Epistemic Diversity and Editor Decisions: A Statistical Matthew Effect

Remco Heesen[1] Jan-Willem Romeijn[2]

[1]Faculty of Philosophy
University of Cambridge
remcoheesen.eu

[2]Faculty of Philosophy
University of Groningen
www.philos.rug.nl/~romeyn/

PEERE International Conference on Peer Review
7 March 2018

The Importance of Epistemic Diversity

> The history of science has been and should be a history of competing research programmes (or, if you wish, 'paradigms'), but it has not been and must not become a succession of periods of normal science: the sooner competition starts, the better for progress. (Lakatos 1978, p. 69)

The Importance of Epistemic Diversity

> The history of science has been and should be a history of competing research programmes (or, if you wish, 'paradigms'), but it has not been and must not become a succession of periods of normal science: the sooner competition starts, the better for progress. (Lakatos 1978, p. 69)

- Journals (editors/peer reviewers) should promote epistemic diversity
- Bias in favor of monoculture is detrimental to progress

Editorial Biases

- Editors' cognitive biases may favor established research program
  - Confirmation bias
  - Anchoring

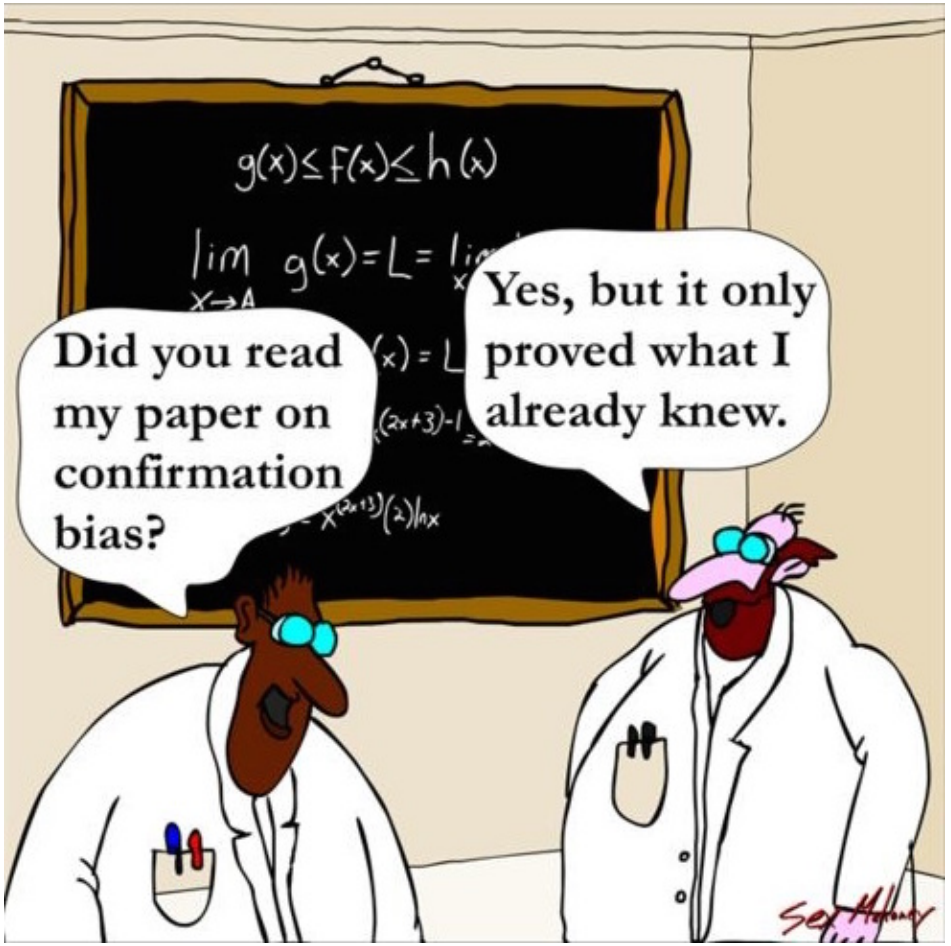


Image source: `http://sexmahoney.blogspot.co.uk`

A Statistical Matthew Effect

Our claim:

- Suppose editor selects only for quality
- "Strictly statistical" biases in peer review
- Favor established research programs

A Statistical Matthew Effect

Our claim:

- Suppose editor selects only for quality
- "Strictly statistical" biases in peer review
- Favor established research programs
- We call this a statistical Matthew effect (Merton 1968)

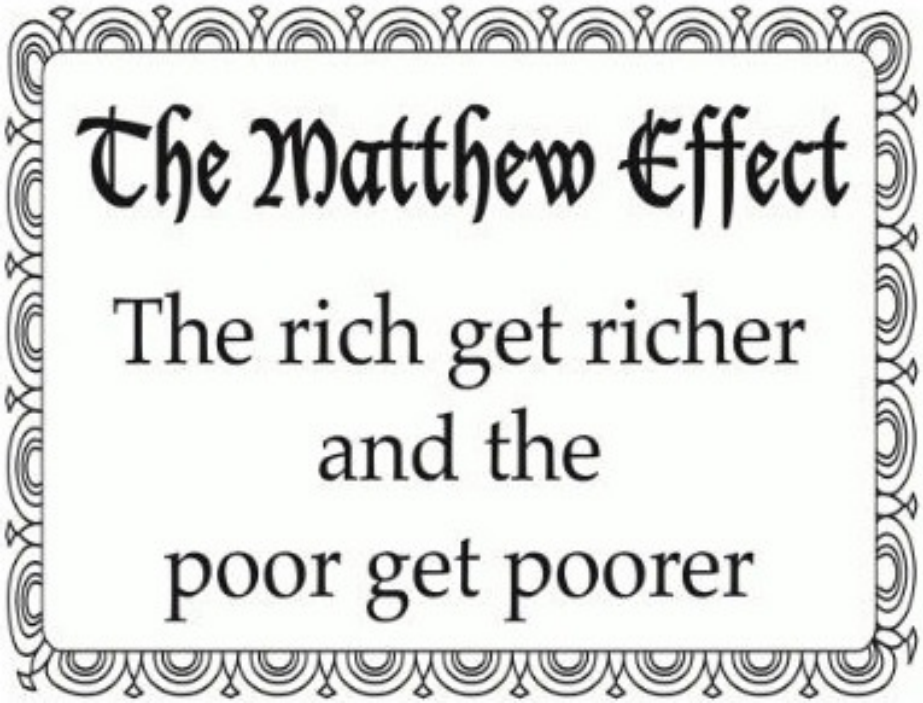


Image source: `http://theliteracywiki.wikispaces.com`

Outline

Outline

Quality and Information

Assumptions of the model: ▸ Formal details

- Each paper has latent quality q
- Less uncertainty for known authors

Image source: `www.blachford.com`

Quality and Information

Assumptions of the model: ▸ Formal details

- Each paper has latent quality q
- Less uncertainty for known authors
- Distribution of quality the same for two research programs
- But: authors from established research program more likely to be known

Image source: `www.blachford.com`

Quality and Information

Assumptions of the model: ▸ Formal details

- Each paper has latent quality q
- Less uncertainty for known authors
- Distribution of quality the same for two research programs
- But: authors from established research program more likely to be known
- Reviewer(s) estimate quality
- Editor accepts papers of high (posterior) quality

Image source: `www.blachford.com`

Bias Favors the Established Research Program

Result

Higher acceptance rate or higher average quality for papers from established research program

Bias Favors the Established Research Program

Result

Higher acceptance rate or higher average quality for papers from established research program

Dilemma for the editor: despite equal quality distributions

- Either established program receives more exposure
- Or published work from established program is seen to be better

Discussion

- Due to information asymmetry, editor treats programs differently
- Justified?
  - Maximum use of information given goal of selecting for quality

Discussion

- Due to information asymmetry, editor treats programs differently
- Justified?
  - Maximum use of information given goal of selecting for quality
- But: epistemic diversity suffers
- How to prevent this?

Discussion

- Due to information asymmetry, editor treats programs differently
- Justified?
  - Maximum use of information given goal of selecting for quality
- But: epistemic diversity suffers
- How to prevent this?
- Suggestion: role of editor's prior is unjustified

Outline

Latent Quality Differences

- In this model, problems arise from latent quality differences
- Plausibly, established research program produces higher quality on average
- Novel program may have startup problems

A Different Model of Peer Review

Assumptions of the model: ▸ Formal details

- Each paper has latent quality q
- Reviewer(s) estimate quality

A Different Model of Peer Review

Assumptions of the model: ▸ Formal details

- Each paper has latent quality q
- Reviewer(s) estimate quality
- Editor accepts based only on reviewer estimate
- Goal is to accept suitable (high quality) papers

A Different Model of Peer Review

Assumptions of the model: ▸ Formal details

- Each paper has latent quality q
- Reviewer(s) estimate quality
- Editor accepts based only on reviewer estimate
- Goal is to accept suitable (high quality) papers
- Quality follows "log-concave" probability distribution
- Average quality in established program higher than in novel program

Peer Review Favors the Established Research Program

Result

Peer review works better for established program: greater proportion of accepted papers is suitable, and suitable papers are accepted at a higher rate

Corollary

Higher acceptance rate and higher average quality of published papers for established program

Peer Review Favors the Established Research Program

Result

Peer review works better for established program: greater proportion of accepted papers is suitable, and suitable papers are accepted at a higher rate

Corollary

Higher acceptance rate and higher average quality of published papers for established program

- Despite "unbiased" peer review, established program better off

Outline

Purely Statistical Biases Versus Other Biases

- Safeguarding epistemic diversity is difficult
- Efforts to curtail cognitive biases must continue, but. . .

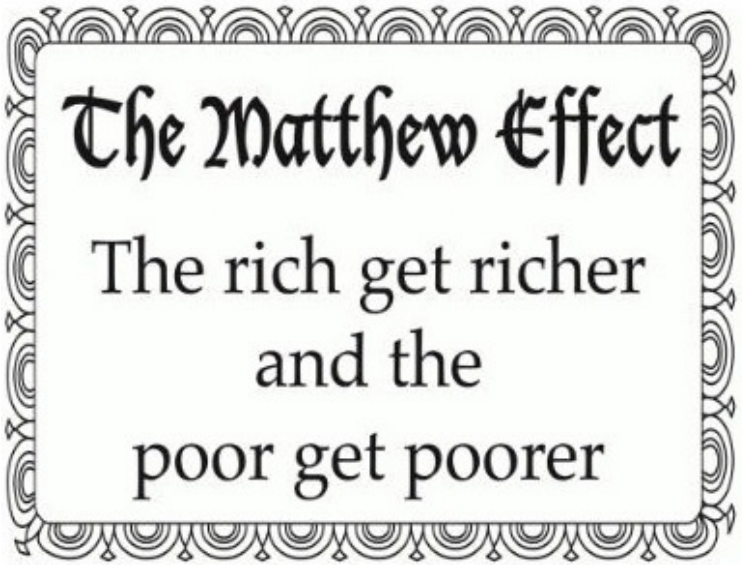


Image source: `http://theliteracywiki.wikispaces.com`

Purely Statistical Biases Versus Other Biases

- Safeguarding epistemic diversity is difficult
- Efforts to curtail cognitive biases must continue, but. . .
- Peer review may favor established research programs even in their absence
- What can be done about this?

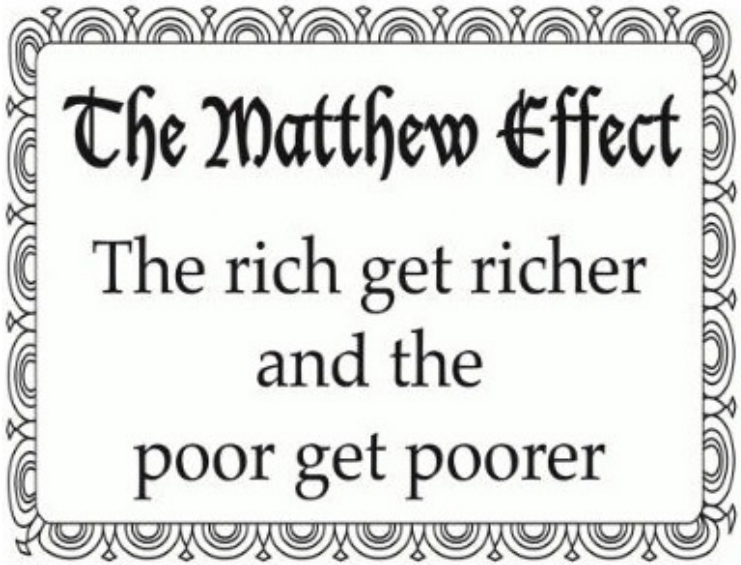


Image source: `http://theliteracywiki.wikispaces.com`

Differential Treatment

- Proposal: solicit extra reviews for close calls

Differential Treatment

- Proposal: solicit extra reviews for close calls
- Additional reviews required more often for novel research program
- Safeguarding epistemic diversity requires differential treatment

Multiple Dimensions of Evaluation

- Objection: notion of quality is too idealized
- Could multidimensional evaluation avoid bias?

Multiple Dimensions of Evaluation

- Objection: notion of quality is too idealized
- Could multidimensional evaluation avoid bias?
- Reply: selection involves implicit unidimensional scale
- Does not avoid bias

Abolish Peer Review

- Proposal: abolish peer review altogether
- ArXiv model of publishing

Thank You!

Thank you for your attention!

Questions?

References I

Denny Borsboom, Jan-Willem Romeijn, and Jelte M. Wicherts. Measurement invariance versus selection invariance: Is fair selection possible? *Psychological Methods*, 13(2):75–98, Jun 2008. doi: 10.1037/1082-989X.13.2.75. URL http://dx.doi.org/10.1037/1082-989X.13.2.75.

Philip Kitcher. The division of cognitive labor. *The Journal of Philosophy*, 87(1):5–22, 1990. ISSN 0022362X. URL http://www.jstor.org/stable/2026796.

Imre Lakatos. *The Methodology of Scientific Research Programmes*. Cambridge University Press, Cambridge, 1978.

Robert K. Merton. The Matthew effect in science. *Science*, 159(3810): 56–63, 1968. ISSN 00368075. URL http://www.jstor.org/stable/1723414.

Quality and Uncertainty

- Submitted paper has latent quality q
- Identity of author is relevant to quality
  - Editor's prior for known author: $\pi(q \mid K)$
  - Editor's prior for unknown author: $\pi(q)$
- Distribution of quality is the same for research programs H and L
  - Research program of author is irrelevant to quality:
  - $\pi(q \mid K, H) = \pi(q \mid K, L)$ and $\pi(q \mid H) = \pi(q \mid L)$
- But authors from program H more likely to be known
  - Editor may belong to program H

Peer Review

- Editor solicits reviews
- Reviewer report R independent of research program and identity of author (given q)
- Editor updates beliefs about q
  - Posterior for known author: $\pi(q \mid K, R)$
  - Posterior for unknown author: $\pi(q \mid R)$

Acceptance and Utility

- Editor must accept (A) or reject ($\neg A$) submission
- Editor selects only for quality
  - Utility of acceptance equals quality q
  - Utility of rejection is some fixed value q^*

$\Longrightarrow$ Editor accepts if and only if posterior mean exceeds q^*

A Different Model of Peer Review

- Quality q follows the same log-concave distribution in both programs
  - $f(tq + (1-t)q') \geq f(q)^t f(q')^{1-t}$
  - E.g., normal, uniform, exponential, gamma
- Reviewer report R unbiased: independent of research program (given q)
- Editor must accept (A) or reject ($\neg A$) submission
- "Frequentist" editor: accept if and only if reviewer report exceeds q^*
  - Identical decision procedures:
  - $D_H = A$ if $R > q^*$ and $D_L = A$ if $R > q^*$
- No distributional assumption on R except: conditional probability of acceptance increasing in q
  - $\Pr(R > q^* \mid q)$ increasing in q

The Result

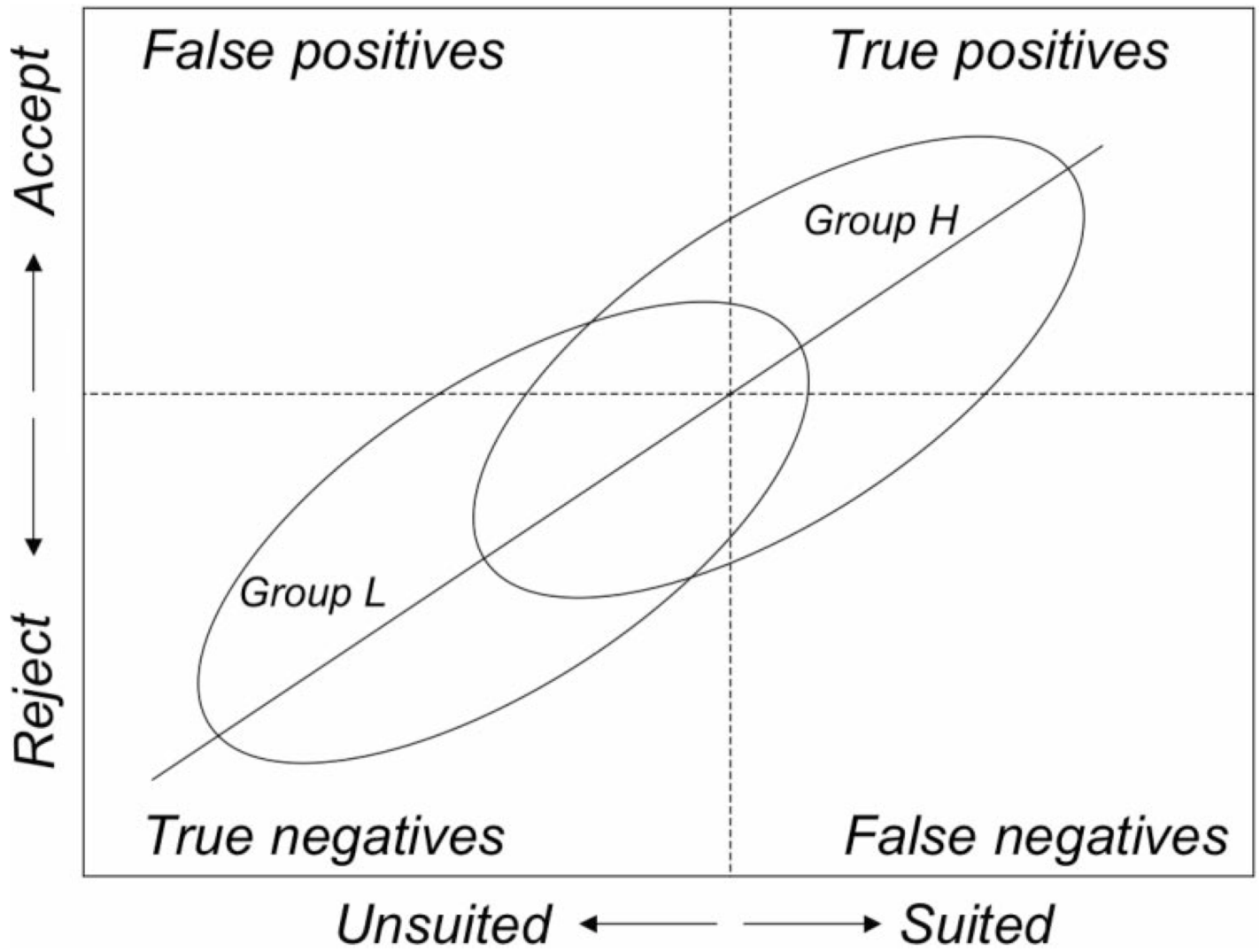
A submission is suitable (S) if its quality q exceeds threshold t

Theorem 1

A greater proportion of accepted papers from program H is suitable: $\Pr(S \mid D_H = A) > \Pr(S \mid D_L = A)$. *Conversely, suitable papers from program H are accepted at a higher rate:* $\Pr(D_H = A \mid S) \geq \Pr(D_L = A \mid S)$, *with strict inequality unless the distribution of quality is exponential.*

- Generalizes Borsboom et al. (2008)
- Generalization also considers different variances

Why Does This Happen?



Addressing the Problem