically commensurable

The study made in

Introduction – The goal

- SISOB project is developing a methodology to systematically evaluate possible biases in different kinds of peer review system.

- We have developed a toolkit of techniques to detect social network effects on peer review outcomes.

- We would like to detect whether reviewer outcomes are affected
  - by authors' prestige (their "centrality" in their respective communities),
  - by their social relationships with reviewers (the distance between authors and reviewers in coauthoring and in author-reviewer networks) and
  - by their membership of specific subcommunities.

Hypotheses

- Mean reviewer scores for papers by a given author are directly related to the lead author's position (centrality) in these networks.

- Mean reviewer scores for papers by a given author are inversely related to the reviewer's distance from the lead author.

- Reviewers belonging to the same subcommunity as an author will give higher scores than reviewers belonging to different communities.

Data

- **Frontiers database**
  - Frontiers Open Access Publishing House (N=4550)
  - Period: June 25, 2007 – March 19, 2012
  - The data included:
    - the name of the journal,
    - to which the paper was submitted,
    - the article type (review, original research etc.),
    - the name and institutional affiliations of the authors and reviewers of specific papers,
    - individual reviewers scores and the overall review result (accepted/rejected)

- **WebConf database**
  - Six computer science conferences (AH2002, AIED2003, CAEPIA2003, ICALP2002, JITEL2007, SINTICE07,UMPAP2011) (N=1204)
  - Period 2002-2011
  - The data included:
    - name of conference,
    - type of contribution,
    - name, gender and institutional affiliations of the authors and reviewers of specific contributions,
    - individual reviewers scores and the final decision (accepted/rejected)

Modelling author-reviewer relations

- Co-authorship networks

- We constructed a list of all authors in the two databases.

- For each author, we generated a list of other authors with whom the author had previously published at least one paper referenced in the Scopus database.

- On this basis, we identified co-authorship relationships present in the two databases.

- Two unweighted (undirected) co-authoring graphs.

- Frontiers: # nodes = 15 842; WebConf: #nodes = 2149

- Both graphs included a giant "connected" component (Frontiers: N = 8 690; WebConf: N=543;) and disconnected "islands".

- Hereafter we use the connected component.

Indicators

- To test our hypotheses:
  - we have used:
    - author centralities
    - author – reviewer distances

  - We have defined:
    - paper centralities
    - paper distances
    - sub communities

Author centrality

- For each author, we computed the following centrality measures showing the author's position in the coauthoring network:

  - Degree centrality,
  - Betweenness centrality,
  - Closeness centrality,
  - Eigen centrality
  - Page Rank centrality.

Author centrality

- For each author, we computed the following centrality measures showing the author's position in the coauthoring network:

  - Degree centrality,
  - Betweenness centrality,
  - Closeness centrality,
  - Eigen centrality
  - Page Rank centrality.

- H1: Mean reviewer scores for papers by a given author are directly related to the lead author's position (centrality) in these networks

- C_{paper} = max (AC_i)

Where

- C_{paper} is paper centrality
- Ac_i is the value of a particular centrality measure for the ith author of the paper

- Rank correlation between different measures of paper centrality and mean review scores

	bw	d	c	e	p	Score
bw	1	0.83	0.68	0.55	0.40	-0,01
d		1	0.64	0.53	0.41	0.01
c			1	0.87	0.20	-0,00
e				1	0.16	-0.01
p					1	-0.01

Frontiers

	bw	d	c	e.	p.	Score
bw	1	0.83***	0.72***	0.69	0.34	0.10
d		1	0.74***	0.71	0.39.	0.22*
c			1	0.87	0.58.	0.22*
e				1	0.43	0.12
p					1	0,18*

Webconf

bw: betweenness centrality
d: degree centrality;
c: closeness centrality;
e: eigen centrality;
p: Page rank centrality
Signif. codes: 0 '***' 0.001 '**' 0.01 '*'

Results - Paper centrality affects reviewer scores

- Rank correlation between different measures of paper centrality and mean review scores

Frontiers

	bw	d	c	e	p	Score
bw	1	0.83	0.68	0.55	0.40	-0,01
d		1	0.64	0.53	0.41	0.01
c			1	0.87	0.20	-0,00
e				1	0.16	-0.01
p					1	-0.01

bw: betweenness centrality
d: degree centrality;
c: closeness centrality;
e: eigen centrality;
p: Page rank centrality
Signif. codes: 0 '***' 0.001 '**' 0.01 '*'

Results - Paper centrality affects reviewer scores

- Rank correlation between different measures of paper centrality and mean review scores

Webconf

	bw	d	c	e.	p.	Score
bw	1	0.83***	0.72***	0.69	0.34	0.10
d		1	0.74***	0.71	0.39.	0.22*
c			1	0.87	0.58.	0.22*
e				1	0.43	0.12
p					1	0,18*

bw: betweenness centrality
d: degree centrality;
c: closeness centrality;
e: eigen centrality;
p: Page rank centrality
Signif. codes: 0 '***' 0.001 '**' 0.01 '*'

Results – Communities

- We conjectured that large-scale patterns might be suppressing possible network biases.

- We identified subcommunities

- We repeated the experiment described in earlier for each individual subcommunity.

- The results were qualitatively similar.

  - The Frontiers data showed no sign of correlation.

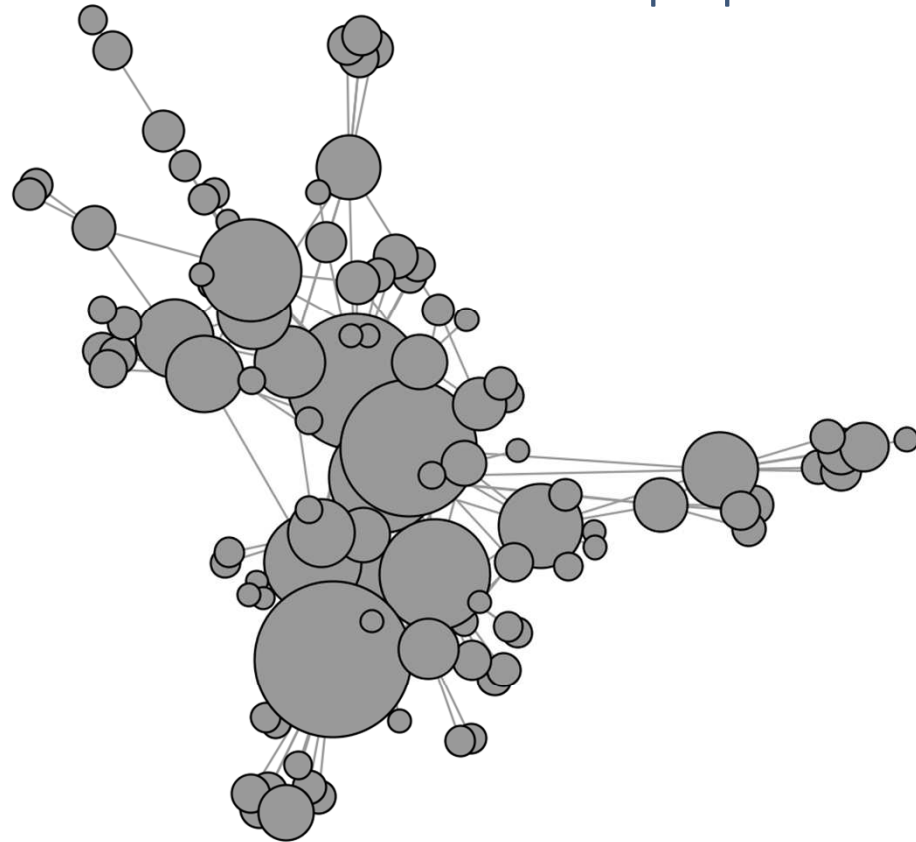 - The WebConf data showed signs of a small but significant correlation.

Layout of a subcommunity in the Frontiers co-authoring network

size of node is proportional to page-rank centrality

Largest community detected by the "fast greedy" community detection algorithm when applied to the WebConf database

size of node is proportional to page-rank centrality

Results – distance vs. score

- Author-reviewer distance affects review scores (H2)
- We interpreted "distance" as the length of the shortest path connecting between authors and reviewers.
- We defined paper distance as the smallest distance between the reviewer and the authors of the paper int he network.
- Compared the paper distances against the score given by the reviewer to the paper.
- The comparison showed no correlation between the two variables either
  - for Frontiers (rank correlation = -0,05) or
  - for WebConf (rank correlation= -0.07)

Results – distance vs. score in subcommunities

- We submitted the giant component of the two networks
- We interpreted reviewer–author proximity as a binary variable
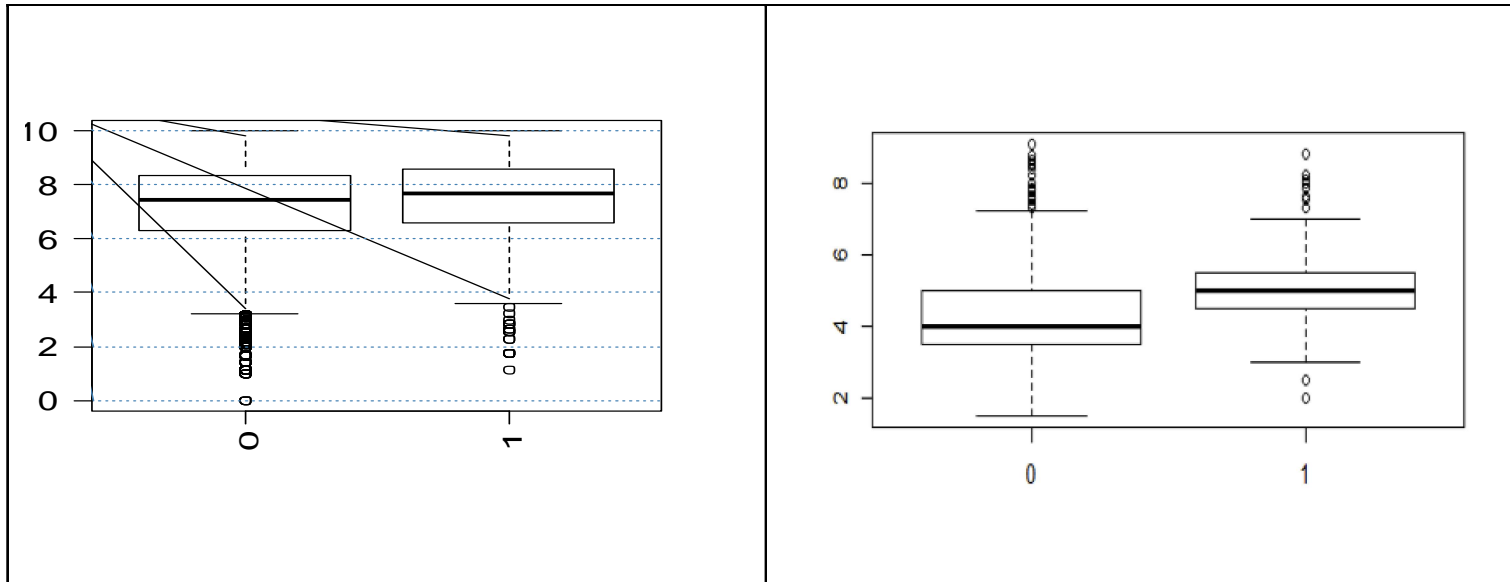
Frontiers   **Webconf**

Mean scores review scores when authors and reviewers are
in different subcommunities (column 0)
or in the same community (column 1).

Summary

- There were no detectable correlations between author centrality and review scores in the Frontiers data

- The WebConf data showed small but significant correlations.

- Neither dataset showed any significant relationship between author-reviewer distance and review scores. This suggests an absence of favoritism.

Summary

- In both systems, reviewers belonging to the same community as authors gave (slightly) higher scores than reviewers coming from different subcommunities - an effect that was stronger for WebConf than for Frontiers.

  - This effect does seem to signal some kind of bias – perhaps because reviewers are most familiar with the language, style and concepts of their own sub-communities.

  - The effect detected was almost certainly too small to affect the outcome of the review process.

- It is possible, however, that biases in other peer review systems are stronger than those registered for Frontiers and WebConf. In such cases, the methodology presented here has the power to detect the bias.

Social networks as a potential source of bias in peer review

- *Thank you for your attention!*